Logic-based Conflict Analysis
and Resolution

David Frieder Georg Lempp

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy
in Logic

2009
To Frida,
Liam,
and Eli
The aim of this thesis is to explore the extent to which formal logic can be applied to the topic of conflict analysis and conflict resolution. It is motivated by the idea that conflicts can be understood as inconsistent sets of goals, beliefs, norms, emotions, or the like. To achieve this aim, two formal frameworks are presented. Conflict Modelling Logic (CML) is a logical system, based on branching-time temporal logic, which can be used to describe and interpret conflicts. Conflict Resolution Logic (CRL) is a set of five algorithms, inspired by the AGM model of belief revision, which can be used to generate possible solutions to conflicts. Furthermore, two numerical measures for the ‘potential conflict power’ of propositional formulae and the ‘degree of inconsistency’ of sets of propositional formulae are introduced. The two measures allow one to assess the role of particular elements within a conflict and the depth of a conflict. The formal framework is illustrated with the example conflict of the Second Congo War.
ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor Ed Mares for his constant support and guidance from the first day of my thesis on. Ed, I appreciate how you always gave me the advice I needed but also the freedom to explore my own ideas as much as I wanted.

During the last months, I enjoyed the company of another logic student at the department, David Gilbert, who, although knowing me only for a short time, proofread my entire thesis. Thank you so much for that, David! Without your help, my thesis would be full of missing commas and weird grammatical constructions.

For a while, my regular midnight chess and philosophy meeting with Tony Scott was the highlight of my day (or night?). With you, Tony, I not only played probably more than 3000 chess games, but also solved most of the big problems in philosophy and had a lot of fun. Thanks!

The conversations I had with Dan Turton, Alice Monro, and Matt Gers about all sorts of philosophical and non-philosophical topics always inspired me. Thank you for all the things I learned from you and for the great time I had with you at the philosophy retreat in Otaki and the AAP in Auckland.

My girlfriend, Kerstin Maurus, gave me a lot of love and prevented me from turning into a freak. Kerstin, there are not enough words to thank you for your support and for making my life colourful at times when I saw everything around me in black and white.

I want to thank all the staff members of the Philosophy Programme which became an academic home to me during the last three years. In particular, thanks to Cei Maslen, David Eng, Sondra Bacharach, Ramon Das, Kim Sterelny, and Nick Agar for useful comments on my project and your social support.

Thanks to the members of the Centre for Logic, Language and Computation, in particular, to Rob Goldblatt, Neil Leslie, and Villemijn Vermaat for your support and for giving me the opportunity to present my ideas in the CLLC seminar series.

There are many more people who helped and supported me significantly during the last three years. I cannot mention how each of you helped me specifically, but I want to thank you all at this place: Tim McVicar, Anna Herbert, Esther Woodbury, Jenny and Tom Wiggley, Lucas Kengmana, Arjun Harindranath, Vivian Rodriguez, Nick Parkin, Pascal Reiling, Alan and Dennis Poole, Chris Rodgers, Vanessa Scouten, Ralf Rombach, Johanna Seibt, Martin Gray, Michael Kane, Esther Werth, Dieter, Ilse and Britta Maurus, Kathrin Knabe, Felix Schaller, Lucy Weston-Taylor, Joana Baber, Kate Tappenden, Susanne Michel, David Stadtmüller, Joel Woodbury, Vanessa Scholes, Martin Lempp, Margarethe Lempp, Linda Weijers, Jakob Lempp, Maria Lempp, Audrey Hasson, Gerhard Lempp, Robin Aldridge-Sutton, Felix Lempp, Sadaf Momeny, Atsushi Iseki, Steven Riley, Adrian Currie, Kim Shaw-Williams, and Rhydian Thomas.

Thanks also to the New Zealand tax payer who made it possible for me to receive a Victoria Postgraduate Scholarship over three years.
## CONTENTS

**Introduction** ................................................................................................................. 17

1 Reflecting on Conflicts:  
Theories of Conflict ............................................................................................................. 21

1.1 Introduction .................................................................................................................. 21
1.2 Formal Theories ............................................................................................................ 22
   1.2.1 Decision Theory .................................................................................................. 24
   1.2.2 Game Theory ....................................................................................................... 36
   1.2.3 Robert Kowalski’s Logic-based Approach ......................................................... 48
   1.2.4 Lewis Richardson’s Equational Model ............................................................... 52
1.3 Psychological Theories ............................................................................................... 56
   1.3.1 Theories of Aggression ....................................................................................... 57
   1.3.2 Muzafer Sherif’s Realistic Conflict Theory ....................................................... 62
   1.3.3 The Theory of Social Dilemmas ....................................................................... 64
1.4 Political Theories ......................................................................................................... 67
   1.4.1 Liberalism and Neoliberal Institutionalism ......................................................... 68
   1.4.2 Realism, Neorealism, and Neoclassical Realism .............................................. 74
1.5 Theories about Attitudes and Social Groups .............................................................. 80
   1.5.1 Theories about Attitudes .................................................................................... 81
   1.5.2 Theories about Social Groups .......................................................................... 84
1.6 Summary ...................................................................................................................... 87

2 Describing Conflicts:  
The Syntax of Conflict Modelling Logic ....................................................................... 89

2.1 Introduction .................................................................................................................. 89
2.2 Symbols ........................................................................................................................ 90
   2.2.1 Propositional Constants .................................................................................... 92
   2.2.2 Logical Symbols .............................................................................................. 94
   2.2.3 Modal Symbols ............................................................................................... 95
   2.2.4 Temporal Symbols ......................................................................................... 95
   2.2.5 Doxastic Symbol ............................................................................................. 98
   2.2.6 Symbol for Goals ............................................................................................ 99
   2.2.7 Deontic Symbol .............................................................................................. 99
   2.2.8 Symbol for Emotions ...................................................................................... 101
2.3 Formulae ........................................................................................................................ 102
   2.3.1 Formation Rules ............................................................................................... 102
   2.3.2 Interpretation of Basic Formulae ...................................................................... 102
2.4 Complex Expressions ................................................................................................. 103
   2.4.1 Complex Expressions for Specific Temporal Aspects ....................................... 104
   2.4.2 Complex Expressions for Conscious Attitudes and False Beliefs .................... 105
   2.4.3 Complex Expressions for Emotions .................................................................. 106

Example: Second Congo War ............................................................................................. 106
2.5 Summary ...................................................................................................................... 111
3 Interpreting Conflicts:
The Semantics of Conflict Modelling Logic ........................................... 113

3.1 Introduction ........................................................................................................... 113
3.2 CML-structures .................................................................................................... 124
  3.2.1 Time .................................................................................................................. 126
  3.2.2 Conflict States and Conflict Histories .............................................................. 128
  3.2.3 Agents and Attitudes ....................................................................................... 133
  3.2.4 The Term Assignment Function ....................................................................... 136
3.3 Further Definitions ............................................................................................... 137
  3.3.1 Immediate Predecessor of a Time Point ............................................................ 138
  3.3.2 Conflict States of a Fixed Time Point ............................................................... 138
  3.3.3 The Time Index Function ................................................................................. 140
  3.3.4 The Accessibility Relation ............................................................................... 141
3.4 Truth Conditions .................................................................................................. 143
  3.4.1 Conditions for the Logical Connectives, Identity, and the Quantifier ............... 143
  3.4.2 Conditions for the Modal Operator ................................................................... 144
  3.4.3 Conditions for the Temporal Operator and other Temporal Symbols ............ 146
  3.4.4 Conditions for the Propositional Attitude Operators ....................................... 149
3.5 Validity .................................................................................................................. 150
  3.5.1 Validity at States ............................................................................................... 151
  3.5.2 Validity in CML-Structures .............................................................................. 151
  3.5.3 Validity in CML ................................................................................................ 152
3.6 Satisfiability .......................................................................................................... 152
3.7 The Semantic Consequence Relation ................................................................. 153
3.8 Summary ............................................................................................................... 157

4 Axiomatising Conflicts:
The Axiomatics of Conflict Modelling Logic ................................................. 159

4.1 Introduction .......................................................................................................... 159
4.2 The Modal Fragment ............................................................................................ 160
  4.2.1 Semantic Characterisation of □-CML ............................................................... 162
  4.2.2 Axioms for □-CML ........................................................................................ 171
  4.2.3 Soundness of □-CML ..................................................................................... 172
  4.2.4 Completeness of □-CML ............................................................................... 173
4.3 The Temporal Fragment ....................................................................................... 173
  4.3.1 Semantic Characterisation of t-CML ............................................................... 175
  4.3.2 Axioms for t-CML ........................................................................................ 176
  4.3.3 Soundness of t-CML ..................................................................................... 178
  4.3.4 Completeness of t-CML ............................................................................... 181
4.4 The Propositional Attitude Fragment ................................................................... 198
  4.4.1 Consistent Believers ....................................................................................... 199
4.5 Axioms for CML .................................................................................................. 202
4.6 Summary ............................................................................................................... 205
5 Defining Conflicts:
The Classification System of Conflict Modelling Logic ..........207

5.1 Introduction .................................................................................................................................207
Background: Conflict Definitions ........................................................................................................208
5.2 General Conflict Definition ..........................................................................................................212
5.3 Synchronic versus Diachronic Conflicts .......................................................................................216
5.4 Conflicts on the Individual, Micro, and Macro Level .....................................................................217
5.5 Intra-Agent and Inter-Agent Conflicts .........................................................................................219
5.6 Subjective versus Objective Conflicts ..........................................................................................220
5.7 Conflict Issues .................................................................................................................................221
5.7.1 Factual Disputes .......................................................................................................................222
5.7.2 Conflicts over Goals ..................................................................................................................222
5.7.3 Conflicts over Values .................................................................................................................223
5.8 The Emotional Load of Conflicts ..................................................................................................223
5.9 Combining the Definitions ............................................................................................................224
5.10 Conflicts in Propositional Form ...................................................................................................225
Example: Second Congo War .............................................................................................................226
5.11 Summary .........................................................................................................................................231

6 Measuring Conflicts:
Potential Conflict Power and Degrees of Inconsistency ..........233

6.1 Introduction .....................................................................................................................................233
Background: Belief Revision ...............................................................................................................234
6.2 The Semantic Definition of Potential Conflict Power .....................................................................245
6.3 The Syntactic Definition of Potential Conflict Power .....................................................................248
6.4 Some Properties of Potential Conflict Power ...............................................................................251
6.5 Degrees of Inconsistency ..............................................................................................................257
Example: Second Congo War .............................................................................................................263
6.6 Summary .........................................................................................................................................267

7 Generating Solutions:
The Algorithms of Conflict Resolution Logic .........................269

7.1 Introduction .....................................................................................................................................269
Background: Dialogue Logic ..............................................................................................................270
7.2 General Characterisation of the Resolution Algorithms ................................................................277
7.3 The Algorithm Generating Solution Sets .......................................................................................281
7.4 Algorithms Selecting Value Assignments .....................................................................................283
7.4.1 The Density Minimising Algorithm .........................................................................................284
7.4.2 The Density Minimising, Δ Preserving Algorithm ....................................................................285
7.4.3 The Minimax Algorithm ..........................................................................................................286
7.4.4 The Difference Minimising Algorithm ....................................................................................287
7.4.5 The Conflict Power Balancing Algorithm ...............................................................................288
7.5 Computability and Complexity .....................................................................................................289
Example: Second Congo War .............................................................................................................292
7.6 Summary .........................................................................................................................................297
8 Evaluating Solutions:
The Resolution Principles of Conflict Resolution Logic ..........299

8.1 Introduction .................................................................................................................. 299
8.2 Minimally Invasive Solutions ..................................................................................... 300
  8.2.1 Solutions Minimising Collective Change .......................................................... 300
  8.2.2 Solutions Minimising Individual Change ......................................................... 303
8.3 Legal Solutions .......................................................................................................... 305
8.4 Compromises ............................................................................................................. 307
  8.4.1 Solutions Equalling Individual Change .......................................................... 308
  8.4.2 Solutions Equalling Individual Non-Change .................................................. 310

Example: Second Congo War ......................................................................................... 312
8.5 Summary .................................................................................................................... 315

Conclusion ....................................................................................................................... 317

Bibliography .................................................................................................................... 325
FIGURES

Figure 1: Rooted Tree ................................................................. 38
Figure 2: Extensive Form .......................................................... 40
Figure 3: AND-OR Tree ............................................................ 50
Figure 5: Discrete, Strict Linear Order without Endpoints .......... 127
Figure 6: Conflict Histories .................................................... 132
Figure 7: Propositional Attitude Functions .............................. 135
Figure 8: Partition of the Conflict Universe ............................ 139
Figure 9: Truth Condition of the Necessity Operator ................ 145
Figure 10: Truth Condition of the Temporal Realisation Operator (Future) .... 147
Figure 11: Truth Condition of the Temporal Realisation Operator (Past) .... 148
Figure 12: Transformation of a Finite S5-Structure into a □-Structure .... 167
<table>
<thead>
<tr>
<th>Table 1: Lotteries</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2: Utilities</td>
<td>31</td>
</tr>
<tr>
<td>Table 3: Probability Distributions</td>
<td>32</td>
</tr>
<tr>
<td>Table 4: Expected-Utilities</td>
<td>32</td>
</tr>
<tr>
<td>Table 5: 2-Person Game in Strategic Form</td>
<td>41</td>
</tr>
<tr>
<td>Table 6: Normal Form</td>
<td>43</td>
</tr>
<tr>
<td>Table 7: Domination</td>
<td>44</td>
</tr>
<tr>
<td>Table 8: Nash Equilibrium in Randomised Strategies</td>
<td>47</td>
</tr>
<tr>
<td>Table 9: Nash Equilibrium in Pure and Randomised Strategies</td>
<td>48</td>
</tr>
<tr>
<td>Table 10: Prisoner’s Dilemma</td>
<td>64</td>
</tr>
<tr>
<td>Table 11: Basic Formulae of CML</td>
<td>103</td>
</tr>
<tr>
<td>Table 12: Absolute and Relative Density</td>
<td>260</td>
</tr>
<tr>
<td>Table 13: Truth Table for $C_g$</td>
<td>264</td>
</tr>
<tr>
<td>Table 14: Relative Densities of $C_g$</td>
<td>266</td>
</tr>
<tr>
<td>Table 15: Relative Densities of $C_b$</td>
<td>267</td>
</tr>
<tr>
<td>Table 16: Particle Rules</td>
<td>272</td>
</tr>
<tr>
<td>Table 17: Dialogue for $\neg p \supset p$</td>
<td>275</td>
</tr>
<tr>
<td>Table 18: Dialogue for $\neg p \land q$</td>
<td>276</td>
</tr>
<tr>
<td>Table 19: Dialogue for $\neg(p \land \neg p)$</td>
<td>276</td>
</tr>
<tr>
<td>Table 20: Information for CRL-1</td>
<td>293</td>
</tr>
<tr>
<td>Table 21: Information for CRL-2</td>
<td>294</td>
</tr>
<tr>
<td>Table 22: Information for CRL-3 and CRL-4</td>
<td>295</td>
</tr>
<tr>
<td>Table 23: Information for CRL-5</td>
<td>296</td>
</tr>
<tr>
<td>Table 24: Input and Output of the Five Algorithms</td>
<td>297</td>
</tr>
<tr>
<td>Table 25: Solutions of CRL-1</td>
<td>301</td>
</tr>
<tr>
<td>Table 26: Solutions of CRL-3</td>
<td>304</td>
</tr>
<tr>
<td>Table 27: Solutions of CRL-2</td>
<td>306</td>
</tr>
<tr>
<td>Table 28: Solutions of CRL-4</td>
<td>309</td>
</tr>
<tr>
<td>Table 29: Solutions of CRL-5</td>
<td>311</td>
</tr>
</tbody>
</table>
Introduction

“A conflict exists when two people wish to carry out acts which are mutually inconsistent. [...] A conflict is resolved when some mutually compatible set of actions is worked out.”¹

As the above definition by political scientist Michael Nicholson shows, conflicts can be understood as inconsistent sets of goals, values, or the like. Conflict resolution, then, is the process of eliminating the inconsistencies of these sets.

The aim of this thesis is to explore the extent to which formal logic can be applied to the topic of conflict modelling and conflict resolution. It is based on the assumption that a logical representation of conflicts helps to identify and, finally, eliminate the inconsistencies that constitute them. To achieve this aim, we present two formal, logic-based frameworks: Conflict Modelling Logic (CML), a logical system which can be used to model conflicts, and Conflict Resolution Logic (CRL), a set of five algorithms which can be used to generate possible solutions to conflicts.

¹ (Nicholson 1992, p. 11)
Introduction

CML consists of a syntax, a semantics, and an axiomatics. The syntax combines operators from various logics, such as classical propositional and first-order logic, branching-time temporal logic, and modal logic. It allows one to express the propositional attitudes of the parties involved in a conflict, such as their beliefs, goals, norms, and emotions, in terms of formulae and to track their temporal development. The semantics provides an interpretation for the formulae by specifying their truth conditions in a universal conflict model. According to this model, conflicts are specific types of social processes passing through a forwards-branching, backwards-linear, tree of conflict states. The axiomatics characterises CML in a proof theoretical way and makes it possible to define the notions of consistency and inconsistency. The completeness theorem, in which we prove that the axiom system adequately characterises the semantics of CML, links the proof theoretic notion of consistency with the semantic notion of satisfiability.

After having introduced the syntax, semantics, and axiomatics of CML, we present a general conflict definition based on our formal framework and establish a classification scheme which allows one to identify the specific type of a conflict modelled by CML. We argue that the property distinguishing conflicts from other social processes is the existence of at least one inconsistency between the parties’ propositional attitudes. Depending on whether the inconsistency occurs between beliefs, goals, or norms, conflicts can be classified as factual disputes, goal conflicts, value conflicts, or combinations of these.

Conflicts described by CML can be entered as an input into the five algorithms of CRL which then generate possible solutions. Each solution has the form of a consistent set of formulae and expresses a combination of propositional attitudes that are mutually compatible. Hence, the solutions generated by the algorithms determine those propositional attitudes which need to be changed in order to resolve the conflict. For a
given input of a set of formulae, each of the algorithms systematically replaces formulae in the set until a consistent set is obtained. The process of replacement is guided by the principle of replacing formulae with a high degree of potential conflict power by formulae with a lower degree of potential conflict power. The ‘potential conflict power’ of a formula is a measure of how likely it is to be inconsistent with arbitrary other formulae. Each of the five algorithms follows a different principle of conflict resolution and, hence, produces a different type of solutions. Thus, we can compute two types of minimally invasive solutions, solutions that are compatible with a set of pre-defined legal or moral norms, and two types of compromise solutions.

An outline of the thesis is as follows: In chapter 1, we review current theories of conflict with a focus on formal theories, political theories, and psychological theories. In chapter 2, we introduce the syntax of CML. Chapter 3 deals with the semantics of CML. In chapter 4, we present a sound and complete axiomatisation of CML. In chapter 5, we introduce a general conflict definition and a classification scheme for conflicts based on CML. In chapter 6, we define the potential conflict power of propositional formulae and the degree of inconsistency of sets of propositional formulae, two numerical measures for assessing propositional formulae and sets of propositional formulae with respect to the role they play within a conflict. Chapter 7 and chapter 8 deal with the resolution algorithms of CRL. First, we define the five algorithms of CRL and address their computability and complexity, and then we evaluate the types of solutions they generate.

The chapters are augmented by six example sections, in which we illustrate the introduced concepts with the example conflict of the Second Congo War, and four background sections, in which we provide expositions of theories underlying our own model. The background sections address the topics of branching-time temporal logic, conflict definitions, belief revision, and dialogue logic.
CHAPTER 1

Reflecting on Conflicts

Theories of Conflict

1.1 Introduction

The aim of chapter 1 is to provide an overview of theories of conflict. Theories of conflict have been developed in a large number of diverse academic disciplines and research traditions including the social sciences, the humanities, the formal sciences, and, to a lesser extent, the natural sciences. In our analysis, we focus on formal, psychological, and political theories of conflict because these disciplines cover a significant amount of theoretical conflict research and provide the most relevant contributions to the development of our own logic-based approach.

As a result of this chapter we will be able to give answers to the following questions.

- What are the most important formal, psychological, and political theories of conflict?

---

2 In the natural sciences, conflict theories are focused on evolutionary explanations of aggressive behaviour and co-operation/defection.
Chapter 1  Reflecting on Conflicts

- How are conflicts described and defined by mathematicians, psychologists, and political scientists, respectively?
- What are the main factors determining the dynamics of conflicts according to the three disciplines?

An outline of the chapter is as follows: In section 1.2, we review formal theories of conflict. In particular, we look at decision theory (1.2.1), game theory (1.2.2), Robert Kowalski’s logic-based model of conflict resolution (1.2.3), and Lewis Richardson’s equational approach (1.2.4). In section 1.3, we describe psychological theories of conflict. First, we concentrate on theories of aggression (1.3.1), and then we look at Muzafer Sherif’s realistic conflict theory (1.3.2), and the theory of social dilemmas (1.3.3). In section 1.4, we present two political approaches to conflict theory. We start describing theories developed in the liberalist/institutionalist tradition (1.4.1), and then we describe realist theories of conflict (1.4.2). In section 1.5 we provide general background theories about attitudes (1.5.1) and social groups (1.5.2). The concept of attitudes will be used later in the thesis to capture various conflict elements, such as beliefs, goals, norms, and emotions. Theories about social groups are relevant to our model because conflicts often involve agents above the individual level.

1.2  Formal Theories

Political scientist James Schellenberg nicely characterises formal theories of conflict as follows:

“A formal theory of conflict [...] places primary emphasis on the logical and mathematical nature of its generalizations. While ordinary language may be used for part of the exposition of these theories, the central ideas are expressed in the language of mathematics. In such a framework, social conflict is seen primarily as a manifestation of
Formal conflict theories can be grouped into four categories: decision theory, game theory, computational approaches, and equational approaches. We provide general descriptions of decision theory and game theory and present example theories for the computational approach and the equational approach.

Many concepts of game theory, such as the notion of utility, have their origin in decision theory. As far as conflicts are concerned, decision theory focuses on explaining how individual agents involved in a conflict make their decisions.

Game theory, which constitutes an independent research field in mathematics, can be seen as the contribution of mathematicians to the analysis of conflicts. Indeed, it has been argued that game theory is a general mathematical theory of conflict. Game theory is not only the most predominant formal theory of conflict, but has also influenced conflict theories developed in other disciplines, such as political science, psychology, biology and sociology.

A computational theory of conflict and conflict resolution has been developed by Kowalski in his paper A logic-based approach to conflict resolution. We describe his model as an example of the computational approach to the theory of conflict.

Finally, we introduce the pioneering work of Richardson as an example of the equational approach. More than the other approaches, his attempt to model various aspects of

---

3 (Schellenberg 1996, p. 115)
4 Game theory is classified under the code 91A as its own mathematical research area in the Mathematics Subject Classification of the American Mathematical Society.
5 Duncan Luce and Howard Raiffa point out that “[i]n some ways the name “game theory” is unfortunate, for it suggests that the theory deals with only the socially unimportant conflicts found in parlour games, whereas it is far more general than that.” (Luce and Raiffa 1957, p. 2)
conflicts in terms of differential equations has opened the door for a statistical analysis of conflicts because equations express most directly the quantitative relationships between the underlying factors of a conflict.

1.2.1 Decision Theory

Decision theory provides formal concepts for modelling a rational and intelligent agent’s decisions and, therefore, allows one to analyse an agent’s decisions in a conflict situation. Decision theory constitutes the conceptual basis for game theory in which two or more agents’ decisions, and their mutual interdependencies, are analysed.

Its aim is to state intuitive requirements for a decision-maker to count as rational and intelligent in terms of axioms expressing the properties a rational and intelligent decision-maker is expected to have with regard to his preferences. Its central result is the expected-utility maximisation theorem. This theorem claims that any agent who complies with the axioms, i.e. whose preferences can be considered as rational and intelligent, makes his decisions in a way maximising his expected-utility.

In the context of decision theory, the words “rational” and “intelligent” are used in a special sense. “Rational” applies to an agent whose preferences satisfy certain conditions, such as transitivity, monotonicity, or continuity. “Intelligent” means that an agent knows all the rules and conditions relevant to his decision and their logical consequences.6

Decision theory is usually divided into descriptive and prescriptive decision theory. Descriptive decision theory aims to model real agents’ decisions and its predictions can

---

6 Roger Myersons claims that “[...] if we develop a theory that describes the behaviour of intelligent players in some game and we believe that this theory is correct, then we must assume that each player in the game will also understand this theory and its applications.” (Myerson 1991, p. 4)
be tested empirically. Prescriptive decision theory provides a normative standard of rationality. However, the distinction between descriptive and prescriptive decision theory is vague, and it is an open question whether to interpret results, such as the expected-utility maximisation theorem, as an empirical law or rather as a normative condition for rationality.

Decisions can be classified according to the quality of information available to the decision-maker. The classification suggested by Luce and Raiffa in their book *Games and Decisions* is widely accepted. According to them, decisions “under certainty” are those in which the decision-maker knows, exactly, the outcomes of his actions. In decisions “under risk”, the decision-maker knows the outcomes of his actions with a certain probability. If the decision-maker’s actions result in certain outcomes, but he has no knowledge about the probabilities of these outcomes, the decision is to be made “under uncertainty”. Decisions under risk make up the core of decision theory.\(^7\)

As decision theory uses concepts of probability theory, we need to introduce some probability theoretical concepts before we can define the basic concepts of decision theory.

Probability theory describes a situation in terms of sample spaces, outcomes, and events.

**Definition 1  (Sample Space, Outcome, Event)**

A sample space is a set \(X = \{x_1, x_2, \ldots\}\).\(^8\) The elements of \(X\) are called outcomes. An event \(S\) over a sample space \(X\) is a subset \(S \subseteq X\). The set of all events over \(X\) is denoted by \(\Xi(X) = \{S \mid S \subseteq X\} = \mathcal{P}(X)\).

The core concept of probability theory is the concept of probability distributions.

\(^7\) (Luce and Raiffa 1957, p. 13)

\(^8\) In the context of concrete conflicts we are only dealing with finite sample spaces. For a brief overview of finite probability theory, see (Mendelson 2004, p. 207ff).
Definition 2  (Probability Distribution)
A probability distribution $p$ over a sample space $X$ is a function $p: X \rightarrow [0, 1]$ from $X$ into the real interval $[0, 1]$ such that $\sum_{x \in X} p(x) = 1$. The set of all probability distributions over $X$ is denoted by $\Delta(X) = \{ p \in [0, 1]^X \mid \sum_{x \in X} p(x) = 1 \}$.

For example, the expression $p(x) = 0.7$ is to be interpreted as “the outcome $x$ occurs with probability 0.7”.

Probability is usually interpreted as the relative frequency of the occurrence of an outcome $x$. However, in decision theory, probabilities are also interpreted as credence functions, i.e. a decision-maker’s subjective assignments of probabilities to outcomes.

The concept of conditional probability, which is defined for pairs of outcomes and events, is crucial to decision theory.

Definition 3  (Conditional Probability Function)
A conditional probability function $p$ over a sample space $X$ is a function $p: \Xi(X) \rightarrow \Delta(X)$ assigning probabilities to every outcome $x \in X$ and every event $S \in \Xi(X)$ such that $p(x \mid S) = 0$ if $x \not\in S$ and $\sum_{x \in S} p(x \mid S) = 1$. The conditional probability of an event $R \in \Xi(X)$ is defined as $p(R \mid S) = \sum_{x \in R} p(x \mid S)$.

For example, the expression $p(x \mid S) = 0.7$ is to be interpreted as “given that the event $S$ has occurred, the outcome $x$ occurs with probability 0.7”.

Conditional probabilities, as well as probabilities for the union, the intersection, and the complement of events, can be determined by the usual equations.

Theorem 1  (Union, Intersection, Complement, Conditional Probability)
If $X$ is a sample space and $R, S \in \Xi(X)$ are events, then following equations hold:
1. $p(R \cup S) = p(R) + p(S) - p(R \cap S)$;
2. $p(R \cap S) = p(R \mid S) \cdot p(S)$;
3. $p(X/R) = p(\neg R) = 1 - p(R)$;
4. $p(R \mid S) = p(R \cap S)/p(S)$.
Proof
The proof is standard and can be found in introductions to probability theory.

QED

To model a situation in terms of decision theory or to solve a decision problem, the following five concepts need to be identified.

Definition 4 (Decision Problem)
A decision problem is described by the following five elements:

1. A set of lotteries \( L = \{f_1, f_2, \ldots\} \);
2. A set of states of nature \( \Omega = \{t_1, t_2, \ldots\} \);
3. A set of prices \( X = \{x_1, x_2, \ldots\} \);
4. A conditional probability function \( p: \Xi(\Omega) \rightarrow \Delta(\Omega) \);
5. A utility function \( u: X \times \Omega \rightarrow [0, 1] \).

The set of states of nature \( \Omega = \{t_1, t_2, \ldots\} \) expresses all external factors relevant to the decision problem and grouped into a number of cases \( t_1, t_2, \ldots \). For example, when deciding whether or not to intervene in a conflict, we can consider the set \( \{t_1 = \text{high level of violence}, t_2 = \text{low level of violence}\} \) containing two states of nature which express the possible levels of violence of the parties’ reaction to the intervention.

Depending on his decision and on the state of nature that becomes true, the decision-maker gets a certain price \( x \in X \). The set of prices \( X = \{x_1, x_2, \ldots\} \) expresses all possible outcomes of the decision. For example, possible outcomes of intervening or not in a violent conflict could be the four prices \( \{x_1 = \text{casualties/ unsuccessful}, x_2 = \text{no casualties/ unsuccessful}, x_3 = \text{casualties/ successful}, x_4 = \text{no casualties/ successful}\} \). Both \( \Omega \) and \( X \) must be exclusive and exhaustive descriptions of the situation in question.

Lotteries express the decision-maker’s alternatives or options. By choosing a lottery \( f \), he determines his chances of getting the prices \( x_1, x_2, \ldots \) for every state \( t_1, t_2, \ldots \). Therefore, each lottery specifies a probability distribution over the set of prices \( X \) for every state \( t \in \Omega \). A formal definition of lotteries can be given as follows.
Chapter 1 Reflecting on Conflicts

Definition 5 (Lottery)
A lottery $f$ is a function $f : \Omega \rightarrow \Delta(X)$ from the set of states of nature $\Omega$ into the set of all probability distributions $\Delta(X)$ over the set of prices $X$. The set $L = \{ f \mid f \in \Delta(X)^\Omega \}$ denotes the set of all lotteries.

For example, if the decision-maker chooses a lottery $f$ with $f(x \mid t) = 0.7$, and $t$ will become the true state of nature, he gets price $x$ with probability 0.7.

A decision-maker can also choose a compound lottery $\alpha f_i + (1 - \alpha)f_j$, where $f_i, f_j \in \mathbb{L}$ and $\alpha \in [0, 1]$. In this case, he chooses $f_i$ with probability $\alpha$ and $f_j$ with probability $(1 - \alpha)$. A lottery that gives a decision-maker a certain price $x$ in every state of nature with probability 1 is denoted by the expression $[x]$, i.e. $[x](x \mid t) = 1$ for all $t \in \Omega$.

Assumptions about which state of nature will become true are expressed by a conditional probability function $p$. Taking into account that the decision-maker may already have some knowledge about the world to the extent that he knows that the true state of nature is in a certain event $S$, $p$ depends on the set $\Xi(\Omega)$ of events over $\Omega$. In the context of decision theory, we can define the conditional probability function $p$ of a decision-maker as follows.

Definition 6 (Conditional Probability Function of a Decision-Maker)
A conditional probability function of a decision-maker is a conditional probability function $p$ over the set $\Omega$ of states of nature, i.e. $p(t \mid S) = 0$ if $t \notin S$ and $\sum_{t \in S} p(t \mid S) = 1$.

In the context of decision theory, we interpret an expression of the form $p(t \mid S) = 0.7$ as “if the decision-maker knows that the true state of nature is in the event $S$, then he believes that $t$ will become the true state of nature with probability 0.7”.

---

9 This notation is also used in (Myerson 1991).
The set of prices $X$ defines the possible outcomes of a decision, but it does not evaluate them. In order to model a decision problem, we also need to know how the decision-maker values the different outcomes. In our example, the price $x_1 = \text{casualties/ unsuccessful}$ is, for instance, probably judged worse than the price $x_4 = \text{casualties/ successful}$. Introducing a utility function $u$ is one way of representing the values a decision-maker assigns to the possible outcomes of a decision problem. A utility function $u$ assigns a number of the real interval $[0, 1]$ to every price $x \in X$ and every state of nature $t \in \Omega$ and measures the gains from getting a certain price in a particular state of nature.

**Definition 7 (Utility Function)**
A utility function $u$ is a function $u: X \times \Omega \to [0, 1]$ from the Cartesian product $X \times \Omega$ into the real interval $[0, 1]$. $u$ is called state independent if and only if $u(x, t_i) = u(x, t_j)$ for all $x \in X$ and all $t_i, t_j \in \Omega$.

The definition shows that one price can be judged differently in two different states of nature unless the utility function is state independent.

Having introduced all the necessary elements of a decision problem, i.e. states of nature, prices, lotteries, conditional probability functions, and utility functions, we can now look at the concept of expected-utility. Expected-utility is calculated for lotteries. It expresses the utility a decision-maker can rationally expect to get if he chooses a certain lottery $f$. Expected-utility depends on the decision-maker’s conditional probability function $p$, his utility function $u$, and the event $S$, which he knows contains the true state of nature.

**Definition 8 (Expected-Utility)**
If $p$ is a conditional probability function, $u$ a utility function, $f \in L$ a lottery, and $S$ an event, then the expected-utility $E_p(u(f) \mid S)$ of $f$ for a given $u, p$, and $S \in \Xi(\Omega)$ is defined as $E_p(u(f) \mid S) = \sum_{t \in S} (p(t) \mid S) \cdot \sum_{x \in X} (f(x \mid t) \cdot u(x, t))$. 
In the following, we illustrate the basic concepts of decision theory with an example modelling the decision of whether to intervene in a violent conflict by means of ground forces, ground and air forces, air forces, or a diplomatic mission.\(^\text{10}\)

In the example, the set of lotteries is given by \(L = \{f_1 = \text{ground forces}, f_2 = \text{ground and air forces}, f_3 = \text{air forces}, f_4 = \text{diplomatic mission}\}\). The set of states of nature is given by the set \(\Omega = \{t_1 = \text{high level of violence}, t_2 = \text{low level of violence}\}\), and reflects the possible levels of violence of the parties’ reaction to the intervention. We look at four possible outcomes of the intervention and characterise them by the set of prices \(X = \{x_1 = \text{casualties/unsuccessful}, x_2 = \text{no casualties/unsuccessful}, x_3 = \text{casualties/successful}, x_4 = \text{no casualties/successful}\}\). By “casualties” we mean casualties of the intervention force and by “successful” that the intervention could end the violent conflict.

The assignment of probability distributions over the set of prices to every lottery and every state of nature is guided by the following six considerations.

1. The probability of having casualties is zero if we launch a diplomatic mission and is highest if we deploy ground forces.

2. Success is more likely if we deploy ground forces than if we deploy air forces.

3. If we launch a diplomatic mission and the level of violence is high, the chance of success is almost zero.

4. If we launch a diplomatic mission and the level of violence is low, the chance of success is high.

5. For all lotteries, the chance of success is higher if the level of violence is low than if the level of violence is high.

\(^{10}\) The example is just hypothetic. Similar examples, modelling strategic decisions in military contexts, can be found in (Straffin 1993, p. 27ff; Taylor 1995, p. 166ff; Ordeshook 1986, p. 220ff).
For all lotteries, the probability of having casualties is higher if the level of violence is high than if the level of violence is low.

A set of lotteries that satisfies all these conditions is shown in the following table.

<table>
<thead>
<tr>
<th>f₁ = ground forces</th>
<th>t₁ = high level of violence</th>
<th>t₂ = low level of violence</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₁(x₁</td>
<td>t₁) = 0.27</td>
<td>f₁(x₁</td>
</tr>
<tr>
<td>f₁(x₂</td>
<td>t₁) = 0.63</td>
<td>f₁(x₂</td>
</tr>
<tr>
<td>f₁(x₃</td>
<td>t₁) = 0.03</td>
<td>f₁(x₃</td>
</tr>
<tr>
<td>f₁(x₄</td>
<td>t₁) = 0.07</td>
<td>f₁(x₄</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f₂ = ground and air forces</th>
<th>t₁ = high level of violence</th>
<th>t₂ = low level of violence</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₂(x₁</td>
<td>t₁) = 0.3</td>
<td>f₂(x₁</td>
</tr>
<tr>
<td>f₂(x₂</td>
<td>t₁) = 0.3</td>
<td>f₂(x₂</td>
</tr>
<tr>
<td>f₂(x₃</td>
<td>t₁) = 0.2</td>
<td>f₂(x₃</td>
</tr>
<tr>
<td>f₂(x₄</td>
<td>t₁) = 0.2</td>
<td>f₂(x₄</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f₃ = air forces</th>
<th>t₁ = high level of violence</th>
<th>t₂ = low level of violence</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₃(x₁</td>
<td>t₁) = 0.24</td>
<td>f₃(x₁</td>
</tr>
<tr>
<td>f₃(x₂</td>
<td>t₁) = 0.06</td>
<td>f₃(x₂</td>
</tr>
<tr>
<td>f₃(x₃</td>
<td>t₁) = 0.56</td>
<td>f₃(x₃</td>
</tr>
<tr>
<td>f₃(x₄</td>
<td>t₁) = 0.14</td>
<td>f₃(x₄</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f₄ = diplomatic mission</th>
<th>t₁ = high level of violence</th>
<th>t₂ = low level of violence</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₄(x₁</td>
<td>t₁) = 0.0</td>
<td>f₄(x₁</td>
</tr>
<tr>
<td>f₄(x₂</td>
<td>t₁) = 0.0</td>
<td>f₄(x₂</td>
</tr>
<tr>
<td>f₄(x₃</td>
<td>t₁) = 0.95</td>
<td>f₄(x₃</td>
</tr>
<tr>
<td>f₄(x₄</td>
<td>t₁) = 0.05</td>
<td>f₄(x₄</td>
</tr>
</tbody>
</table>

Table 1: Lotteries

Weighing the various outcomes in terms of utility is a difficult task. However, x₁ = casualties/unsuccessful is definitely the worst price, and x₄ = no casualties/successful is the best price. Therefore, they should get the lowest and highest utility value, respectively. The prices x₂ = casualties/successful and x₃ = no casualties/unsuccessful are weighed equally here, although one might consider own losses less important if a violent conflict can successfully be stopped or vice versa. A state independent utility function expressing this evaluation of prices is shown in the following table.

<table>
<thead>
<tr>
<th>x</th>
<th>u(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>x₁ = casualties/unsuccessful</td>
<td>0.0</td>
</tr>
<tr>
<td>x₂ = casualties/successful</td>
<td>0.5</td>
</tr>
<tr>
<td>x₃ = no casualties/unsuccessful</td>
<td>0.5</td>
</tr>
<tr>
<td>x₄ = no casualties/successful</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 2: Utilities
Before we can calculate the expected-utilities for our lotteries, we have to guess probabilities for the two states of nature $t_1$ and $t_2$. We can do this by using a simple probability distribution over $\{t_1, t_2\}$ as there is no event $S \subset \{t_1, t_2\}$ which we know to contain the true state of nature, i.e. we assume $S = \Omega$.

We will calculate expected-utilities for three different probability distributions. The probability function $p_1$ characterises the situation in which the parties’ reacting with a high level of violence is considered equally likely to their reacting with a low level, i.e. $p_1(t_1) = p_1(t_2) = 0.5$. According to $p_2$, it is certain that the level of violence will be high, i.e. $p_2(t_1) = 1$ and $p_2(t_2) = 0$, whereas in $p_3$ the level of violence will certainly be low, i.e. $p_3(t_1) = 0$ and $p_3(t_2) = 1$.

<table>
<thead>
<tr>
<th></th>
<th>$p_1(t)$</th>
<th>$p_2(t)$</th>
<th>$p_3(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$ = high level of violence</td>
<td>0.5</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$t_2$ = low level of violence</td>
<td>0.5</td>
<td>0.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 3: Probability Distributions

Expected-utility can now be computed for each of the four lotteries and each probability function.

|          | $E_{p_1}(u(f) | \Omega)$ | $E_{p_2}(u(f) | \Omega)$ | $E_{p_3}(u(f) | \Omega)$ |
|----------|--------------------------|--------------------------|--------------------------|
| $f_1$ = ground forces | 0.5                      | 0.4                      | 0.6                      |
| $f_2$ = ground and air forces | 0.55                    | 0.45                     | 0.65                     |
| $f_3$ = air forces | 0.55                     | 0.45                     | 0.65                     |
| $f_4$ = diplomatic mission | 0.75                    | 0.525                    | 0.975                    |

Table 4: Expected-Utilities

The example shows that for the given utility function $u$ and lotteries $f_1$, $f_2$, $f_3$, and $f_4$, launching a diplomatic mission is the option that maximises expected-utility regardless of how probability is distributed over the two states of nature.
Utility and conditional probability functions provide a means to evaluate lotteries with regard to their expected-utility. In the previous example, we can conclude that we should choose lottery $f_4$ first, lotteries $f_2$ and $f_3$ second, and lottery $f_1$ last if we want to maximise expected-utility.

A different way of looking at decision problems is to ask a decision-maker directly about his preferences with regard to the available lotteries. In this case, decision theory provides criteria that allow one to assess whether or not a decision-maker’s preferences can be considered rational. The most straightforward criterion is transitivity: a decision-maker fails to be considered rational if he prefers a lottery $f_i$ over a lottery $f_2$, $f_2$ over a lottery $f_3$, and $f_3$ over $f_i$. Further criteria are monotonicity, completeness, continuity, and independence. Preferences are expressed by means of a binary preference relation $\leq_S$.

This relation, which is defined over the set of lotteries $L$, i.e. $\leq_S \subseteq L \times L$, depends also on the set $\Xi(\Omega)$ of events. An expression $f_i \leq_S f_j$ is to be interpreted as “if the decision-maker knows that the true state of nature is in $S$, he prefers lottery $f_i$ over lottery $f_j$”.

Two relations are derived from the original preference relation. The indifference relation $\sim_S$ is defined by the condition $f_i \sim_S f_j$ iff $f_i \leq_S f_j$ and $f_j \leq_S f_i$. The strict preference relation $<_S$ is defined by $f_i <_S f_j$ iff $f_i \leq_S f_j$ and not $f_i \sim_S f_j$.

In the following definition we state the axioms for rational decision-makers’ preferences.

**Definition 9**  (Axioms for Rational Decision-makers’ Preferences)

A preference relation $\leq_S$ is a preference relation of a rational decision-maker iff it satisfies the following axioms:

1. $\forall f_i, f_j (f_i \leq_S f_j \lor f_j \leq_S f_i)$ (Completeness);
2. $\forall f_i, f_j, f_k (f_i \leq_S f_j \land f_j \leq_S f_k \Rightarrow f! \leq_S f_k)$ (Transitivity);
3. $\forall f_i, f_j, \forall \alpha, \beta (f_i <_S f_j \land \alpha < \beta \Rightarrow \alpha f_i + (1 - \alpha) f_j \leq_S \beta f_i + (1 - \beta) f_j)$ (Monotonicity$^{11}$);

$^{11}$ The axiom of monotonicity is not an independent axiom as it can be derived from the other axioms. Cf. (Myerson 1991, p. 33).
The preference relation is complete if, for any two lotteries \( f_i \) and \( f_j \), the decision-maker either prefers \( f_i \) over \( f_j \) or \( f_j \) over \( f_i \). When the decision-maker preferring \( f_i \) over \( f_j \) and \( f_j \) over \( f_k \) implies he prefers \( f_i \) over \( f_k \), the relation is transitive. The preference relation is monotone if the decision-maker strictly preferring \( f_i \) over \( f_j \) implies he prefers those compound lotteries assigning a higher probability of getting \( f_i \) over those compound lotteries assigning a lower probability of getting \( f_i \). The preference relation is continuous in the sense that if the decision-maker prefers \( f_i \) over \( f_j \) and \( f_j \) over \( f_k \), then there is a compound lottery of \( f_i \) and \( f_k \) such that the decision-maker is indifferent between this lottery and \( f_j \). It is independent if the decision-maker’s preference of \( f_i \) over \( f_j \) implies that he prefers every compound lottery involving \( f_i \) over the corresponding compound lottery involving \( f_j \).

The axioms are supposed to provide a formalisation of intuitive principles of rationality. Therefore, a decision-maker who violates one of them cannot be considered as rational, according to standard decision theory. Although most of them seem intuitively plausible at first glance, several authors have questioned their status as axioms of rationality. In empirical tests, it can be shown that people’s preferences do not necessarily comply with some of the axioms, such as completeness and independence.\(^{12}\) However, the axioms provide a normative framework for rational decision-making. Moreover, they provide a procedure for finding appropriate utility functions in a decision-problem.

\(^{12}\) A famous counterexample to the axiom of independence is the Allais Paradox. See (Allais 1953). For further studies exploring the empirical status of the axioms, see (Ross 2006; Hansson 2005, p. 35f).
The expected-utility maximisation theorem was proven by John von Neumann and Oskar Morgenstern in their joint book *Theory of Games and Economic Behaviour*. It links a rational and intelligent decision-maker’s preference relation to the concept of utility functions as introduced earlier in the section. Indeed, it says that these two concepts are equivalent. If a decision-maker has a complete, transitive, continuous, and independent preference relation over the set \( L \) of all lotteries, then this preference relation provides an effective procedure to construct a utility function \( u \) and a conditional probability function \( p \) such that the expected-utility of any lottery \( f_i \) is higher than the expected-utility of lottery \( f_j \) if and only if the decision-maker prefers lottery \( f_i \) over \( f_j \).

On the other hand, it can be shown that the ordering over the set \( L \) of all lotteries induced by the expected-utility function \( E_p(u(f) | S) \) satisfies the axioms of completeness, continuity, monotonicity, transitivity, and independence and is, therefore, equivalent to a preference relation \( \leq_S \).

The theorem is given as follows.

**Theorem 2 (Expected-Utility Maximisation)**

In a decision problem there exists a preference relation \( \leq_S \) that satisfies the axioms of completeness, monotonicity, continuity, transitivity, and independence iff there exists a utility function \( u \) and a conditional-probability function \( p \) such that \( f_i \leq_S f_j \) iff \( E_p(u(f_i) | S) \leq E_p(u(f_j) | S) \).

**Proof**

The proof is standard and can be found in introductions to decision-theory. Here, we sketch only the basic idea, which consists of the construction of a utility function \( u \) and a conditional-probability function \( p \) on the basis of a given preference relation \( \leq_S \).

First, we have to transfer the preference relation onto the set of prices \( X \) for every state \( t \in \Omega \). This can be done by restricting \( \leq_S \) to \( t \), i.e. by looking at the preference relation \( \leq_{(t)} \). The new ordering over the set of prices \( X \) allows one to

---

13 (von Neumann and Morgenstern 1953)

14 For instance, in (Myerson 1991, p. 14f).
pick one of the best prices $x_{\max}$ and one of the worst prices $x_{\min}$ such that for all $x \in X$, we have $x_{\min} \leq x \leq x_{\max}$.

In a second step, we define two lotteries $f_{\max}$ and $f_{\min}$, where $f_{\max}$ certainly results in the best price $x_{\max}$ for every state $t \in \Omega$, and $f_{\min}$ certainly results in the worst price $x_{\min}$ for every $t \in \Omega$, i.e. $f_{\max} = \lceil x_{\max} \rceil$ and $f_{\min} = \lfloor x_{\min} \rfloor$. For every state $t \in \Omega$ and every price $x \in X$, we ask the decision-maker to give a compound lottery $\alpha f_{\max} + (1 - \alpha) f_{\min}$ such that he is indifferent between the lottery $[x]$, which gives price $x$ in every state, and the compound lottery given that $t$ is the true state of nature. According to the axiom of continuity, such lottery exists. We then take $\alpha$ as the utility of price $x$ in state $t$, i.e. $u(x, t) = \alpha$.

By a similar procedure we can construct a conditional-probability function $p$. Finally, $u$ and $p$ can be shown to satisfy the condition $f_i \leq S f_j$ if and only if $E_p(u(f_i) | S) \leq E_p(u(f_j) | S)$.

QED

The theorem implies that if a decision-maker chooses the lottery that is highest with regard to his preference relation $\leq_S$, then he simultaneously chooses the lottery with the highest value of expected-utility, i.e. he maximises his expected-utility.

1.2.2 Game Theory

In game theory, the participants of a game are called players. They can be characterised as social units such as individuals, social groups, social institutions etc. In the context of conflicts, typical players are, for instance, states, guerrilla groups, trade unions, firms, colleges, or spouses. It is assumed that the number of players is finite and greater or equal to 2. The set of all players is abbreviated by $N = \{1, 2, \ldots, n\}$.

In a game, each player chooses a strategy out of a fixed number of available strategies. Strategies are possible actions that the players can carry out within a game. If a player chooses a strategy, his behaviour in the game is completely determined by this strategy. The strategy determines a move for every state of the game. The set of all strategies available to a player $i$ is abbreviated by $C_i = \{c_{i1}, c_{i2}, \ldots, c_{ik}\}$ with every $c_{ij} \in C_i$ symbolising one strategy of the player $i$. Besides choosing a pure strategy, i.e. one single
element of $C_i$, a player can also choose a randomised strategy, i.e. a combination of elements of $C_i$, each of which he then plays with a certain probability $\alpha$.

If every player $i \in N$ chooses a strategy $c_{ij} \in C_i$, the game can be played leading to an outcome in which each player gets a certain price. The subjective value of the price a player gets in a certain outcome is measured by an ordinal or cardinal utility function $u_i$. Hence, utility functions specify the utility of the price a player gets in a certain outcome for every player $i \in N$ and every possible outcome of the game.

It may be the case that the outcomes of a game depend on some arbitrary factors outside the players’ influence. However, from a game theoretic point of view, each player must decide how to behave in every possible outcome of the arbitrary factor before the game is played. Games without arbitrary factors are called deterministic games.

Another parameter that is modelled by game theory is information. In each phase of a game, the players have information about the game, such as information about the other players’ moves, their own moves, and the outcomes of arbitrary factors. This information is expressed by information states assigned to the players in each state of the game. If a player’s information states are identical in two different states of the game, he is not able to decide in which state the game is. Games in which all players always know the current state of the game are called games with perfect information. Otherwise, the game is a game with imperfect information.

Most of the literature deals with 2-person games, i.e. games with $N = \{1, 2\}$. An important subclass of these games is the class of 2-person zero-sum games. In these
games the sum of the utilities of the two players is zero in any outcome, i.e. the gain of the winner equals the loss of the looser.\footnote{2-person games are simple and cover a wide range of applications. Zero-sum games provide a natural model for situations in which the competitive aspect is very high. Cf., for instance, (Mendelson 2004, p. 53ff).}

In its extensive form, a game is displayed as a rooted tree. Rooted trees are a certain type of graph, where a graph is a finite set of nodes \( \{x_1, x_2, \ldots, x_k \} \) together with a set of pairs of nodes. In the context of game theory, a pair of nodes \( <x_i, x_j> \) is also called a branch. A path is a set \( \{<x_1, x_2>, <x_2, x_3>, \ldots, <x_{m-1}, x_m>\} \) of pairs of nodes, where all \( x_i \) occurring in the path are distinct nodes and \( m \geq 2 \). A path connects the two nodes \( x_i \) and \( x_{m} \). A rooted tree can be defined as follows.

**Definition 10 (Rooted Tree)**

A rooted tree is a graph in which any two nodes are connected by exactly one path, and one node is designated as the root of the tree.

The unique path connecting a node \( x_i \) and the root of the tree is referred to as the path of \( x_i \). Branches connecting a node \( x_i \) with other nodes are called alternatives of \( x_i \) if they are not in the path of \( x_i \). Nodes without alternatives are called terminal nodes. The following figure illustrates the concept of a rooted tree.

![Figure 1: Rooted Tree](image-url)
A game in extensive form consists of a rooted tree in which all nodes and branches are labelled in the following way: each node that is not a terminal node has exactly one player label. Player labels are natural numbers \( \{0, 1, \ldots, n\} \) expressing which player controls the node. Nodes labelled by 0 are called chance nodes. Each alternative of a chance node is labelled by a number of the real interval \([0, 1]\) specifying the probability of the alternative. The sum of all probabilities assigned to the alternatives of a chance node equals one, i.e. we assign a probability distribution over its alternatives to every chance node. To every node that is controlled by a player, a second label is assigned expressing the information state of the player. If two nodes have the same information label, the player is not able to distinguish between the two states represented by these nodes. Consequently, the set of alternatives of two nodes with the same information label must be identical. A move label is assigned to each alternative of a node that is controlled by a player. To every terminal node, we assign a utility vector \((u_1, u_2, \ldots, u_n)\) representing the utility payoffs of the players 1, 2, \ldots, n if the game terminates at this node.

Figure 2 shows a game in extensive form between the two players 1 and 2.\(^{16}\) The game starts with a chance node at which two possible outcomes can occur each with equal probability 0.5. Then, player 1 can choose between the two moves r and s. The two information states A and B indicate that player 1 knows which of the two outcomes of the chance node has occurred at the beginning of the game. If player 1 chooses r, the game is over. Player 1 gets the utility payoff 1 or -1, and player 2 gets the payoff -1 or 1, depending on the outcome of the chance node. If player 1 chooses the alternative s, then player 2 can choose between the two moves u and v, and the game is over. However, when choosing between u and v, player 2 does not know which of the two outcomes of

\(^{16}\)This example shows a game-theoretic model of a simple card game. For an exhaustive description of this game, see (Myerson 1991, p. 37ff).
the chance node has occurred at the beginning of the game as both nodes have the same information label C. Utility payoffs for the two players are shown at the terminal nodes, respectively.

![Figure 2: Extensive Form](image)

The extensive form of a game explicitly shows its dynamic and informational structure. The players’ strategies are better reflected in the strategic form of a game. However, they can also be obtained from the extensive form by specifying a move for each information state. For example, a strategy for player 1 in the game of Figure 2 could be to choose r if he is in information state A and s if he is in B.\(^{17}\)

The strategic form of a game specifies a set of players \(N = \{1, 2, \ldots, n\}\) participating in the game, a set of strategies \(C_i\) for each player \(i \in N\), and a set of expected-utility functions \(u\) from the Cartesian product \(\times_{j \in N} C_j\) into the set of real numbers for each player \(i \in N\). Elements of \(\times_{j \in N} C_j\) are called strategy profiles. They represent the outcomes of the game by specifying for each player the strategy that he chooses. If \(c \in \times_{j \in N} C_j\) is a strategy profile, then \(u_i(c)\) expresses the expected-utility payoff of player \(i\) if all players choose

\(^{17}\) A procedure constructing the strategic form of a game in extensive form is described at the end of the section.
their strategy according to $c$. Formally, we can define a game in strategic form as follows.\footnote{A similar notation is used in (Myerson 1991).}

**Definition 11 (Games in Strategic Form)**
A game in strategic form is a structure $\Gamma = (N, (C_i)_{i \in N}, (u_i)_{i \in N})$, where $N$ is a finite set of players, $(C_i)_{i \in N}$ is a sequence of sets $C_i$ such that every $C_i$ is a set of strategies of player $i$, and $(u_i)_{i \in N}$ is a sequence of expected-utility functions $u_i$ such that every $u_i$ is a function from $\prod_{j \in N} C_j$ into $\mathbb{R}$.

A game in strategic form can be displayed as a matrix. If there are only two players 1 and 2, the matrix has the simple form shown in the following table.

<table>
<thead>
<tr>
<th></th>
<th>$c_{11}$</th>
<th>$c_{12}$</th>
<th>$c_{13}$</th>
<th>$c_{1m}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$u_1(c_{11}, c_{21})$, $u_2(c_{11}, c_{21})$</td>
<td>$u_1(c_{11}, c_{22})$, $u_2(c_{11}, c_{22})$</td>
<td>$u_1(c_{11}, c_{23})$, $u_2(c_{11}, c_{23})$</td>
<td>$u_1(c_{11}, c_{2m})$, $u_2(c_{11}, c_{2m})$</td>
</tr>
<tr>
<td>2</td>
<td>$u_1(c_{12}, c_{21})$, $u_2(c_{12}, c_{21})$</td>
<td>$u_1(c_{12}, c_{22})$, $u_2(c_{12}, c_{22})$</td>
<td>$u_1(c_{12}, c_{23})$, $u_2(c_{12}, c_{23})$</td>
<td>$u_1(c_{12}, c_{2m})$, $u_2(c_{12}, c_{2m})$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$c_{1n}$</td>
<td>$u_1(c_{1n}, c_{21})$, $u_2(c_{1n}, c_{21})$</td>
<td>$u_1(c_{1n}, c_{22})$, $u_2(c_{1n}, c_{22})$</td>
<td>$u_1(c_{1n}, c_{23})$, $u_2(c_{1n}, c_{23})$</td>
<td>$u_1(c_{1n}, c_{2m})$, $u_2(c_{1n}, c_{2m})$</td>
</tr>
</tbody>
</table>

Table 5: 2-Person Game in Strategic Form

A player can either choose a pure strategy, i.e. an element of $C_i$, or a randomised strategy, i.e. a probability distribution over $C_i$. If a player chooses a randomised strategy, he plays every pure strategy included in the combination with a certain probability. For instance, player 1 in Table 5 could play the randomised strategy $0.5[c_{11}] + 0.3[c_{14}] + 0.2[c_{17}]$. The set of all randomised strategies of a player $i$ can be identified with the set of probability distributions $\Delta(C_i)$ over the set $C_i$ of strategies of $i$. The expression $\sigma(c_i)$ denotes the probability of player $i$ playing strategy $c_i$ in the randomised strategy $\sigma$. For instance, we have $(0.5[c_{11}] + 0.3[c_{14}] + 0.2[c_{17}])_i(c_{14}) = 0.3.$
In the strategic form of a game, we have no information about the temporal structure of the game, nor do we know the players’ information states during the game. Instead, we have information about the various strategies of the players, and we know the players’ expected-utilities for every outcome of the game.

If a game is given in extensive form, we can construct its strategic form $\Gamma$ as follows: the set of players $N$ is identical in both forms. Let $H_i = \{h_1, h_2, \ldots, h_p\}$ be the set of information states of player $i$ in the extensive form, and let $M_i(h)$ be the set of all alternatives of player $i$ if he is in the information state $h$. Then, $C_i$, i.e. the set of strategies of player $i$, can be defined as the set $\times_{1 \leq k \leq p} M_i(h_k)$. A strategy picks exactly one alternative for every information state from the set of available alternatives of that state.

For example, in the game in extensive form of Figure 2, player 1 has the four strategies $rr$, $rs$, $sr$, and $ss$ as we have $H_1 = \{A, B\}$, $M_1(A) = \{r, s\}$, and $M_1(B) = \{r, s\}$ and, therefore, $C_1 = \times_{1 \leq k \leq 2} M_i(h_k) = \{r, s\} \times \{r, s\} = \{rr, rs, sr, ss\}$.

The expected-utility $u_i(c)$ of player $i$, if the players choose their strategies according to the strategy profile $c$, is the sum of the utility payoffs of $i$ in those terminal nodes that can possibly be reached according to $c$ weighed by the probability of their occurrence. The weighing is necessary as the chance nodes in the extensive form have to be taken into account. In the game of Figure 2, we have, for instance, $u_1(rs, u) = 0.5 \cdot 1 + 0.5 \cdot (-2) = -0.5$ and $u_2(rs, u) = 0.5 \cdot (-1) + 0.5 \cdot 2 = 0.5$.

If we transfer a game in extensive form into strategic form, we call the result the normal form of the game. The following table shows the normal form of the game in extensive form of Figure 2.
If a situation is modelled as a game in extensive form or strategic form, we are interested in the question of which strategy each player will choose. An answer to this question can be understood either descriptively or prescriptively. In both cases, the answer relies on the assumption that every player is rational and intelligent, i.e. his goal is to maximise his expected-utility, and he is aware of any consequence of his actions for him and all other players.

A first, simple way to answer this question is the concept of domination. If a strategy of a player results in a smaller amount of utility payoff than another strategy, regardless of the other players’ behaviour, we say that the first strategy is strongly dominated by the second one. A utility maximising player never chooses a dominated strategy as he can certainly obtain a higher amount of expected-utility by choosing the dominating strategy.

Domination holds not only for pure strategies, but also for randomised strategies. In the example of Table 6, the randomised strategy 0.5[sr] + 0.5[ss] strongly dominates the pure strategy rr. This can be shown by comparing the expected-utility payoffs of the two strategies for player 1 for every choice between u and v of player 2. If player 2 chooses u, we get the expected-utility payoffs $u_i(0.5[\text{sr}] + 0.5[\text{ss}], u) = 0.5 \cdot 0.5 + 0.5 \cdot 0 = 0.25$ and $u_i(\text{rr}, u) = 0$. If player 2 chooses v, we get $u_i(0.5[\text{sr}] + 0.5[\text{ss}], v) = 0.5 \cdot 0 + 0.5 \cdot 1 = 0.5$ and $u_i(\text{rr}, v) = 0$. In both cases, player 1 obtains a higher amount of expected-utility by choosing the randomised strategy $0.5[\text{sr}] + 0.5[\text{ss}]$. 

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>u</td>
<td>v</td>
</tr>
<tr>
<td>rr</td>
<td>0,0</td>
<td>0,0</td>
</tr>
<tr>
<td>rs</td>
<td>-0.5,0.5</td>
<td>1.0,-1.0</td>
</tr>
<tr>
<td>sr</td>
<td>0.5,-0.5</td>
<td>0,0,0</td>
</tr>
<tr>
<td>ss</td>
<td>0,0,0</td>
<td>1.0,-1.0</td>
</tr>
</tbody>
</table>

Table 6: Normal Form
In order to define the concept of domination, we have to introduce two further notions. A strategy profile $c \in \prod_{j \in N} C_j$ specifies exactly one strategy for each player. If we delete the set of strategies $C_i$ of player $i$ from $\prod_{j \in N} C_j$, we obtain the set $C_{-i} = \prod_{j \in N \setminus \{i\}} C_j$ expressing all strategy profiles that the players other than $i$ can possibly choose. Elements of $C_{-i}$ are denoted by $c_{-i}$. If player $i$ chooses a certain strategy $d_i$, $(d_i, c_{-i})$ denotes the strategy profile in which $i$ chooses strategy $d_i$ and the other players choose their strategies according to $c_{-i}$.

Now we can define the concept of domination as follows.

**Definition 12 (Domination)**

A strategy $d_i \in \prod_{j \in N} C_j$ is strongly dominated for player $i$ iff there is a randomised strategy $\sigma_i \in \Delta(C_i)$ such that for all strategy profiles $c_i \in C_i$, the following condition holds: $u_i(d_i, c_i) < \sum_{e \in C_i}(\sigma_i(e) u_i(e, c_{-i}))$.

A strategy $d_i \in \prod_{j \in N} C_j$ is weakly dominated for player $i$ iff there is a randomised strategy $\sigma_i \in \Delta(C_i)$ such that for all strategy profiles $c_i \in C_i$, the following condition holds: $u_i(d_i, c_i) \leq \sum_{e \in C_i}(\sigma_i(e) u_i(e, c_{-i}))$.

As rational and intelligent players never choose strongly dominated strategies, we can eliminate them from the strategic form of a game. In the game shown in the following table, we can first eliminate strategy $c_3$, which is dominated by $0.2[c_1] + 0.8[c_2]$, then $d_2$ dominates $d_1$, and finally, $c_1$ dominates $c_2$.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>d_1</strong></td>
<td>1, 0</td>
<td>1, 1</td>
</tr>
<tr>
<td><strong>d_2</strong></td>
<td>3, 1</td>
<td>0, 2</td>
</tr>
<tr>
<td><strong>c_3</strong></td>
<td>2, 3</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

Table 7: Domination
Note that $d_2$ dominated $d_1$ only after $c_3$ was eliminated. As $(c_1, d_2)$ is the only remaining strategy profile after iteratively eliminating dominated strategies, we can predict the outcome of the game. Of course, this procedure does not apply to all games as there are games with no dominated strategies. Furthermore, the iterative elimination of weakly dominated strategies can lead to different predictions, depending on the sequence in which the elimination was carried out.

A more general solution concept is that of a Nash equilibrium. A Nash equilibrium is a strategy profile with the property that no player can improve his expected-utility by switching to another strategy, given that all the other players retain their strategies. For the case that the Nash equilibrium consists of pure strategies, we can define it as follows.

**Definition 13 (Nash Equilibrium in Pure Strategies)**

The pure strategy profile $(d_1, d_2, \ldots, d_n)$ is a Nash equilibrium iff for all $i \in N$ and for all $c_i \in C_i$, the following condition holds:

$$u_i(d_1, \ldots, d_{i-1}, c_i, d_{i+1}, \ldots, d_n) \leq u_i(d_1, \ldots, d_{i-1}, d_i, d_{i+1}, \ldots, d_n).$$

Nash equilibriums in pure strategies can be found by checking all the outcomes in the strategic form of the game. For example, the game in Table 7 has the unique pure Nash equilibrium $(c_1, d_2)$. The game in Table 6 has no pure Nash equilibrium.

If a Nash equilibrium is reached in a game, the players have no incentive to leave it again as this would only worsen their utility payoff. Thus, Nash equilibriums represent stable outcomes of a game. However, they are not necessarily the best outcomes for all the players in the sense of Pareto-optimality. An outcome is Pareto-optimal if there is no other outcome in which at least one player could obtain a higher amount of utility, while the other players obtain utility payoffs at least as high as in the original outcome. The

---

19 John Nash introduced the concept of equilibrium points in his paper *Equilibrium points in n-person games*. See (Nash 1950).
game in Table 7 has two Pareto-optimal outcomes, i.e. \((c_2, d_i)\) and \((c_3, d_i)\). Note that in this game, neither the solution by elimination of dominated strategies nor the Nash equilibrium is Pareto-optimal.

For the case that the Nash equilibrium involves randomised strategies, we have to introduce a further notation. If all players choose a randomised strategy, we obtain a randomised strategy profile \((\sigma_1, \ldots, \sigma_n)\) with \(\sigma_i \in \Delta(C_i)\) for all \(i \in N\). Hence, the set of all randomised strategy profiles can be defined as the Cartesian product \(\prod_{i \in N} \Delta(C_i)\).

Now we can define a Nash equilibrium in randomised strategies as follows.

**Definition 14 (Nash Equilibrium in Randomised Strategies)**

The randomised strategy profile \((\sigma_1, \ldots, \sigma_n)\) is a Nash equilibrium iff for all \(i \in N\) and all \(\tau_i \in \Delta(C_i)\), the following condition holds:

\[
u_i(\sigma_1, \ldots, \sigma_{i-1}, \tau_i, \sigma_{i+1}, \ldots, \sigma_n) \leq \nu_i(\sigma_1, \ldots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \ldots, \sigma_n).
\]

To find Nash equilibriums that involve randomised strategies, we have to simultaneously solve a number of linear equations.\(^{20}\) The equations can be found by making use of the idea that if one player chooses a randomised strategy, the probability distribution of this strategy must ensure that the expected-utility of the other players’ strategies are identical. Otherwise, they were not forced to choose a randomised strategy as they could choose the pure strategy resulting in the highest amount of expected-utility.

We illustrate this idea with the game of Table 6. In this game, player 2 has two strategies \(u\) and \(v\), and player 1 has four strategies \(rr\), \(rs\), \(sr\), and \(ss\). As \(rr\) is strongly dominated by the randomised strategy \(0.5[sr] + 0.5[ss]\), and \(rs\) is weakly dominated by \(ss\), we have to look only at the reduced strategic form of the game shown in the following table.

\(^{20}\) There are several algorithms computing Nash equilibriums. One of them is the Simplex Method which is described, for instance, in (Mendelson 2004, p. 109ff).
Let \( \alpha \) and \((1 - \alpha)\) be the probabilities of the randomised equilibrium strategy of player 1 and \( \beta \) and \((1 - \beta)\) the probabilities of player 2. If player 1 chooses the pure strategy \( sr \), he will receive the expected-utility payoff 
\[
u_1(sr, \beta[u] + (1 - \beta)[v]) = \beta \cdot 0.5 + (1 - \beta) \cdot 0 = 0.5\beta.
\]
If he chooses \( ss \), he gets 
\[
u_1(ss, \beta[u] + (1 - \beta)[v]) = \beta \cdot 0 + (1 - \beta) \cdot 1 = 1 - \beta.
\]
Player 2 has to choose \( \beta \) in such a way that player 1 is forced to randomise between \( sr \) and \( ss \). The only way to do this is by making the expected-utility payoffs of player 1 for the two pure strategies equal, i.e. 
\[
u_1(sr, \beta[u] + (\beta - q)[v]) = \nu_1(ss, \beta[u] + (1 - \beta)[v]).
\]
Therefore, our first equation is:

(I) \[0.5 \cdot \beta = 1 - \beta.\]

The same consideration for player 1 leads to the second equation:

(II) \[-0.5 \cdot \alpha = 1 - \alpha.\]

The only solution to these equations is \( \alpha = 0.66 \) and \( \beta = 0.66 \). Hence, the Nash equilibrium of the game is the randomised strategy profile \((0.66[ sr ] + 0.33[ ss ], 0.66[u] + 0.33[v])\). If any player deviates unilaterally from this strategy profile, his utility payoff will decrease.

It can be proven that for any game in strategic form, there exists at least one Nash equilibrium. The Nash equilibrium may be a pure strategy profile or a randomised strategy profile. However, the Nash equilibrium is not necessarily unique as there are
games in which more than one Nash equilibrium exist. For example, in the game shown in Table 9, there are two pure Nash equilibriums \((a_1, b_1)\), \((a_2, b_2)\) and one randomised Nash equilibrium \((0.75[a_1] + 0.25[a_2], 0.25[b_1] + 0.75[b_2])\).

<table>
<thead>
<tr>
<th></th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>b₁</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
</tr>
<tr>
<td>a₁</td>
<td>3, 1</td>
</tr>
<tr>
<td>a₂</td>
<td>0, 1</td>
</tr>
</tbody>
</table>

Table 9: Nash Equilibrium in Pure and Randomised Strategies

### 1.2.3 Robert Kowalski’s Logic-based Approach

In his 2003 paper *A logic-based approach to conflict resolution*, the British computer scientist Robert Kowalski combines formal logic with decision theory to develop a model of conflict resolution. His model is based on a cognitive model which he calls “unified logic-based agent model”. It unifies Alan Newell’s idea of production-systems, Paul Cohen’s BDI logics, and goal hierarchies, and enables us to represent crucial elements of a conflict, such as goals, beliefs, and actions, within a formal system.

Alternative solutions to a given conflict are generated systematically by means of backward and forward reasoning. Then, decision theory helps “to decide between different solutions, in the attempt to optimise their expected-utility.”

Kowalski defines a conflict in logical terms as an “inconsistency between different goals”. Conflicts can be resolved by finding a way to satisfy all goals or, if this is impossible, to satisfy each goal at least to a certain extent. Such a compromise solution

---

21 (Kowalski 2003, p. 1)
22 (Kowalski 2003, p. 1)
can be found by generalising the goals to higher-level goals which are not inconsistent any more, and then satisfying these higher-level goals.

Kowalski specifies his procedure of conflict resolution by dividing it into five steps. The steps proceed from a given set of beliefs and inconsistent goals constituting the conflict. First, consistent higher-level goals are identified for every original goal. In the second step, consistent ways of satisfying these higher-level goals are identified. In the third step, positive and negative consequences of these alternatives are inferred. Step four consists of estimating the degrees to which the alternatives satisfy the higher-level goals. In step five, a combination of alternatives satisfying the higher-level goals to the greatest extent and each individual’s higher level-goals to some minimal extent, is chosen.

Kowalski uses BDI logic to represent goals and beliefs in terms of declarative statements.\(^{23}\) Declarative statements have well defined meanings and are either true or false. Kowalski mentions the following examples of beliefs and actions.\(^{24}\)

1. Belief: You defend yourself, if whenever someone attacks you, you attack them back;
2. Goal: You defend yourself.

The general pattern of beliefs and goals represented in BDI, has the form $\beta(a, \sigma)$ and $\gamma(a, \sigma)$, respectively. $\beta(a, \sigma)$ expresses that the agent $a$ believes $\sigma$, whereas $\gamma(a, \sigma)$ expresses that $a$ desires $\sigma$. In both cases $\sigma$ is a declarative statement.

Another way of representing goals are production systems. Whereas BDI logic represents goals declaratively, production-systems can represent them procedurally. Production

\(^{23}\) (Kowalski 2003, p. 8)
\(^{24}\) (Kowalski 2003, p. 9)
systems consist of a set of condition-action rules and a set of statements called the working memory. Each condition-action rule has the form: If condition $x$, then do actions $y_1, \ldots, y_n$. Kowalski mentions the following example of a condition-action rule:\textsuperscript{25}

If someone attacks you, then attack them back.

Kowalski intends to unify the declarative and the procedural approach to represent goals. Since the hierarchical relations among goals are of interest, Kowalski uses goal hierarchies to represent goals. Goal hierarchies can either be illustrated graphically by AND-OR trees or linguistically by goal-reduction procedures.

Goal-reduction procedures are instructions consisting of one statement of a goal and a finite set of conditions or sub-goals. Conditions and sub-goals are connected by one of the two connectives AND or OR. A set of goal-reduction procedures makes up an AND-OR tree that represents a goal hierarchy.\textsuperscript{26} Kowalski gives several examples of goal-reduction procedures as well as AND-OR trees. Among his examples is the following AND-OR tree.\textsuperscript{27}

![Figure 3: AND-OR Tree](image_url)

\textsuperscript{25} (Kowalski 2003, p. 6)  
\textsuperscript{26} (Kowalski 2003, p. 12)  
\textsuperscript{27} (Kowalski 2003, p. 10)
This AND-OR tree illustrates the hierarchical relations between goals and sub-goals. The tree shows, for instance, that the goal of improving your enjoyment of life can be achieved by the sub-goal of improving your standard of living which, in turn, can be achieved by earning more money. This tree can be expressed by the following four goal-reduction procedures.

1. To improve enjoyment of life, increase standard of living or work less hard;
2. To increase standard of living, increase pay;
3. To increase pay, go on strike or increase productivity;
4. To provide for old age, increase pay and save money.  

Kowalski presents his unified logic-based agent model as a sequence of five steps: (1) cycle, (2) observe, (3) think, (4) decide (what actions to perform), and (5) act. In step (2), beliefs and goals, represented by goal-reduction procedures and BDI expressions, respectively, are read. Step (3) involves both forward reasoning, i.e. deducing goals from sub-goals, and backward reasoning, i.e. inducing sub-goals from goals. In step (4), a calculation based on decision-theory decides which actions to perform. Finally, the chosen actions are carried out (5).

In our view, Kowalski’s approach, as presented in his paper, does not provide a suitable model for conflict resolution. It remains unclear how the different computational tools can co-operate with each other and together resolve a specific conflict. Also, the process of identifying higher-level goals is not specified, i.e. it is not clear how higher-level goals can be found at all.

---

28 (Kowalski 2003, p. 12)
1.2.4 Lewis Richardson’s Equational Model

Modelling conflicts in terms of mathematical equations is the basic idea of Lewis Richardson’s contributions to the theory of conflict. In contrast to the game theoretical model, according to which conflicts are finite sets of states, Richardson aims at describing conflicts in terms of differential equations. Each equation expresses a relationship between certain conflict elements, such as the extent of accumulated grievance, the tendency to react strongly or weakly to another party’s threat, or the degree of difficulty to produce arms. Once a list of equations is specified, they can be used to derive general conclusions about the general dynamics of conflicts or to make predictions about the development of specific conflicts.

Richardson’s research on conflict was mainly developed during World War II. Observing how the accumulation of arms and mutually perceived threat had led to large scale war, Richardson focused on concepts, such as the level of armament, perceived threat, and the extent of accumulated grievance in his conflict analysis. Later, his models turned out to be particularly applicable to the Cold War situation and triggered a research tradition of scientific conflict analysis including scholars such as Quincy Wright, Kenneth Boulding, and Anatol Rapaport.

Richardson was also the first conflict scholar who collected a large amount of empirical data on the topic. His dataset included statistical data on 108 wars in the period between 1820 and 1949 which he used to test the conclusions made on the basis of his equations and to refine his models. His results related to the statistical analysis of conflict were published posthumously in the book *Statistics of deadly quarrels.*

---

29 (Richardson 1960)

We illustrate Richardson’s theory with his equational model for arms races. Richardson’s aim was to provide equations expressing the underlying regularities of an arms race, i.e. he tried to model the dynamics of armament and disarmament. To achieve this task, he followed a strategy including the following six steps: First, he identified the key concepts relevant for describing arms races. Second, he assigned variables and constants to the concepts. Third, he formulated equations expressing common sense assumptions about the relationships among the variables and constants. Fourth, he derived conclusions from the equations. Fifth, he tested the conclusions against statistical data on actual arms races. Sixth, he adjusted the equations according to the data. We illustrate the first four steps of Richardson’s strategy.

In step one, Richardson identifies four concepts relevant for modelling arms races: the armament level of the states involved in the arms race, the tendency of a state to react strongly or weakly to the threat by other states, the tendency to be strongly or weakly influenced by the difficulties of producing arms, and the extent of accumulated grievances. As a fifth factor, Richardson includes time in his model as he intends to model the development of armament levels over time. The model is tailored to two states x and y. However, it could easily be extended to further states.

In step two, he assigns variables and constants to the five concepts.

(1) \( t \): variable for time;

---

30 (Richardson et al. 1960)
31 The model was fully developed in (Richardson et al. 1960). We use a summary of the model presented in (Schellenberg 1996, p. 103ff).
(2) $x/y$: variable for the armament level of state $x/y$;

(3) $a_{x/y}$: constant for the tendency of state $x/y$ to react strongly or weakly to the threat by state $y/x$;

(4) $m_{x/y}$: constant for the tendency of state $x/y$ to be strongly or weakly influenced by the difficulties of producing arms;

(5) $g_{x/y}$: constant for the extent of accumulated grievances of state $x/y$.

As his model tracks changes of armament levels over time, these two concepts are expressed by variables. The factors influencing the two variables, i.e. the other three concepts, are assumed to be constant.

In step three, Richardson makes three assumptions about the relationships between the identified concepts. We are not assessing the plausibility of his assumptions here as our aim is to illustrate his method rather than to evaluate the content of his assumptions.

(1) The change of the armament level of a state $x$ is proportional to the threat by other states due to their armament level, i.e. the higher the armament level of another state, the higher the armament level of state $x$. This relationship is mediated by the tendency of state $x$ to react strongly or weakly to the threat by other states.

(2) The change of the armament level of a state $x$ is proportional to the tendency of state $x$ to be strongly or weakly influenced by the difficulties of producing arms, i.e. the increase in armament level of a state $x$ is constrained by how easy it is for $x$ to produce arms.

(3) The extent of accumulated grievances of state $x$ positively influences its armament level, i.e. the more grievances $x$ has accumulated, the faster it raises its armament level.
Translating the assumptions into equations, Richardson comes up with two differential equations, one for each state. The equations have on their left a term expressing the change of the armament level over time, and on their right a term combining the three factors: tendency to react strongly or weakly to the threat by other states, tendency to be strongly or weakly influenced by the difficulties of producing arms, and extent of accumulated grievances. The change in armament level is expressed by the first derivative of x with respect to time, i.e. $dx/dt$. The factors on the right are combined by summation, where the tendency to be strongly or weakly influenced by the difficulties of producing arms is negatively taken into account as it lowers the rate of armament according to Richardson’s assumptions. The two constants for the factors: tendency to react strongly or weakly to the threat by other states and tendency to be strongly or weakly influenced by the difficulties of producing arms are multiplied by the armament levels $x$ and $y$, respectively, as the threat perceived by state $x$ is proportional to the armament level of state $y$ and the difficulty of producing arms directly influences the armament level. Altogether, Richardson suggests the following two differential equations.

(I)  \[ \frac{dx}{dt} = a_x y - m_x x + g_x; \]

(II) \[ \frac{dy}{dt} = a_y x - m_y y + g_y. \]

From these two equations, Richardson deduces several conclusions. He derives the conclusions by manipulating the equations and distinguishing various situations characterised by specific combinations of the six constants $a_x$, $a_y$, $m_x$, $m_y$, $g_x$ and $g_y$. In particular, Richardson identifies four prototypical situations characterised by specific conditions for the constants.

(1) If $m_x + m_y > a_x + a_y$, i.e. if the braking factors $(m_x, m_y)$ together outweigh the combined strength of reactivity $(a_x, a_y)$, the states reach a stable equilibrium.
(2) If $m_x + m_y > a_x + a_y$ and $g_x < 0$ and $g_y < 0$, i.e. if the braking factors together outweigh the combined strength of reactivity and both states start with an accumulation of goodwill ($-g_x$ and $-g_y$), the states move to disarmament.

(3) If $m_x + m_y < a_x + a_y$ and $g_x > 0$ and $g_y > 0$, i.e. if the combined strength of reactivity outweighs the braking factors and both states start with an accumulation of grievances, the states engage in an arms race leading to war.

(4) If $m_x + m_y < a_x + a_y$ and $g_x < 0$ and $g_y < 0$, i.e. if the combined strength of reactivity outweighs the braking factors and both states start with an accumulation of goodwill, the situation is indeterminate, i.e. either the states engage in an arms race leading to war or move to disarmament.

The conclusions allow one to make predictions about the outcome of a situation on the basis of data about the braking factors, the strength of reactivity, and the extent of accumulated grievances of the states involved.

1.3 Psychological Theories

Psychological theories of conflict aim at explaining conflicts by looking at characteristics of individual people and the interactions between them. For a monograph on psychological theories of conflict, see (Shalit 1988). A comprehensive anthology on the topic is, for instance, (Fuchs and Sommer 2004).

Psychological theories of conflict aim at explaining conflicts by looking at characteristics of individual people and the interactions between them. First, we look at theories of aggression. We introduce three theories of aggression reflecting the nature-nurture controversy on the topic. We start with Konrad Lorenz’s ethological two-factor theory of aggression, continue with John Dollard’s frustration-aggression hypothesis, and end with Albert Bandura’s social learning theory of aggression.

---

32 For a monograph on psychological theories of conflict, see (Shalit 1988). A comprehensive anthology on the topic is, for instance, (Fuchs and Sommer 2004).
Second, we discuss Muzafer Sherif’s realistic conflict theory, a psychological theory explicitly designed to explain social conflicts.

Finally, we introduce the theory of social dilemmas. The concept of social dilemmas is used to describe paradigmatic conflict situations. We first describe the prisoner’s dilemma, the most widely discussed dilemma in conflict studies, and then the dilemma of the commons which is often seen as a prototype of resource conflicts. Both dilemmas can be understood as specific types of games as defined in game theory.

1.3.1 Theories of Aggression

In his book, *On Aggression*, Konrad Lorenz postulated the existence of a human fighting instinct. For Lorenz, aggressive behaviour is natural and inherent to human nature. He assumes that a pattern of aggressive behaviour is genetically encoded in most animals, including human beings, because such behaviour is evolutionarily advantageous for the species. Referring to a large number of observational animal studies, he concludes that aggressive behaviour is a survival-enhancing instinct.

Lorenz’s two-factor theory of aggression does not imply that there is a constant pressure to behave aggressively or destructively. Aggressive behaviour only occurs when triggered by specific stimuli of the environment called releasers. Thus, the first factor for aggressive behaviour is the innate urge to aggress, and the second factor is the appropriate stimulation by an environmental releaser.

According to Lorenz, the innate fighting instinct is not learned ontogenetically, but has evolved over the human phylogeny. Aggressive behaviour between members of different

---

33 (Lorenz 1971)
species is an evolutionary advantage as it results in an increased supply of resources for the species.

However, the question arises why members of the same species fight each other. Lorenz gives three reasons why intra-specific aggression can also be survival-enhancing to the whole species. First, aggressive behaviour results in greater distribution of the population. If a member of a population defends its resources against another member, the latter will be forced to look for new territory. Second, fights within a population are a means of natural selection: if only the strongest members of a population are selected to survive, the whole species has an evolutionary advantage. He writes,

“\textquote {I return to the theme of the survival value of the rival fight, with the statement that this only leads to useful selection where it breeds fighters fitted for combat with extra-specific enemies as well as for intra-specific duels.}”\textsuperscript{34}

Third, intra-specific fights have the consequence that the offspring is defended by the most “aggressive family defender”.\textsuperscript{35}

When explaining human wars, which are not survival-enhancing, Lorenz identifies two dysfunctions of the fighting instinct. Unlike other animals, humans are badly equipped with natural weapons, such as sharp teeth, poison etc. To compensate for this disadvantage, humans have developed technical weapons aimed at killing others. As a consequence, human fights are more lethal than rival fighting of other species. Furthermore, technical weapons can be used to kill from a long distance. As there is no physical contact between the fighters, they are not able to detect their enemies’ appeasement gestures. Indeed, most species have developed patterns of behaviour indicating surrender or retreat. If this behaviour is shown during a fight, the winner has

\textsuperscript{34} Cf. (Lorenz 1971, p. 40).
\textsuperscript{35} (Lorenz 1971)
no incentive to continue the fight and kill as he has reached his goal. Lorenz argues that, as a consequence of distant fighting, humans have lost their appeasement gestures.

Next, we turn to the frustration-aggression hypothesis, a position which emphasises both the biological and social causes of aggressive behaviour. It was formulated by a group of scholars around John Dollard at Yale University in the 1930s.\textsuperscript{36} In its initial version, the hypothesis claimed that

\begin{quote}
"the occurrence of aggressive behavior always presupposes the existence of frustration and, contrariwise, the existence of frustration always leads to some form of aggression."
\end{quote}

Frustration is understood as the unexpected failure of reaching a goal that one had expected to reach.

Although there is some empirical support of the frustration-aggression hypothesis, its universality was questioned by subsequent scholars. Counterexamples can be found in both directions, i.e. frustration does not always result in aggressive behaviour, and aggressive behaviour is not only the result of frustration. For instance, Martin Seligman showed in a number of studies that frustration can also lead to a feeling of helplessness and resignation, instead of aggression.\textsuperscript{38} On the other hand, soldiers or policemen often behave aggressively in order to obey an order or to prevent other people from aggressive behaviour, rather than because of frustration.

A generalisation of the frustration-aggression hypothesis was proposed by Leonard Berkowitz.\textsuperscript{39} In his theory of aversively generated aggression, frustration is just one example of the more general phenomenon of aversive events. An aversive event is a

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{36} (Dollard 1964)
\item \textsuperscript{37} (Dollard 1964, p. 1)
\item \textsuperscript{38} (Seligman 1975)
\item \textsuperscript{39} (Berkowitz 1962)
\end{itemize}
\end{footnotesize}
situation that is unpleasant and, therefore, avoided. According to Berkowitz, aversive events are the only source of aggressive behaviour. However, aggression is only one possible reaction to an aversive event.

According to Berkowitz, aversive events generate negative feelings, certain behavioural tendencies, and certain thoughts. These reactions make up three different types of responds to aversive events. First, the reactions help to deal with the aversive event rationally and calmly. Second, they make a person escape from the situation. Third, they make a person behave aggressively. Which of these reaction patterns is predominant is determined by genetic influences, the perceived effectiveness of the reaction in the particular situation, and the past learning history of the individual involved.\(^{40}\)

Berkowitz’ theory has been tested in various empirical studies. Most famous are his experiments in which he asked subjects to administer punishments to other subjects, while holding a hand in a tank of water. When the water was painfully cold, test persons gave significantly more punishments than when the water in the tank was warm.

Another generalisation of the frustration-aggression hypothesis is the concept of relative deprivation. Particularly, aggressive behaviour on the macro level, such as collective violence, can be explained by relative deprivation. Relative deprivation is defined as the “[s]ense of having less than that to which one feels entitled”\(^{41}\). Berkowitz argues that relative deprivation leads to individual frustration which, when amplified by an aversive event, leads to aggressive behaviour.

Finally, we present social learning theory as a theory explaining aggressive behaviour. Representing the nurture-end, social learning theory claims that aggressive behaviour is

\(^{40}\) (Berkowitz 1962)

\(^{41}\) (Hogg and Vaughan 2002, p. 288)
learned, i.e. it is not innate to humans but the result of past experiences. This theory, which is closely related to the behaviourist school in psychology, provides a general explanation of why people acquire, instigate, and maintain certain patterns of aggressive behaviour.

Albert Bandura explains the occurrence of a person’s aggressive behaviour by four factors: the person’s past experiences of aggressive behaviour, the success of the person’s aggressive behaviour in the past, the perceived likelihood of the aggression being rewarded or punished, and a complex of other cognitive, social, and environmental factors.42

The learning process involves two types of learning: operant conditioning and observational learning. In operant conditioning, a person experiences either a positive or a negative respond to his behaviour. If the respond is positive, the likelihood that the person shows the behaviour again rises. If the respond is negative, it falls. So, if a person is rewarded for aggressive behaviour, it will become more likely for the person to behave aggressively again. The reward can either be an additional pleasant stimulus (positive operant conditioning) or the removal of an unpleasant stimulus (negative operant conditioning) Observational learning is the process of imitating observed behaviour. For instance, if a child watches somebody else behaving aggressively, it will imitate this behaviour and behave aggressively itself. This effect is particularly strong when the child observes that the observed person is rewarded for his aggressive behaviour. In a number of studies, Bandura found strong empirical support for this effect. In his experiments, children observed somebody punching a doll with a hammer. Afterwards they were left to play with the doll. Most children showed a significantly higher number of aggressive

42 (Bandura 1978)
acts towards the doll than children of a control group who did not observe the violent model.\textsuperscript{43}

1.3.2 Muzafer Sherif’s Realistic Conflict Theory

A social psychological theory of conflict has been developed by Muzafer Sherif in the 1950s and 1960s.\textsuperscript{44} Emphasising scarce resources as the main cause of conflict, he called his theory realistic conflict theory. His theory is based on a series of experiments known as the Robbers Cave Experiment.\textsuperscript{45} Before formulating Sherif’s theory, we describe his experiments. They are classics in psychological conflict studies and provide the empirical basis for his theory.

Sherif observed the behaviour of 20 boys during a boys scout camp. The camp, which was run by experimenters, consisted of four phases. For each of the four phases, Sherif changed certain conditions and observed how these changes influenced the quality of the boys’ interpersonal and intergroup relations.

In the first phase, the boys arrived at the camp and experienced ordinary, everyday activities. Over time, various friendships formed among the boys. In the second phase, the camp was split into two completely isolated groups. Members of the two groups lived in different living quarters, and there was no contact between them at any time of the day. The result was that formerly established friendships deteriorated, and own group norms developed. Some boys expressed a view in which they scaled others with reference to their group and considered their own group as superior. In the third phase, the groups

\textsuperscript{43} (Bandura et al. 1961)
\textsuperscript{44} (Sherif and Sherif 1953)
\textsuperscript{45} (Sherif et al. 1961)
were brought together again, but every contact was organised as an intergroup competition, such as a sports contest. In any competition, only one group could win an attractive price. This led to fierce competition during, and outside the organised competitions. The boys made discriminatory remarks towards members of the other group, engaged in various forms of aggressive behaviour, and the general hostility between the two groups increased significantly. At the same time, solidarity within the two groups increased. In the fourth phase, Sherif exposed the boys to goals which they could only achieve through co-operation. One such superordinate goal was, for instance, to pull out a bogged down truck bringing a movie. All the boys wanted to watch the movie, but the truck could only be pulled by the two groups together. The experimenters observed an improvement in the intergroup relations as a result of the co-operation.

Based on his experiments, Sherif posited a causal relation between individual goals, group goals, interpersonal behaviour, and intergroup behaviour. He further classified goals as either shared goals, i.e. goals that can only be achieved through co-operation, or exclusive goals, i.e. goals that can only be achieved by one individual or group. According to realistic conflict theory, shared individual goals lead to interpersonal co-operation, group formation and solidarity. Shared group goals lead to intergroup co-operation and intergroup harmony. Individual exclusive goals lead to interpersonal competition and conflict, reduced solidarity, and group collapse, whereas exclusive group goals lead to intergroup competition and conflict.

Sherif’s theory was confirmed in various studies. However, critics have pointed out that goals are not the only variables determining interpersonal and intergroup behaviour.

---

46 See, for instance, a study investigating intergroup relations between tribal groups in Africa was carried out by Marilynn Brewer and Donald Campbell. See (Brewer and Campbell 1976).
1.3.3 The Theory of Social Dilemmas

The theory of social dilemmas aims at transforming abstract models of game theory into real world scenarios which can then be tested empirically. If people behave in accordance with game theoretic models, so the idea goes, conflict behaviour can be predicted and explained by these models.

The most famous paradigmatic conflict scenario is the prisoner’s dilemma devised by Merill Flood and Melvin Dresher in 1950, and so named by Albert Tucker. Since then, it has become an influential research paradigm in the social sciences. It can be described as a game between two players, 1 and 2, where both players have two strategies, c_i and d_i, called co-operation and defection, respectively. If both players choose c_i, they both receive the same utility payoff \( u_{1/2}(c_1, c_2) \). If they both defect, they receive the payoff \( u_{1/2}(d_1, d_2) \). However, \( u_{1/2}(c_1, c_2) \) is higher than \( u_{1/2}(d_1, d_2) \). If one of the players defects and the other co-operates, the defector receives a higher utility payoff than the co-operator, i.e. \( u_i(d_i, c_j) > u_j(d_i, c_j) \). Indeed, \( u_i(d_i, c_j) \) is the highest utility payoff and \( u_i(d_i, c_j) \) the lowest, i.e. \( u_i(d_i, c_j) > u_{1/2}(c_1, c_2) > u_{1/2}(d_1, d_2) > u_i(d_i, c_j) \), with \( i, j \in \{1, 2\} \).

The strategic form of a prisoner’s dilemma, with \( u_i(d_i, c_j) = 0, u_{1/2}(c_1, c_2) = -1, u_{1/2}(d_1, d_2) = -5, \) and \( u_i(d_i, c_j) = -10, \) is shown in the following table.

<table>
<thead>
<tr>
<th></th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>c_2</td>
</tr>
<tr>
<td>1</td>
<td>-1, -1</td>
</tr>
<tr>
<td></td>
<td>0, -10</td>
</tr>
</tbody>
</table>

Table 10: Prisoner’s Dilemma

The dilemma arises since defecting strongly dominates co-operating for both players. So, if both players behave rationally, they end up with the utility payoffs \( u_i(d_1, d_2) \) and \( u_2(d_1, d_2) \).
Theories of Conflict

d_j), although they could be better off with the outcomes \(u_i(c_1, c_j)\) and \(u_2(c_1,c_2)\), which they could receive by co-operating. As the only Nash equilibrium of the game is the strategy profile \((d_1, d_2)\), which is not Pareto-optimal, the predicted outcome of the game is not the best outcome for them.

The name of the game stems from a story used to illustrate the dilemma: Two prisoners have the two options of either confessing to a crime jointly committed, or not to confess to the crime. If they both confess, they will be sent to prison for five years. If neither confesses, they will be charged with a minor crime and sent to prison for only one year. If one of them confesses and the other does not, the confessing prisoner will be released as a principal witness, whereas the other will be sentenced for ten years. If the prisoners are aware of these conditions, they rationally both confess the crime and are sentenced for five years.

Test persons faced with variations of the prisoner’s dilemma, where the outcomes are money or other incentives, overwhelmingly opt for defection. This could be shown in a large number of studies with people from various countries.\(^{47}\)

The dilemma of the commons, which can be used to model conflicts over limited resources, is an n-person prisoner’s dilemmas, i.e. an n-person game in which an optimal solution can be reached if all players co-operate. If every player defects, the outcome is worst for all. However, any individual player can increase his payoff by defecting, given that all other players co-operate. Formally, an n-person prisoner’s dilemma can be defined as a game with a set of players \(N = \{1, 2, ..., n\}\) in which each player \(i\) has the two strategies \(c_i\) or \(d_i\), and for each player’s utility function, the following conditions hold.

\(^{47}\) For an overview of these studies up to 1980, see (Dawes 1980).
Chapter 1  Reflecting on Conflicts

(1) \( u_i( c_1, \ldots, c_{i-1}, d_i, c_{i+1}, \ldots, c_n) \) is the highest utility value;
(2) \( u_i( x_1, \ldots, x_{i-1}, c_i, x_{i+1}, \ldots, x_n) < u_i( x_1, \ldots, x_{i-1}, d_i, x_{i+1}, \ldots, x_n), \) where \( x \in \{ c, d \} \);
(3) \( u_i( d_1, \ldots, d_{i-1}, c_i, d_{i+1}, \ldots, d_n) \) is the lowest utility value;
(4) \( u_i( x_1, \ldots, x_{i-1}, c_i/d_i, x_{i+1}, \ldots, x_n) < u_i( y_1, \ldots, y_{i-1}, c_i/d_i, y_{i+1}, \ldots, y_n) \)
   if \( \{ x_1, \ldots, x_n \} \) contains more ds than \( \{ y_1, \ldots, y_n \} \).

Hence, the more players that defect, the lower the utility payoff is for each individual player. The dilemma arises as, for every player, defecting is the better option compared to co-operating, and, therefore, they all defect and end up with the worst outcome.

This model provides an explanation for conflicts about resources, such as natural resources, or public goods. It was originally used by Garrett Hardin to explain a conflict about common pastures typical for English villages. This is why it is called the tragedy of the commons or the commons dilemma. The following description of the dilemma is cited from Hardin's original 1968 paper published in *Science*.

"The tragedy of the commons develops in this way. Picture a pasture open to all. It is to be expected that each herdsman will try to keep as many cattle as possible on the commons. Such an arrangement may work reasonably satisfactorily for centuries because tribal wars, poaching, and disease keep the numbers of both man and beast well below the carrying capacity of the land. Finally, however, comes the day of reckoning, that is, the day when the long-desired goal of social stability becomes a reality. At this point, the inherent logic of the commons remorselessly generates tragedy. As a rational being, each herdsman seeks to maximize his gain. Explicitly or implicitly, more or less consciously, he asks, "What is the utility to me of adding one more animal to my herd?" This utility has one negative and one positive component.

1. The positive component is a function of the increment of one animal. Since the herdsman receives all the proceeds from the sale of the additional animal, the positive utility is nearly + 1.

2. The negative component is a function of the additional overgrazing created by one more animal. Since, however, the effects of overgrazing are shared by all the herdsmen, the negative utility for any particular decision-making herdsman is only a fraction of - 1.

Adding together the component partial utilities, the rational herdsman concludes that the only sensible course for him to pursue is to add another animal to his herd. And another .... But this is the conclusion reached by each and every rational herdsman sharing a commons. Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit -- in a world that is limited. Ruin is the
destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons. Freedom in a commons brings ruin to all."\textsuperscript{48}

Experimental studies simulating the commons dilemma show ambiguous results. For instance, Kaori Sato showed that most subjects defect, and, therefore, commonly destroy the resource.\textsuperscript{49} In other studies, the subjects voluntarily co-operated.\textsuperscript{50}

1.4 Political Theories

International relations scholars have developed theories about the interactions between states. These theories provide explanations for inter-state conflicts, as well as conflicts in general when their conceptual framework is applied to units other than states.\textsuperscript{51} Moreover, because an increasing number of violent conflicts occur within states, specific political theories have been developed to understand intra-state conflicts, such as civil wars, revolutions or ethnic wars.\textsuperscript{52}

We illustrate two political approaches to the theory of conflict. Both approaches originate in the theory of international relations and are primarily intended to explain conflicts among states. First, we introduce a number of theories usually classified under the name of liberalism. In particular, we follow the historical development of liberalist theories starting with classical liberalism and idealism, and continue with neoliberal institutionalism, democratic peace theory, and international liberalism. Second, we look at

\begin{itemize}
\item \textsuperscript{48} (Hardin 1969)
\item \textsuperscript{49} (Sato 1988)
\item \textsuperscript{50} (Caporael et al. 1989)
\item \textsuperscript{51} For political theories on inter-state conflicts, see, for instance, (van Creveld 1991).
\item \textsuperscript{52} For an overview over political theories of civil war, see (Waldmann 2003). A recent socio-political theory of revolution was presented in (Goldstone 2001). For a monograph on ethnic wars, see (Kaufman 2001).
\end{itemize}
Chapter 1 Reflecting on Conflicts

theories developed within the realist paradigm. We start with classical realism and follow with expositions of the two successors of realism: neorealism and neoclassical realism.

1.4.1 Liberalism and Neoliberal Institutionalism

The roots of liberalism can be found in the eighteenth-century Enlightenment movement. Ideas of economic and political liberty in the nineteenth century, and Woodrow Wilson’s idealism of the early twentieth-century, have further contributed to the theory. Post-war liberalism as a general theoretical perspective includes neoliberal institutionalism, democratic peace theory, and international liberalism.

The French Enlightenment philosopher Baron de Montesquieu argues that human nature is generally good. As humans are rational beings, so his argument goes, they are able to understand the universal laws of nature and human society. In virtue of this ability, humans can live together in a peaceful and just society. However, when a person enters civil society, the person might be confronted with a corrupt environment which makes the person behave irrationally. Thus, Montesquieu concludes that it is the society which turns rational human beings into irrational ones, and that conflict is always the product of inadequate social institutions. He emphasises the role of education in the prevention of conflicts.

Immanuel Kant, one of the most comprehensive philosophers of the Enlightenment era, also argues that humans are primarily rational beings. He advocates the idea that nations can overcome difficulties and resolve their conflicts by means of collective action. His vision is a federation of sovereign democratic states that are able to settle their conflicts without war, and, hence, allow their citizens a peaceful life. In Perpetual Peace, Kant argues
for a federal order that structures the relations between states and mirrors the rule of law
which each individual sovereign state is bound by.\textsuperscript{53}

In the nineteenth century liberal thought was prominent both in the political and
economic domain. In \textit{On Liberty} John Stuart Mill highlights liberty as the main condition
for humans to achieve happiness. He argues that each individual knows best what is best
for him and how to achieve this.\textsuperscript{54} Rational human beings should, therefore, be left as
autonomous and free as possible in their pursuit of happiness.

Political liberalism claims that liberal democracies are the best form of government as
social institutions of a liberal democracy serve only the one goal of furthering the ends of
their citizens. Social institutions, such as laws, only exist as a result of free individuals
who decide to abide by them if it furthers their goals. Liberal democracies also do not
favour certain citizens over others, and, hence, guarantee a just society without social
classes. As states only exist for their citizens, individuals must be protected against
excessive and illegitimate government interferences.

Economic liberalism propagates the idea that government interference in economic
matters should be minimal both with regard to the domestic and the international
market. This implies that there should neither be trade barriers nor governmental
subsidies or monopolies. The liberalist claim that market forces automatically bring about
an optimal situation for each individual is expressed by Adam Smith’s famous metaphor
of the invisible hand.

\textsuperscript{53} (Kant 1795(2005))
\textsuperscript{54} (Mill 1859(1975))
With regard to conflicts, 19th century liberalism claims that conflicts originate in the restriction of individual freedom and they can only be avoided if individuals possess a maximal amount of freedom.

In reaction to the excessive human suffering caused by World War I, American president Woodrow Wilson published a list of Fourteen Points for Peace, which led to the creation of the League of Nations. Wilson’s writings express the idea that war is preventable, and that international peace is the ultimate goal nations should strive for. Wilsonian idealism introduces the concept of collective security in international relations theory. Collective security describes the principle that war can be prevented by a multilateral institution consisting of sovereign nations that agree to punish each member that behaves aggressively within the international arena. Nations have little incentive to engage in violent actions or war if they know they are to be punished by a large number of other nations. According to the idealists, conflicts can and should be resolved by various means of collective problem solving, such as mediation, arbitration or judicial settlement. They advocate for the formation of strong international institutions and believe that international law is effective in dealing with conflicts. Idealism also recognizes the role of disarmament in preventing international conflicts from escalating.

Liberal international relations theory was comprehensively reformulated under the name of neoliberal institutionalism in the 1970s by the American political scientist Robert Keohane. Neoliberal institutionalists combine 19th century liberalism and Wilsonian idealism with the principles of game theory.

---

55 The text of the original speech can be found in (Wilson 1918).
56 One of the seminal works of neoliberal institutionalism is Robert Keohane’s book After Hegemony: Cooperation and Discord in the World Political Economy (Keohane 1984).
Neoliberal institutionalism assumes that there are two types of actors critical to the quality of international relations. These are states and non-state groups, such as nongovernmental organisations, multinational corporations or international organisations. States and non-state groups can mutually influence one another and thereby determine the character of international relations.

Actors are assumed to be rational to the extent that they make decisions by evaluating all potential actions relative to their goals. In contrast to classical liberalists, who assume that human beings are naturally good and co-operation is innate to the human species, neoliberalists hold that humans are only interested in pursuing their own goals, and co-operation is the product of rational behaviour.

A further neoliberal assumption is that in pursuing their goals, actors depend on each other, i.e. goals can be achieved best if all actors co-operate with each other. As a result of the inclusive character of goals, the predominant feature of international relations is interdependence.

Finally, neoliberal institutionalism claims that, as a consequence of international interdependency, actors build international institutions because they make it easier for them to co-operate, and, hence, to achieve their goals. International institutions start to develop their own dynamic, i.e. they develop their own rules and norms which, in turn, influence the behaviour of the original actors. Institutions are defined broadly. They can be international organisations, such as the UN, international regimes, i.e. clusters of norms and rules regulating the actors’ behaviour with regard to a certain problem, international networks, i.e. clusters of procedural rules for a certain problem area, or international organising principles which are not limited to a certain problem area but regulate international relations as a whole.
Neoliberal institutionalists stress the repetitive character of games.\textsuperscript{57} If actors can interact with each other only once, the most rational option might be defection. However, neoliberal institutionalists argue that in international relations, actors have the chance to interact with each other repetitively. This makes it more rational for them to co-operate as they can react to the other player’s behaviour. Indeed, as political scientist and game theorist Robert Axelrod’s results seem to show, in a repeated prisoner’s dilemma, the best strategy is to co-operate as long as the opponent co-operates too, and to punish the opponent’s defection by defection.\textsuperscript{58} When dealing with conflicts, rational states and non-state actors rationally co-operate with each other and, therefore, avoid self-damaging behaviour such as wars or arms races.

After the end of the Cold War, liberal international relations theorists focused on two research questions. The first question concerns the relationship between democracies and peace and establishes a research tradition known as democratic peace theory. The second question is based on the assumption that interstate wars are less frequent than they used to be and asks why this is the case. This research tradition has been characterised as international liberalism.\textsuperscript{59}

The basic hypothesis of democratic peace theory states that liberal democracies do not go to war with one another. Conversely, if the hypothesis is true, the main cause of war is the absence of democracy, i.e. some sort of authoritarianism. A strength of democratic peace theory is its empirical character. The Correlates of War project at the University of Michigan, one of the largest databases collecting data on various aspects of wars since

\textsuperscript{57} Iterated prisoner’s dilemmas were originally studied by Robert Axelrod (Axelrod 1984).
\textsuperscript{58} This strategy is called TIT FOR TAT. See, for instance, (Milinski 1987).
\textsuperscript{59} (Mingst and Karns 1995)
Theories of Conflict

1816, has been used to both confirm and disconfirm the hypothesis. Indeed, as Edward Mansfield and Jack Snyder remark:

“[m]uch of the research on the democratic peace has relied on statistical tests, which indicate that democracies become involved in war as frequently as other states, but that by reasonably restrictive definitions, they have never fought each other.”

Although not bound to liberalism, democratic peace theory is usually supported by liberalists. Liberalism provides an explanation for the assumption of democratic peace theory as it identifies causal mechanisms that link democracies to peace. Democratic norms restrain their leaders’ power of decision-making to the extent that they can only choose actions that are in the interest of the majority of their citizens. War, so the argument runs, is never an optimal option for the citizens of a state because of its high costs in terms of human losses, destruction of infrastructure, expensive armament, etc. Hence, democracies do not go to war because their leaders’ can only choose options in the interest of their citizens, and war is never in the interest of their citizens. In addition to this general argument, there is a second argument linking democracy and peace. Democracies are bound by international institutions which advocate norms that ban war as a means of conflict resolution. As democracies, in contrast to authoritarian regimes, obey these norms, democracy is both a necessary and sufficient condition for not engaging in war.

Democratic peace theory has been criticised mainly by scholars of the realist tradition. They question the statistical significance of the findings that support democratic peace theory. Considering all combinatorial possibilities of relations between states, it could be only a coincidence that wars occur less frequently between democracies. John Owen points out that

60 (Brown et al. 1996, p. 304)
“the lack of wars among democracies, even if true, is not surprising. Wars are so rare that random chance could account for the democratic peace, much as it could account for an absence of war among, say, states whose names begin with the letter K.”

The writings of the American political philosophers Francis Fukuyama and John Mueller initiated a controversy about the future of political conflicts. According to them, wars are becoming less likely as states tend to consider them morally unacceptable. They argue that the spread of Western democracy guarantees that there will be fewer wars in the future. This claim is supported by the fact that Western democracies have not fought against each other since the end of World War II. In Retreat from Doomsday: The Obsolescence of Major War, Mueller argues that the inhuman experience of World War I and II has made it morally unacceptable for Western democracies to engage in war again.

Democracy, according to this view, is closely linked to the principles of free market economies. In fact, some definitions of democracy involve the concept of market or private property economics.

There are several points of criticism that have been raised against international liberalism. Common to all of them is the doubt that the idea of Western democracy is universally applicable as assumed by international liberalists.

1.4.2 Realism, Neorealism, and Neoclassical Realism

Realism and its successors, neorealism and neoclassical realism, make up the second school of thought in international relations theory. Realistic thinking is opposed to idealism and institutionalism, and is seen as rather pessimistic with regard to international

---

61 (Owen 1994, p. 88)
62 (Fukuyama 1992)
63 (Mueller 1989)
64 Cf., for instance, (Doyle 1983).
conflicts compared to idealism. Realism dates back to the ancient Greek historian Thucydides who formulated the fundamental assumptions of the realistic view. Further realist thinkers include the Christian bishop Augustine and the English philosopher Thomas Hobbes. In the 20th century realism became the prominent theory of international relations after World War II and influenced both the American and the European foreign policy of the Cold War era. After the Cold War, political scientists, such as Kenneth Waltz, simplified the realist assumptions and eliminated their essentialist prerequisites. This simplification of the classical approach has become known as the neorealist or structuralist school of realism. In the 1990s, a third generation of realists intended to unify the classical and the neorealist views under the new theoretical framework of neoclassical realism.

Classical realism originates in Thucydides’s *History of the Peloponnesian War*. In his analysis Thucydides formulates four assumptions that are still shared by contemporary realists. First, he identifies states as the main actors within the international system. Other actors, such as individuals or groups of interest, are assumed to have little control over a state’s relation to other states. Second, Thucydides believes that states act unitarily, i.e. they speak with one voice. Third, he claims that states act rationally. Every state weighs the costs and benefits of all available actions and chooses the one that best suits its national interest. Fourth, Thucydides specifies what it is for a state to pursue its national interest by identifying national interest with security. The most important goal of a states’ foreign policy is its protection against enemies. State behaviour is not determined by moral principles, but by the goal to increase security. According to Thucydides, this can be achieved by increasing domestic capacities, by increasing economic power, or by building alliances with other states.

65 (Thucydides BC431(1972))
St. Augustine provides an explanation of Thucydides’s assumption that states act egoistically by postulating that human nature itself is egoistic and selfish and, consequently, states are egoistic and selfish.

This link between state behaviour and human nature is further elaborated by Thomas Hobbes. He argues that states have the right to struggle for their survival similar to individuals who have the right to defend themselves and protect their own safety if left alone in a war of all against all. His main contribution to realism is his observation that the international system is anarchic. Lacking any higher entity that determines and controls the behaviour of states within the international system, each state is forced to regulate its interactions with other states on its own. Its potential actions are determined by its national capacities. Powerful states are more likely to survive than weaker states as they have more control over their external relations. From the fact that the international system is anarchic, Hobbes concludes that states have the right to protect themselves against other states even if this implies war; thus, conflicts and wars are inevitable.

In 1948, the German immigrant Hans Morgenthau wrote Politics among Nations, the first comprehensive standard volume of the realist approach. Morgenthau, as Hobbes, assumes that international politics is a struggle for power among sovereign nations. This struggle for power manifests itself on three different levels. On the individual level, people struggle for their survival. On the state level, unitary and independent states act in a way that maximises their national interests. On the international system level, anarchy is the predominant condition making it impossible for states to end their constant struggle for power.

---

66 (Hobbes 1651(2003))
67 (Morgenthau 1967)
Morgenthau identifies various options for states to increase their power. States can aggressively try to amass resources by means of an expansive foreign policy. Following this strategy, a state has to spend a large amount of his resources on building a strong attacking force. Offensive realism is the view according to which states naturally aim at expanding their territory and are only limited by their amount of military resources. In contrast to offensive realists, defensive realists, such as Stephen Walt, assume that states aim at protecting their own territory by investing in a strong defensive armament. A third option for states to gain power is to build military alliances with other states.

According to Morgenthau, the first two options, i.e. investing in a strong offensive or defensive force, lead to a security dilemma. A security dilemma evolves if all states try to increase their security through increasing their military power. As a whole, this leads to a more insecure situation than before. The fact that security gains of one state imply security losses for another can be described by a zero-sum game in which the players have the option to increase their military expenditure or not to increase it. If one state opts for increasing and the other does not, the security gained by the first player is lost by the second player. John Herz describes the security dilemma as follows:

“Striving to attain security from attack, [states] are driven to acquire more and more power in order to escape the power of others. This in turn renders the others more insecure and compels them to prepare for the worst. Since none can ever feel entirely secure in such a world of competing units, power competition ensues, and the vicious circle of security and power accumulation is on.”

If states choose the third option of building alliances with other states, the realist concept of a balance of power applies. It refers to an equilibrium state in which all blocks of allied states have an equal amount of power, and, consequently, no state has disproportionately

---

68 (Mearsheimer 2001)
69 (Walt 2005)
70 (Herz 1950, p. 157)
more power than any other state. In order to avoid war, the foreign policy of states should be driven by the goal of balancing power by supporting allied states and containing enemies. This concept, which assumes that state relations are exclusively determined by their relative military and economic power, has shaped the foreign policy of the Cold War era. Various examples can be found, in which either the USA or the Soviet Union supported allied states in order to weaken the other superpower’s power. Realists frequently argue that nuclear war has only been prevented by a balance of power within the international system.

Several assumptions of Morgenthau’s realism have been criticised. The main point of criticism concerns the lack of empirical testability of the theory and the connected problem of falsifiability. One of the critics, the American political scientist Kenneth Waltz, developed a simplification of classical realism known as neorealism. In his 1979 book *Theory of International Politics*, Waltz rejects the realist assumption that bad human nature is the cause of international anarchy. He also weakened the realist focus on states as the only actors within the international system. Although he recognises the dominant role of states, he points out that there are actors above and below the state level that influence state behaviour. He shares the realist view that the only goal of a state is to increase its power in order to survive within international anarchy. However, states are restrained in their actions by the structure of the international system.

According to Waltz, the international system is ordered by the specific distribution of capabilities among states. As individual states are not able to control the structure of the international system, or to change their position within it, their actions are restricted by the system. For instance, the options available to a superpower and those available to a

---

71 (Waltz 1979)
developing country differ significantly, just because they have a fundamentally different position within the international system.

Depending on their position within the international system, it can be disadvantageous for states to co-operate with other states for two reasons. First, even if both states are able to increase their power through co-operation, the question of whether this is advantageous for both of them depends on how much each state gains from the co-operation. If one state’s power increases much more than the other state’s power, the latter has a loss in relative power. What counts, according to the neorealist view, is relative power rather than absolute power. Second, co-operation makes states dependent on one another, and, hence, it restrains their potential actions. Therefore, too much dependence lowers a state’s power. As a result of these negative effects of co-operation, states distrust one another, and, generally, avoid co-operating. Weak states have two strategies according to the neorealist view. When confronted with a stronger state’s power, they can either join alliances with other weak states against the stronger state, or they can accept the stronger state’s influence and try to gain from supporting it. The former strategy is called balancing, the latter bandwagoning. The decision of which strategy is to be chosen is a cost-benefit calculation.

Similar to the realist view, neorealists also are pessimistic with regard to the occurrence of violent conflicts. In an atmosphere of mutual distrust and competition for relative power, violent conflicts are inevitable.

Neoclassical realists, the third generation of realists, aim at reunifying classical realism and neorealism. They add domestic policy as an intervening variable between the international system level and the specific form of foreign policy chosen by states. Political scientists like Gideon Rose points out that the foreign policy of a state is
determined not only by the actual amount of relative power the state has within the international system, but also by the national perception of the system and the specific national incentives of the state.²² Hence, in order to predict a state’s foreign policy, we first have to look at the distribution of power, and second, we have to consider its perception of the international system and its domestic interests. These factors together determine, for instance, whether a state goes to war or not, according to neoclassical realism.

1.5 Theories about Attitudes and Social Groups

Extracting the basic elements of conflicts from the various formal, psychological, and political theories introduced in this chapter, we can identify entities such as goals, beliefs, emotions, and values as the main elements constituting conflicts. None of these elements is specific to conflicts, i.e. they all occur in other human situations. Hence, there are general theories dealing with these concepts. In this section we look at theories about some of the basic elements that constitute conflicts.

First, we provide an exposition of theories about attitudes. The reason for this is that crucial conflict elements, such as goals, beliefs, emotions, and values, can be understood as certain types of attitudes. Indeed, in our logic-based model, we will reconstruct these elements as propositional attitudes. First, we introduce the concept of attitudes as developed by Gordon Allport. Then, we look at Icek Ajzen and Martin Fishbein's theory of planned action which links attitudes and actions.

²² (Rose 1998)
Second, we address the question of how individual behaviour is linked to group behaviour. This is necessary as we want to extrapolate psychological characteristics, such as emotions, values, cognitions, etc. to social groups. Without this step, we would not have an explanatory basis for intergroup conflicts. We first describe the theory of minimal groups and introduce the concept of social norms. Then, we introduce social identity theory, the standard theory about the relationship between individual behaviour and group behaviour.

1.5.1 Theories about Attitudes

The concept of attitudes is a core concept of social psychology. A large number of studies in social psychology consist of measuring attitudes by means of questionnaires. The importance of studying attitudes is motivated by the fact that they play a major role in explaining human interactions.

The main function of attitudes is their ability to provide orientation towards objects for the person holding the attitude. An example illustrating this function is, for instance, the attitude that snakes are dangerous. Although not true for every snake, this attitude provides a guideline for people to deal with snakes.

Attitudes are not directly accessible, as they are cognitive structures encoded in the human brain. Gordon Allport, an early social psychologist, defined an attitude as

“[…] a mental and neural state of readiness, organised through experience, exerting a directive or dynamic influence upon the individual’s response to all objects and situations with which it is related.”  

73 (Allport 1935, p. 810)
Allport’s definition expresses the passive character of attitudes. They only become active when a person faces the object or situation, which the attitude is about. Then, the attitude determines the person’s reaction to the object or situation.

Leon Thurstone suggests that attitudes are holistic and consist of only one, predominantly affective component. He defines them as a person’s positive or negative affect towards a psychological object.\textsuperscript{74} Alice Eagly and Shelly Chaiken identify three components of an attitude: an affective component, a cognitive component, and a behavioural component.\textsuperscript{75} The affective component involves all emotional reactions a person has towards an object or situation. The cognitive component summarises thoughts, beliefs, judgements, opinions, schemes etc. the person holds towards the attitude object. The behavioural component can either be the person’s behavioural reaction to the attitude object or the person’s intention to react in a certain way, when facing the attitude object.

The dynamics of attitudes, i.e. their formation and change, is explained by a number of theories subsumed under the name “cognitive consistency theories”. The basic idea of these theories is that a person tries to keep his attitudes consistent with each other. So, if someone holds a positive attitude towards pacifism, for instance, and believes at the same time that a military intervention is justified, his attitudes are inconsistent or dissonant, as Leon Festinger calls it. As dissonance is subjectively aversive, the person tends to drop or at least to modify one of his attitudes. The two most prominent cognitive consistency

\textsuperscript{74} (Thurstone and Chave 1929)
\textsuperscript{75} (Eagly and Chaiken 1993)
Theories of Conflict

theories are Leon Festinger’s theory of cognitive dissonance and Fritz Heider’s balance theory.76

The standard theory about the relationship between attitudes and behaviour is Icek Ajzen and Martin Fishbein’s theory of reasoned action and its extension the theory of planned behaviour (TPB).77 According to TPB there are three variables, in addition to a person’s attitudes, that determine his behaviour towards an object. In order to explain, for instance, why somebody joins a paramilitary group, we have to refer to the person’s attitudes, his subjective norms, his perceived behavioural control, and his behavioural intention, according to TPB. The first condition is that his attitude towards joining the group is positive. Second, the person has to believe that relevant other persons in his environment appreciate his decision, i.e. that they believe joining the paramilitary group is a proper thing to do (subjective norm). Third, the person has to believe that he has control over the performance and consequences of his action. So, he has to believe that he has the required resources to join the paramilitary group, and that there will be an opportunity for him to do so (perceived behavioural control). If these three conditions are met, the person generates an intention to join the paramilitary group. If the intention is stable over a certain interval of time, he will finally carry out the action of joining the group. Predicting behavioural intentions and behaviour itself on the basis of TPB could be shown to be relatively reliable.78

76 (Festinger 1957; Heider 1958)
77 See (Ajzen and Fishbein 1980; Fishbein and Ajzen 1975).
78 See, for instance, (Madden et al. 1992).
1.5.2 Theories about Social Groups

The question of whether people behave differently as group members as opposed to individuals has been answered both positively and negatively. Connected to this question is the question of reductionism: Can group behaviour be completely reduced to the behaviour of its members or is there a fundamental difference between the behaviour of groups and the behaviour of individuals?

A simple but surprisingly robust theory about group behaviour, known as the minimal group paradigm, has been developed by Henri Tajfel. The theory shows that people engage in intergroup competition and ethnocentric behaviour as soon as they enter a group, even one arbitrarily formed. Both types of behaviour are examples of group behaviour, i.e. individuals show this kind of behaviour only as group members.

From the perspective of an individual, groups which the individual is a member of are called ingroups. All other groups are outgroups. Tajfel tried to state the minimal conditions for a person to show group behaviour. His studies have shown that a randomly chosen categorisation distinguishing an ingroup from an outgroup is sufficient to make people engage in intergroup competition and ethnocentric behaviour. In his experiment he randomly categorised subjects as X- or Y-group members. The subjects then had to distribute points between X- and Y-group members. Overwhelmingly, the subjects chose distributions that either maximised the number of points for members of their ingroup or maximised the difference in the number of points awarded in favour of the ingroup. The maximum ingroup profit strategy and the

---

79 For an overview article on the minimal group paradigm, see (Diehl 1990).
80 (Tajfel et al. 1971)
maximum difference strategy show that the subjects favoured members of their ingroup, although the subjects knew that the groups were formed randomly.

Tajfel’s result was supported in a number of subsequent studies. Michael Hogg and Graham Vaughan conclude from these experiments that

“[t]he robust finding, from scores of minimal group experiments conducted with a wide range of participants, is that the mere fact of being categorised as a group member seems to be necessary and sufficient to produce ethnocentrism and competitive intergroup behaviour”\(^{81}\)

In the following, we present the concept of social norms. Social norms are behavioural and attitudinal uniformities of a group, such as the behaviour of speaking Spanish typically associated with the group Mexicans or the expectation to wear a tie as a member of the group politicians. Different social norms define different social groups. They describe how members of a social group typically behave, and they prescribe how people are expected to behave if they intend to join a group.

Some social norms, such as the legal system of a country codifying the various norms about right and wrong behaviour held in that country, are explicit. Illegal behaviour is punished, whereas norm compliant behaviour is often rewarded. Other social norms are implicit. They only become explicit if they are violated by a group member.

The formation of social norms is often explained by majority influence. New group members conform to the norms of the majority. Majority influence and group conformity were studied in a famous experiment by Solomon Asch.\(^{82}\) He asked subjects to compare the lengths of three lines with the length of a reference line. The subjects had to tell which of the three lines was of equal length to the reference line. The tasks were

\(^{81}\) (Hogg and Vaughan 2002, p. 297)
\(^{82}\) (Asch 1951)
unambiguous, i.e. the subjects could find the correct answer straightforwardly. Nevertheless, if a number of other group members, who were confederates of the experimenter, consistently gave a wrong answer, a significant number of subjects also gave the wrong answer.

The minimal group paradigm is explained by, and integrated into, a larger theoretical framework called social identity theory.\(^8\) This theory, which was developed by Henri Tajfel and John Turner, aims at linking group behaviour on the macro level with processes on the individual level by means of the concept of social identity.

Its first assumption is that in any society, there are social groups, such as workers, men, tennis players, politicians, Hindu castes etc. Social identity can then be defined as the part of a person’s self-concept that derives from the person’s membership in a social group. A person, so the idea goes, has as many social identities as there are groups the person identifies with. For example, somebody’s social identity might consist of his identification as a man, a tennis player, a politician, a worker, and a member of a certain Hindu caste. By distinguishing between personal and social identity, social identity theorists avoid having to explain group behaviour merely in terms of personality and interpersonal relations.

Social identity theory claims that parts of a person’s behaviour can only be explained by taking into account the person’s social identity. This kind of behaviour can then be classified as group behaviour. Social identity theory identifies five characteristics of group behaviour. First, members of a group usually see the world through the lenses of this group. They evaluate their environment relative to the group’s preferences and, hence, engage in ethnocentric behaviour. Second, group members favour other ingroup

---

\(^8\) (Tajfel and Turner 1979; Tajfel and Turner 1986)
members over outgroup members. Therefore, ingroup favouritism is a characteristic of group behaviour. Third, people distinguish their ingroups from other groups. Doing this, they emphasise differences between their own group and other groups, i.e. they make intergroup differentiations. The fourth characteristic of group behaviour is the phenomenon of stereotyping which is connected to intergroup differentiations. The members of a group widely share simplified pictures of other social groups and their members as a means to differentiate between groups. The last feature of ingroup behaviour is that people usually tend to conform to ingroup norms.

Various intergroup processes on the macro level, such as social movements or societal changes, could be successfully explained by social identity theory.  

1.6 Summary

The aim of chapter 1 was to provide an overview of theories of conflict. This has been achieved to the extent that we have given descriptions of formal theories, psychological theories, and political theories of conflict. We have identified game theory as the most predominant formal theory of conflict and have seen how this theory reoccurs in psychological theories of conflict, such as the theory of social dilemmas, as well as political theories, such as neoliberal institutionalism or the security dilemma described by realists.

We have identified the concepts ‘agent’, ‘goal’, ‘attitude’, ‘value’, ‘strategy’, ‘utility’, ‘emotion’, and ‘probability’ as the concepts used by game theorists, psychologists, and political scientists to describe conflicts. Also, we have identified concepts like ‘goal’

---

84 (van Knippenberg and Ellemers 1993)

The theories give us a conceptual background for the development of our own logic-based approach. In particular, we will use the idea of tree structures, as defined in game theory, as the underlying temporal structures of conflicts in our model. Also, we will use the concepts ‘state’ and ‘agent’ from game theory and the concepts ‘goal’, ‘norm’, ‘belief’ and, ‘emotion’ from the psychological theories.

Having extracted the main concepts of existing conflict theories, we will develop a general language to describe conflicts in the following chapter.
CHAPTER 2

Describing Conflicts

The Syntax of Conflict Modelling Logic

2.1 Introduction

The aim of chapter 2 is to introduce the syntax of Conflict Modelling Logic (CML). As a logical system intended to model conflicts, CML consists of a syntax, a semantics, and a proof theory. The syntax of CML is constituted by a formal language, which can be used to describe conflicts.

In this chapter we characterise the language of CML, \( L_{\text{CML}} \), by defining the set of its primitive symbols and stating the formation rules for its formulae. This will allow us to express statements about conflicts in terms of formulae of the language of CML.

As a result of the chapter, we will be able to give answers to the following questions.

- Which primitive symbols are used in the language of CML to describe conflicts and what are their natural language interpretations?
- How can the primitive symbols of CML be used to construct formulae expressing statements about conflicts?
Chapter 2  Describing Conflicts

- How can CML-formulae be used to express specific aspects of temporal expressions, attitudes, beliefs, and emotions?
- How can elements of the Second Congo War be described by CML-formulae?

An outline of the chapter is given as follows: In section 2.2, we introduce the primitive symbols of the language of CML and provide natural language interpretations for them. In particular, we introduce symbols for propositions (2.2.1), logical connectives and quantifiers (2.2.2), alethic modalities (2.2.3), various constituents of time (2.2.4), beliefs (2.2.5), goals (2.2.6), norms (2.2.7), and emotions (2.2.8). In section 2.3, we define the set of CML-formulae. First, we state their formation rules (2.3.1), and then we provide natural language descriptions for some basic CML-formulae (2.3.2). In section 2.4, we show how a number of statements, typically occurring in discussions about time, knowledge and belief, and emotions, can be formalised in the language of CML. In particular, we formalise expressions for specific aspects of temporal statements (2.4.1), attitudes and beliefs (2.4.2), and emotions (2.4.3). Finally, we describe aspects of the Second Congo War in the language of CML in order to illustrate the language of CML.

2.2  Symbols

The language of CML, $L_{CML}$, is constituted by a set SYM of primitive symbols. SYM consists of three sets of constant symbols: p-CON, t-CON, and a-CON, two sets of variable symbols, t-VAR and a-VAR, an indexical symbol n, and a set of operator symbols, OPE.

**Definition 15  (Symbols of $L_{CML}$)**

$SYM = p\text{-}CON \cup t\text{-}CON \cup a\text{-}CON \cup t\text{-}VAR \cup a\text{-}VAR \cup \{n\} \cup OPE$ is the set of primitive symbols of $L_{CML}$, where:
The Syntax of Conflict Modelling Logic

(1) $p\text{-CON} = \{p_1, p_2, \ldots \}$ is a countable set of propositional constants;
(2) $t\text{-CON} = \{t_1, t_2, \ldots \}$ is a countable set of temporal constants;
(3) $a\text{-CON} = \{a_1, a_2, \ldots, a_n\}$ is a finite set of agent constants;
(4) $t\text{-VAR} = \{x_1, x_2, \ldots \}$ is a countable set of temporal variables;
(5) $a\text{-VAR} = \{y_1, y_2, \ldots \}$ is a countable set of agent variables;
(6) $OPE = \{\neg, \land, \forall, =, R, <, \Box, B, G, N, E\}$ is a set of operators.

Informally, elements of $p\text{-CON}$ refer to propositions, elements of $t\text{-CON}$ refer to specific time points, and elements of $a\text{-CON}$ refer to specific agents. Elements of $t\text{-VAR}$ range over time points, and elements of $a\text{-VAR}$ range over agents. The indexical symbol $n$ stands for “now”. $OPE$ includes the truth-functional connectives $\neg$ and $\land$, the universal quantifier $\forall$, the identity symbol $=$, the temporal realisation operator $R$, the symbol for the temporal precedence relation $<$, the alethic modal operator $\Box$, the belief operator $B$, the goal operator $G$, the norm operator $N$, and the emotion operator $E$.

We also use the letters $p$, $q$, and $r$ for propositional constants and the letters $t$, $x$, $a$, and $y$ for elements of $t\text{-CON}$, $t\text{-VAR}$, $a\text{-CON}$, and $a\text{-VAR}$, respectively. Note that the letter $x$ is reserved for temporal variables, whereas the letter $y$ is reserved for agent variables. Furthermore, we use $CON$ for the set of all temporal and agent constants, i.e. $CON = t\text{-CON} \cup a\text{-CON}$, and $VAR$ for the set of all variables, i.e. $VAR = t\text{-VAR} \cup a\text{-VAR}$.

As there are constants and variables for time points and agents, we can distinguish temporal terms from agent terms. The indexical symbol $n$ is not a term of $L_{CML}$. Terms are defined as follows.

**Definition 16 (Terms of $L_{CML}$)**

$TER = t\text{-TER} \cup a\text{-TER}$ is the set of terms of $L_{CML}$ where:

(1) $t\text{-TER} = t\text{-CON} \cup t\text{-VAR};$
(2) $a\text{-TER} = a\text{-CON} \cup a\text{-VAR}.$

Apart from their semantic category, elements of $SYM$ can also be classified according to the domain they are intended to model. This categorisation of $SYM$ will be used in
chapter 4 to single out various fragments of CML. The following definition distinguishes between modal, temporal, and attitude symbols.

**Definition 17 (Classification of SYM)**

\[ \text{SYM} = \square\text{-SYM} \cup \text{t-SYM} \cup \text{a-SYM} \]

is the set of symbols of \( L_{\text{CML}} \) where:

1. \( \square\text{-SYM} = p\text{-CON} \cup \{\neg, \land, \square\} \) is the set of modal symbols;
2. \( \text{t-SYM} = p\text{-CON} \cup p\text{-TER} \cup \{n\} \cup \{\neg, \land, \forall, =, R, <\} \) is the set of temporal symbols;
3. \( \text{a-SYM} = p\text{-CON} \cup p\text{-TER} \cup \{\neg, \land, \forall, =, B, G, N, E\} \) is the set of attitude symbols.

Note that the three categories are not mutually exclusive but jointly exhaustive.

**2.2.1 Propositional Constants**

Conflicts can be described by declarative sentences. In such a description, the sentences refer to propositions that are in some sense relevant to the conflict. For example, when describing the war in Bosnia and Herzegovina the sentence

“Bosnia and Herzegovina is a recognised, independent state”

refers to the proposition that Bosnia and Herzegovina is a recognised, independent state. This proposition is either true or false, depending on whether Bosnia and Herzegovina is, indeed, a recognised, independent state or not.\(^85\)

Following the example, a proposition can be defined as the entity expressed by a declarative sentence or the entity a declarative sentence refers to.\(^86\) If a proposition is true, it is also called a fact.

\(^85\) In fact, the sentence was false before April 6, 1992 and true afterwards.

\(^86\) This view of propositions as the sense or thought expressed by sentences is due to Gottlob Frege who distinguished in his classic paper *On Sense and Reference* between the sense and the reference of a sentence.
The example shows that propositions are the bearers of truth-values leading to another definition of propositions as the primary bearers of truth-values.\footnote{A historical overview over theories of propositions is given in (Nuchelmans 1973).} In CML a bivalent view is adopted with regard to propositions, i.e. every proposition must either have the truth-value ‘true’ or the truth-value ‘false’. Furthermore, it is excluded that a proposition is true and false at the same time, has no truth-value at all, or has a truth-value other than ‘true’ or ‘false’.

Propositions have the ability to occur as the content of propositional attitudes. Propositions can be believed, wanted, known, liked, brought about, morally claimed, etc. The basic structure of a propositional attitude is $A(a, p)$, where $A$ designates an attitude relation, such as believes-that, wants-that, knows-that, etc., the letter $a$ stands for the agent holding the propositional attitude, and $p$ is the proposition towards which $a$ holds the attitude.\footnote{(Salmon and Soames 1988)}

Propositional constants, contained in p-CON, refer to specific propositions relevant to the conflict that is modelled. The question whether a proposition is relevant to a conflict or not has only a vague answer. If a particular conflict is to be modelled, one can ask the parties, which propositions they consider relevant to the conflict. Alternatively, one can list all those propositions which are relevant to the conflict from an outside perspective, such as the perspective of an expert of the conflict or a research institute.\footnote{The two methods correspond to two views in conflict studies. According to the subjective view, conflicts are determined by what is conceived, valued, desired etc. by the parties. See, for instance, (Deutsch 1973, p. 11). Advocates of the objective view, such as Karl Marx or John Galtung, claim that conflicts exist regardless of what the conflicting parties think or feel. See (Galtung 1990, p. 10).}

---

\footnote{According to Frege, the sense of a sentence is the thought expressed by the sentence, whereas the referent of a sentence is its truth-value. See (Frege 1892(1980)).}
2.2.2 Logical Symbols

Logical symbols used in $L_{CML}$ are either truth-functional connectives, quantifiers, or the identity symbol. We use the classical negation symbol, $\neg$, and the classical conjunction symbol, $\land$, as a complete base of truth-functional connectives. Further connectives for classical disjunction ($\lor$), material implication ($\supset$), and the material bi-conditional ($\equiv$) are defined from the base $\{\neg, \land\}$. The truth-functional connectives are read in the usual way, i.e. $\neg$ as “it is not the case that”, $\land$ as “and”, $\lor$ as “or”, $\supset$ as “implies”, and $\equiv$ as “if and only if”. The identity symbol, $=$, connects terms of $L_{CML}$ or the indexical symbol $n$. We also use it to define the inequality symbol $\neq$.

The basic quantification symbol of $L_{CML}$ is the universal quantifier $\forall$ which is read as “for all”. We can quantify over both temporal and agent variables. Quantification leads to the definition of bound and free occurrences of variables in formulae. These definitions are given in the usual way, i.e. a variable $x$ or $y$ occurs bound in a formula just in case it occurs within the scope of a quantification with respect to $x$ or $y$, and its occurrence is free if it is not bound. Formulae in which all temporal and agent variables occur bound are called sentences of $L_{CML}$. The existential quantifier $\exists$, which is read as “for some”, is defined from $\forall$ in the usual way.

Defined logical symbols used in CML are listed in the definition below.

**Definition 18 (Defined Logical Symbols of $L_{CML}$)**

1. $\varphi \lor \psi$ for $\neg(\neg\varphi \land \neg\psi)$;
2. $\varphi \supset \psi$ for $\neg(\varphi \land \neg\psi)$;
3. $\varphi \equiv \psi$ for $\neg(\varphi \land \neg\psi) \land \neg(\psi \land \neg\varphi)$;
4. $\exists \nu \varphi$ for $\neg\forall \nu \neg \varphi$ where $\nu \in \text{VAR}$;
5. $(\tau_1 \neq \tau_2)$ for $\neg(\tau_1 = \tau_2)$ where $\tau_1, \tau_2 \in \text{TER} \cup \{n\}$. 

2.2.3 Modal Symbols

In order to express alethic modal properties of propositions, two unary operators are used in $L_{CML}$. The box symbol $\Box$ expresses necessity, whereas the diamond symbol $\Diamond$ expresses possibility. Hence, $\Box$ is read as “necessarily” and $\Diamond$ as “possibly”.

We only introduce the necessity operator $\Box$ as a primitive symbol and define the possibility operator $\Diamond$ in the usual way.

**Definition 19 (Defined Modal Symbol of $L_{CML}$)**

$\Diamond \varphi$ for $\neg \Box \neg \varphi$.

2.2.4 Temporal Symbols

Temporal symbols used in $L_{CML}$ are either temporal terms, the indexical symbol, or temporal operators. Temporal terms, summarised in the set $t\text{-TER}$, denote time points. The indexical symbol $n$ also denotes a time point. However, it is not classified as a term because its value is not determined by a term assignment function. The binary relation symbol $<$ expresses the temporal precedence relation between time points, and the binary operator $R$ is used to express that a formula is true at a certain time point.

The temporal symbols of $L_{CML}$ are based on the symbols used in Nicholas Rescher and Alastair Urquhart’s R/U-calculus.\(^90\) In contrast to the tense operators F and P of Arthur Prior’s tense logic,\(^91\) the operators of the R/U-calculus explicitly link time points to propositions. The temporal realisation operator $R$ allows one to build formulae that assert the truth of a sub-formula at a specific time point. The temporal precedence

\(^90\) See (Rescher and Urquhart 1971, p. 35).
\(^91\) See (Prior 1967).
relation U makes it possible to compare time points with regard to their chronological order. In the R/U-calculus, it is possible, for instance, to explicitly represent sentences like

“Bosnia and Herzegovina was recognised as an independent state on April 6, 1992”.

In contrast, if we use tense operators like F or P, we can only represent sentences like

“At some time in the past, Bosnia and Herzegovina was recognised as an independent state”.

The expressive power of the R/U-calculus is stronger than the expressive power of tense logic as the tense operators F, P, G, and H can be defined within the R/U-calculus. Therefore, every formula of tense logic can be translated into a formula of the R/U-calculus.92

When speaking about time points, it is possible to single out one preferred position: the position of “now”. Then, all other time points can be compared with this position and identified as past time points or future time points. However, as time passes, so does the position of “now”. In L_CML, we refer to the preferred position of “now” by the symbol n. This symbol is classified as an indexical symbol as it is neither a constant nor a variable. Rescher and Urquhart describe the syntactical representation of “now” as follows.

“The status of n = “now” is, of course somewhat peculiar from the logical standpoint. It is neither fish nor fowl, neither constant nor variable. It is not a constant, since it does not denote the same T-element [i.e. time point - F.I.] on each occasion of its use, and not a variable, since it cannot meaningfully be quantified […]. If we were to assign n to any semantical category, it would have to be that of indexicals, i.e., those parts of speech...

---

92 For the definability of tense operators in the R/U-calculus, see Rescher and Urquhart 1971, p. 50ff.
The Syntax of Conflict Modelling Logic

which denote something uniquely in a given context, though not the same thing in every context.\footnote{Rescher and Urquhart 1971, p. 35}

With the syntactical representation of “now”, it is possible to divide the temporal domain into three disjoint classes: past, present, and future. Thereby, we obtain descriptions of time in the style of John McTaggart’s A-series.\footnote{McTaggart 1908} Time can also be described in terms of series of earlier and later, which McTaggart called B-series. Here, the basic concept is the notion of temporal precedence expressed in $\mathcal{L}_{\text{CML}}$ by the binary temporal relation symbol $\prec$.\footnote{Rescher and Urquhart use the symbol U for temporal precedence.} The symbol $\prec$ stands between any two temporal terms, or between a temporal term and the indexical symbol $n$, and is read as “temporally precedes”. We use the infix notation for $\prec$.

In order to express that two time points temporally coincide, we use the identity symbol $=.$ In this context, $=$ is read as “temporally coincides with”. Similar to the precedence symbol, $=$ stands between temporal terms or a temporal term and the indexical symbol $n$.

In order to express that a proposition is true at a certain time point, we use the binary temporal realisation operator $R$. As propositions, expressed by formulae of $\mathcal{L}_{\text{CML}}$, can be true at any time point including “now”, $R$ links temporal terms or the indexical symbol $n$ with formulae of $\mathcal{L}_{\text{CML}}$. For the operator $R$, we use the prefix notation, i.e. statements like “$\varphi$ is realised at the time point $t$” or “now it is true that $\varphi$” are expressed by formulae of the form $Rt\varphi$ and $Rn\varphi$, respectively.

Further temporal ordering relations, such as temporal succession, weak temporal precedence, and weak temporal succession can be defined as follows.
Definition 20 (Defined Temporal Symbols of $L_{\text{CML}}$)

(1) $\tau_1 > \tau_2$ for $\tau_2 < \tau_1 \wedge \neg(\tau_1 = \tau_2)$ where $\tau_1, \tau_2 \in t\text{-TER} \cup \{n\}$;

(2) $\tau_1 \leq \tau_2$ for $\tau_1 < \tau_2 \vee \tau_1 = \tau_2$ where $\tau_1, \tau_2 \in t\text{-TER} \cup \{n\}$;

(3) $\tau_1 \geq \tau_2$ for $\neg(\tau_1 < \tau_2)$ where $\tau_1, \tau_2 \in t\text{-TER} \cup \{n\}$.

2.2.5 Doxastic Symbol

In $L_{\text{CML}}$, we use the doxastic operator $B$ to express that an agent believes that a certain proposition, expressed by a formula of $L_{\text{CML}}$, is true. The operator is read as “believes that”. $B$ is assumed to express a weak belief predicate, i.e. $B$ is neither related to the truth/falsity of the believed proposition nor to the agent’s reasons or justifications for the belief.\(^6\)

A formula of the form $B\alpha \varphi$, with $\alpha \in a\text{-TER}$, expresses that the agent referred to by $\alpha$ believes that $\varphi$ in the sense that the agent would answer affirmatively to the question “Is $\varphi$ true?”. If the agent were not asked, he would not be aware of the fact that he believes that $\varphi$. Hence, $B$ is a reactive and unconscious belief operator.

For instance, the statement

“Radovan Karadzic believes that the Republic of Srpska should have close relations with Serbia”

can be expressed by a formulae of the form $B\alpha p$, where the agent constant $\alpha$ refers to the agent Radovan Karadzic and $p$ to the proposition expressed by the sentence

“The Republic of Srpska should have close relations with Serbia”.

\(^6\) For a detailed analysis and classification of doxastic and epistemic predicates, see (Stelzner 1984).
2.2.6 Symbol for Goals

Goals pursued by the agents involved in a conflict are expressed in $L_{CML}$ by the goal operator $G$ which links agent terms to propositions. $G$ is read as “wants that” or “has the goal that”. $G$ expresses that an agent wants that a certain proposition, expressed by a formula $\varphi$ of $L_{CML}$, is true.

For example, the statement

“The citizens of Srebrenica want Srebrenica to be independent from the Republic of Srpska”

can be expressed by the formula $Gap$, where the agent constant $a$ refers to the agent “the citizens of Srebrenica” and $p$ to the proposition expressed by the sentence

“Srebrenica is independent from the Republic of Srpska”.

Again, $G$ is assumed to be a reactive and unconscious operator. The fact that an agent has a certain goal does not imply that the agent is aware his goal. Formulae of the form $G\alpha\varphi$ merely express that the agent $\alpha$ would answer affirmatively to the question “Do you want $\varphi$ to be true?”.

2.2.7 Deontic Symbol

Unlike deontic logic, CML does not deal with logical interdependencies of normative statements. Our only aim is to syntactically represent norms, and values held by agents involved in a conflict. This is reflected by the basic norm operator $N$ of $L_{CML}$. $N$ links agent terms to propositions and is read as “considers it a moral/legal norm that”.
N can be used to express moral or religious values. For instance, the fact that the third Pillar of Islam is considered a religious value by religious Bosnian Muslims, which is expressed by the sentence

“Religious Bosnian Muslims believe that one is morally obliged to spend 2.5% of his wealth for the benefit of the poor or needy”

can be modelled by the formula Nap, where the agent constant a refers to the agent “the religious Bosnian Muslims” and p is the proposition expressed by the sentence

“One spends 2.5% of his wealth for the benefit of the poor or needy”.

N can also be used to model legal norms held by a party. For instance, if we want to represent the fact that the European Union considers Article 44 of Protocol I, additional to the Geneva Conventions, a legal norm, which is expressed by the sentence

“The European Union thinks that it should be the case that any combatant who falls into the power of an adverse party while not engaged in an attack or in a military operation preparatory to an attack shall not forfeit his rights to be a combatant and a prisoner of war by virtue of his prior activities”

we can do this with the formula Nap where the agent constant a refers to the agent “European Union” and p to the proposition expressed by the sentence

“Any combatant who falls into the power of an adverse party while not engaged in an attack or in a military operation preparatory to an attack shall not forfeit his rights to be a combatant and a prisoner of war by virtue of his prior activities”.

The difference between the goal operator G and the norm operator N is a difference in the degree of generality. G only expresses that one particular agent wants a certain
proposition to be true, whereas \( N \) represents the case in which an agent claims that every agent is obliged to bring about the truth of a certain proposition.

2.2.8 Symbol for Emotions

Emotional excitement can be reconstructed as a propositional attitude.\(^9^7\) The emotional operator \( E \) expresses that an agent is in a state of emotional excitement, and that this state is caused by a certain proposition, expressed by a formula of \( L_{\text{CML}} \). The operator links agent terms to propositions and is read as “is in a state of emotional excitement caused by” or “is emotionally excited about”.

For example, the statement

“The citizens of Srebrenica are emotionally excited (anger) because Srebrenica belongs politically to the Republic of Srpska”

can be modelled by the formula \( E a p \) where the constant \( a \) refers to the agent “the citizens of Srebrenica” and \( p \) refers to the proposition expressed by the sentence

“The citizens of Srebrenica are emotionally excited (anger) because Srebrenica belongs politically to the Republic of Srpska”.

The polarity of the excitement, i.e. whether the excitement is positive or negative, as well as further parameters determining the specific character of the emotion, are not captured by the basic operator \( E \). They can, however, be defined by formulae that combine the emotional operator with other attitude operators like \( B \) or \( G \).

\(^9^7\) The view of emotions as certain types of propositional attitudes is closely related to a cluster of theories of emotion summarised as cognitive appraisal theories. Among the advocates of the cognitive appraisal approach to emotions are Richard Lazarus, Nico Frijda, Magda Arnold, and Rainer Reisenzein. See (Lazarus and Lazarus 1994; Frijda 1986; Frijda et al. 1989; Shields and Kappas 2006; Reisenzein and Hofmann 1993).
2.3 Formulae

Using the primitive symbols of $L_{CML}$, we can build CML-formulae. The set of all formulae of the language $L_{CML}$ is referred to by FOR and defined recursively by a definition in Backus Naur Form (BNF).\textsuperscript{98} The BNF makes it possible to decide for every string of elements of SYM, whether it represents a well formed CML-formula or not.

2.3.1 Formation Rules

Formulae of $L_{CML}$ are defined by the following BNF.

\begin{definition} (Formulae of $L_{CML}$)
FOR is the set of all strings of elements of SYM satisfying the following BNF:
\[ \phi ::= p | \neg \psi | \psi \land \chi | \forall \nu \phi | \tau_1 = \tau_2 | \Box \psi | R\tau_1 \phi | \tau_1 < \tau_2 | O\alpha \phi, \text{ where } p \in \text{p-CON}, \nu \in \text{VAR}, \tau_1, \tau_2 \in \text{TER} \cup \{n\}, \tau_3, \tau_4 \in \text{t-TER} \cup \{n\}, \alpha \in \text{a-TER}, O \in \{B, G, N, E\}. \]
\end{definition}

2.3.2 Interpretation of Basic Formulae

The following table displays all types of basic formulae of $L_{CML}$, shows how to read them, and gives an example for every formula.

\textsuperscript{98} BNF definitions were introduced by John Backus and Peter Naur. For the original paper, see (Naur and Backus 1964).
Table 1: Basic Formulae of CML

<table>
<thead>
<tr>
<th>Formula</th>
<th>How to read it</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>p</td>
<td>Srebrenica belongs to the Republic of Srpska.</td>
</tr>
<tr>
<td>¬p</td>
<td>it is not the case that p</td>
<td>Srebrenica does not belong to the Republic of Srpska.</td>
</tr>
<tr>
<td>p ∧ q</td>
<td>p and q</td>
<td>Srebrenica belongs to the Republic of Srpska and 30% of the citizens of Srebrenica are Bosnian Muslims</td>
</tr>
<tr>
<td>∀xRxp</td>
<td>for all time points Rxp</td>
<td>Srebrenica belongs always to the Republic of Srpska.</td>
</tr>
<tr>
<td>∀yByp</td>
<td>for all agents Byp</td>
<td>Everybody believes that Srebrenica belongs to the Republic of Srpska.</td>
</tr>
<tr>
<td>□(r ⊃ p)</td>
<td>it is necessarily the case that (r ⊃ p)</td>
<td>It is necessarily the case that if the Dayton Agreement is implemented, Srebrenica belongs to the Republic of Srpska.</td>
</tr>
<tr>
<td>n = t₁</td>
<td>n temporally coincides with t₁</td>
<td>Today is April 6, 1992.</td>
</tr>
<tr>
<td>t₁ &lt; t₂</td>
<td>t₁ temporally precedes t₂</td>
<td>April 6, 1992 is earlier than December 14, 1995.</td>
</tr>
<tr>
<td>Rt₂r</td>
<td>r is realised at the time point t₂</td>
<td>The Dayton Agreement was signed on December 14, 1995.</td>
</tr>
<tr>
<td>Ba₁s</td>
<td>a₁ believes that s</td>
<td>R. Karadzic believes that S. Milosevic is dead.</td>
</tr>
<tr>
<td>Ga₂¬p</td>
<td>a₂ wants that ¬p</td>
<td>The citizens of Srebrenica want Srebrenica to be independent from the Republic of Srpska.</td>
</tr>
<tr>
<td>Na₃u</td>
<td>a₃ considers it a moral/legal norm that u</td>
<td>The Serbian government considers it a legal norm that war criminals should be detained.</td>
</tr>
<tr>
<td>Ea₃p</td>
<td>a is emotionally excited because of p</td>
<td>The citizens of Srebrenica are emotionally excited because Srebrenica belongs to the Republic of Srpska.</td>
</tr>
</tbody>
</table>

2.4 Complex Expressions

In the following, we provide formalisations for some expressions typically occurring in the context of time, knowledge and belief, goals, norms, and emotions. The statements express a number of specific temporal relationships, conscious attitudes, true and false beliefs, and the polarity of emotions.
2.4.1 Complex Expressions for Specific Temporal Aspects

First, we look at formulae that express certain temporal aspects. If a statement $\varphi$ was true in the past, is true at present, or will be true in the future, we can express this by the following formulae, respectively.

$$\text{Past}(\varphi) = \forall x (x < n \supset Rx\varphi);$$
$$\text{Present}(\varphi) = Rx\varphi;$$
$$\text{Future}(\varphi) = \forall x (n < x \supset Rx\varphi).$$

If a statement $\varphi$ is true during a certain interval in time, this can be formalised by the following expression.

$$\text{Interval}(t_1, t_2, \varphi) = \forall x (t_1 \leq x \land x \leq t_2 \supset Rx\varphi).$$

Here $t_1$ is the starting point and $t_2$ is the endpoint of the interval.

If we want to express that a statement $\varphi$ will be true until a time point at which $\psi$ becomes true, we can do this by the following formula.

$$\text{Until}(\varphi, \psi) = \exists x_1 (n < x_1 \land Rx_1\psi \land \forall x_2 (n \leq x_2 \land x_2 < x_1 \supset Rx_2\varphi)).$$

Similarly, we can express that $\varphi$ has been true since a time point at which $\psi$ was true by the following formula.

$$\text{Since}(\varphi, \psi) = \exists x_1 (x_1 \leq n \land Rx_1\psi \land \forall x_2 (x_1 \leq x_2 \land x_2 \leq n \supset Rx_2\varphi))^{99}$$

---

$^{99}$ The “Since” and “Until” operators were introduced into temporal logic by Hans Kamp in his PhD thesis. See (Kamp 1968).
As the semantics of CML is based on a discrete model of time, we can formalise statements about immediate temporal predecessors and successors. For instance, if we want to say that \( \varphi \) is true at the next time point or that \( \varphi \) was true at the previous time point, we can do this by the following formulae.

\[
X(\varphi) = \exists x_1 (n < x_1 \land Rx_1 \varphi \land \neg \exists x_2 (n < x_2 \land x_2 < x_1));
\]

\[
Y(\varphi) = \exists x_1 (x_1 < n \land Rx_1 \varphi \land \neg \exists x_2 (x_1 < x_2 \land x_2 < n)).
\]

### 2.4.2 Complex Expressions for Conscious Attitudes and False Beliefs

Consciousness conditions for statements involving propositional attitudes can be expressed as conjunctions with the belief operator. We can contrast any unconscious belief, goal, norm, and emotion, with its conscious counterpart defined as follows.

\[
\text{Bconscious}(a, \varphi) = BaBa \varphi \land Ba \varphi;
\]

\[
\text{Gconscious}(a, \varphi) = BaGa \varphi \land Ga \varphi;
\]

\[
\text{Nconscious}(a, \varphi) = BaNa \varphi \land Na \varphi;
\]

\[
\text{Econscious}(a, \varphi) = BaEa \varphi \land Ea \varphi.
\]

The first expression, for instance, expresses that the agent a believes \( \varphi \) and believes that he believes that \( \varphi \), i.e. he is consciously aware of his belief.

False beliefs and true beliefs can be expressed by formulae of the form

\[
\text{Bfalse}(a, \varphi) = Ba \varphi \land \neg \varphi;
\]

\[
\text{Btrue}(a, \varphi) = Ba \varphi \land \varphi.
\]
2.4.3 Complex Expressions for Emotions

Finally, we provide two definitions which make it possible to distinguish between positive and negative emotions. Following the goal-relevance approach of emotions, an emotion is positive if its attitude object is wanted by the agent, and an emotion is negative if the attitude object is not wanted by the agent. A positive emotion is, therefore, defined by the following formula.

\[ E_{\text{positive}}(a, \varphi) = E_a \varphi \land G_a \varphi. \]

A negative emotion is defined by the formula

\[ E_{\text{negative}}(a, \varphi) = E_a \varphi \land G_a \neg \varphi. \]

In both cases, the agent is emotionally excited about \( \varphi \). \( \varphi \) is goal-congruent in the first case and goal-incongruent in the second case.

**Example: Second Congo War**

In this section, we illustrate how \( L_{\text{CML}} \) can be used to describe a conflict by applying it to the Second Congo War, which took place in the Democratic Republic of Congo (DRC) between 1998 and 2003. Obviously, this conflict is far too complex to be comprehensively represented here. We just focus on some aspects of the conflict. The choice of elements is motivated by the aim of illustrating the various operators of \( L_{\text{CML}} \).

---

100 According to the goal-relevance approach, which was developed in the cognitive appraisal paradigm of emotions, emotions only occur as reactions to events that are relevant for an agent. Goal congruent events cause positive emotions and goal incongruent emotions cause negative emotions. See (Lazarus and Lazarus 1994; Arnold and Gasson 1954, p. 294ff).

101 Comprehensive analyses on the War in the Democratic Republic of Congo can be found in (Weiss 2000; Nest et al. 2006; ICG 2008a; ICG 2008b).
The example is intended to show how $\mathcal{L}_{\text{CML}}$ works in general rather than to give a comprehensive representation of the Second Congo War.

During the First Congo War, from December 1996 to May 1997, an alliance of various rebel groups headed by Laurant-Desire Kabila, and supported by Rwanda, Uganda, and Angola, overthrew the Mobutu regime that had ruled the country for 31 years.\(^{102}\) In the Second Congo War, from August 1998 to July 2003, Kabila’s former allies Rwanda and Uganda turned against him and unsuccessfully tried to topple him as president. The Kabila government received support from Zimbabwe, Angola, and Namibia, as well as other African countries.\(^{103}\) In the course of the war, Rwanda and Uganda turned against each other and fought both direct battles as well as proxy battles through aligned rebel groups on the territory of the DRC. Altogether, the Second Congo War was a large scale war that cost approximately four million lives.\(^{104}\)

The factions in the Second Congo War can be grouped as follows: Kabila-aligned forces, Rwanda-aligned forces, Uganda-aligned forces, and Hutu-aligned forces.\(^{105}\)

| a\(_1\): Kabila-aligned forces |
| a\(_2\): Hutu-aligned forces |
| a\(_3\): Rwanda-aligned forces |
| a\(_4\): Uganda-aligned forces |

Each of these parties can be represented by an agent constant. The agent $a_1$ includes the Congolese national army under Kabila, his allies Zimbabwe, Namibia and Angola, and a number of Congolese nationalistic groups grouped under the name Mai-Mai. The agent

\(^{102}\) (ICG 2008a)  
\(^{103}\) (Clark 2002)  
\(^{104}\) (Coghlan et al. 2006)  
\(^{105}\) (ICG 2008b; ICG 2008a)
a₂ is mainly represented by the Hutu rebel organisation “Forces democratiques pour la Libération du Rwanda” (FDLR)¹⁰⁶, a group formed by former perpetrators of the 1994 genocide in Rwanda who flew into the neighbouring Congolese Kivu provinces after the Tutsi government under Kagame came into power in Rwanda. Furthermore, the agent includes Burundian rebels, Congolese Hutu, and aligned Mai-Mai groups. The agent a₃ includes the national armies of Rwanda and Burundi, and the rebel organisation “Rassemblement Congolais pour la Democratie” (RCD), formed by the Banyamulenge, the Congolese Tutsi people. The agent a₄ includes the national Ugandan army, and the rebel organisation “Mouvement pour la Liberation du Congo” (MLC).

The following list represents propositions relevant to the Second Congo War, i.e. issues the conflict was fought over.¹⁰⁷ Each of them is represented by a propositional constant.

<table>
<thead>
<tr>
<th>Proposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>p₁: The DRC is a uniform and sovereign state.</td>
</tr>
<tr>
<td>p₂: The national security of Rwanda is protected.</td>
</tr>
<tr>
<td>p₃: The national security of Uganda is protected.</td>
</tr>
<tr>
<td>p₄: Banyamulenge are attacked in the Kivu provinces.</td>
</tr>
<tr>
<td>p₅: Banyamulenge are eradicated.</td>
</tr>
<tr>
<td>p₆: Foreign Tutsi forces are expelled from the territory of the DRC.</td>
</tr>
<tr>
<td>p₇: The Tutsi government of Rwanda is overthrown.</td>
</tr>
<tr>
<td>p₈: The government of the DRC controls the exploitation of the natural resources in the Kivu provinces.</td>
</tr>
<tr>
<td>p₉: The government of Rwanda controls the exploitation of the natural resources in the Kivu provinces.</td>
</tr>
<tr>
<td>p₁₀: The government of Uganda controls the exploitation of the natural resources in the Kivu province.</td>
</tr>
<tr>
<td>p₁¹: The FDLR controls the exploitation of the natural resources in the Kivu provinces.</td>
</tr>
</tbody>
</table>

¹⁰⁶ The FDLR is still active in the Eastern Congo until now. Their website is (FDLR 2008). For a description of the rebel movement see (ICG 2005).

¹⁰⁷ The list represents issues identified in the Second Congo War by looking at annual reports of the Heidelberg Institute on International Conflict Research. See (HIIK 1999; HIIK 1998; HIIK 2000; HIIK 2001; HIIK 2002; HIIK 2003).
An attack is launched against the Kabila government.

Using the propositional attitude operators $B$, $G$, $N$, and $E$, we can connect the four agent constants with the 13 propositional constants and, thereby, describe some of the beliefs, goals, norms, and emotions held by the four factions. The following list contains a number of formulae and the respective statements that they express.

- $Ga_1p_1$: The Kabila-aligned forces want the DRC to be a uniform and sovereign state.
- $Ga_2p_2$: The Rwanda-aligned forces want the national security of Rwanda to be protected.
- $Ga_3p_3$: The Uganda-aligned forces want the national security of Uganda to be protected.
- $Ga_3\neg p_4$: The Rwanda-aligned forces want Banyamulenge not to be attacked in the Kivu provinces.
- $Ea_3p_4$: The Rwanda-aligned forces are emotionally affected by the fact that Banyamulenge are attacked in the Kivu provinces.
- $Na_3p_4$: The Rwanda-aligned forces consider it morally/legally wrong to attack Banyamulenge in the Kivu provinces.
- $Ba_3Ga_2p_5$: The Rwanda-aligned forces believe that the Hutu-aligned forces want to eradicate the Banyamulenge.
- $Ga_2p_6$: The Hutu-aligned forces want foreign Tutsi forces to be expelled from the territory of the DRC.
- $Ga_3\neg p_7$: The Rwanda-aligned forces want the Tutsi government of Rwanda not to be overthrown.
- $Ga_3(p_8 \land \neg p_9 \land \neg p_{10} \land \neg p_{11})$: The Kabila government wants the exploitation of the natural resources in the Kivu provinces to be exclusively controlled by the government of the DRC.
- $Ga_3p_{13}$: The Rwanda-aligned forces want an attack to be launched against the Kabila government.
- $Ba_3Ga_2p_5$: The Rwanda-aligned forces believe that the Hutu-aligned forces want the Banyamulenge in the Kivu provinces to be eradicated.
- $Ea_3Ga_2p_5$: The Rwanda-aligned forces are emotionally affected by the fact that the Hutu-aligned forces want the Banyamulenge in the Kivu provinces to be eradicated.

Using the possibility operator $◊$ and the conditional $⊢$, we can build more complex formulae expressing hypothetical and conditional statements about the conflict.
Chapter 2  Describing Conflicts

◊¬p₅: It is possible that the Banyamulenge in the Kivu provinces are not eradicated.
p₅ ⊃ p₁₃: If the Banyamulenge in the Kivu provinces are eradicated, then an attack is launched against the Kabila government.

If we want to describe the temporal component of the conflict we have to introduce temporal constants. For the sake of simplicity, we introduce here only one temporal constant \( t_1 \), standing for the date “August 3, 1998”.

\[ t_1: \text{August 3, 1998} \]

Using the realisation operator \( R \) we can temporally locate our statements. The following list displays three simple temporal statements and two more complex statements about the Second Congo War.

\[ R_{t_1}p_{13}: \text{On August 3, 1998 an attack is launched against the Kabila government.} \]
\[ R_{t_1}G_{a₃}p_{13}: \text{On August 3, 1998 the Rwanda-aligned forces want an attack to be launched against the Kabila government.} \]
\[ R_{t_1}E_{a₃}p_{13}: \text{On August 3, 1998 the Kabila government is emotionally affected by the fact that an attack is launched against them} \]
\[ B_{a₃}\exists x(n < x \land R_{x₃})$: The Kabila government believes that the DRC will be a uniform and sovereign state at some point in the future.
\[ B_{a₃}\forall x(n < x \land R_{x₃})$: The Kabila government believes that it is possible that the DRC will be a uniform and sovereign state at some point in the future.

The above statements, formalised in the language of CML, express various statements that can be made about the Second Congo War. The statements also show how different levels of particularity can be expressed depending on the level of complexity of the formulae representing them.
2.5 Summary

The aim of this chapter was to introduce the syntax of CML. This has been achieved to the extent that we have defined the set $\text{SYM}$ of primitive symbols of $L_{\text{CML}}$ and have stated the formation rules for the set $\text{FOR}$ of CML-formulae. $\text{SYM}$ includes symbols for important elements of a conflict. In particular, we have introduced symbols for agents, various temporal entities, propositions, modalities, beliefs, goals, norms, and emotions. With the BNF formation rules we have given a precise definition of CML-formulae.

Designing $L_{\text{CML}}$ was guided by two considerations. On the one hand, $L_{\text{CML}}$ had to be expressive enough to express all relevant aspects of a conflict. On the other hand, it had to be simple enough to be accessible to computational manipulations. The expressive power of $L_{\text{CML}}$ was illustrated by formalising statements about the Second Congo War.

Having introduced the syntax of CML, we are now able to describe conflicts in a uniform way by means of a formal language. The syntax of CML allows one to represent conflicts in terms of sets of CML-formulae. Finding a uniform and formal description is the first step in the process of modelling and resolving a conflict. In the following chapter, we will show how the formulae of the language of CML can be interpreted in a general semantics.
CHAPTER 3

Interpreting Conflicts

*The Semantics of Conflict Modelling Logic*

3.1 Introduction

The aim of chapter 3 is to introduce the semantics of CML. The semantics of CML provides the basis for interpreting CML-formulae and thereby makes it possible to interpret statements about conflicts. CML-formulae are evaluated in algebraic structures called CML-structures relative to conflict states.

In this chapter, we define CML-structures and state the conditions under which CML-formulae are true relative to a conflict state in a CML-structure. If we take the meaning of a statement to be the truth condition of the formula expressing the statement, then the semantics of CML provides a general theory of meaning for statements about conflicts. The concepts of validity, satisfiability, and semantic consequence, which are closely related to the truth conditions of CML, are defined at the end of the chapter.

As a result of the chapter we will to give answers to the following questions.

- What are the components of CML-structures?
Chapter 3 Interpreting Conflicts

- How can CML-formulae be evaluated in CML-structures?
- What are the truth conditions for the operators used in CML?
- What does it mean to say that a CML-formula is valid or satisfiable, and when does a CML-formula logically follow from other CML-formulae?
- How can aspects of the Second Congo War be reconstructed as a CML-structure?

An outline of the chapter is as follows: We start with a background section on branching-time temporal logic as the semantics of CML is developed in the context of these logics. In section 3.2, we introduce the components of CML-structures starting with the set of time points (3.2.1), followed by the set of conflict states and the set of conflict histories (3.2.2), the set of agents and attitude functions (3.2.3), and the term assignment function (3.2.4). In section 3.3, we define some further concepts: the immediate predecessor of a time point (3.3.1), the set of conflict states at a fixed time point (3.3.2), the time index function (3.3.3), and the accessibility relation (3.3.4). In section 3.4, we state the truth conditions for the logical symbols (3.4.1), the modal operator (3.4.2), the temporal symbols (3.4.3), and the propositional attitude operators (3.4.4). In section 3.5 we define three levels of validity of CML-formulae. First we define validity in states (3.5.1), followed by definitions of validity in CML-structures (3.5.2) and general validity in CML (3.5.3). Sections 3.6 and 3.7 deal with the concept of satisfiability and the semantic consequence relation, respectively. Finally, we apply the concept of CML-structures to the second Congo War.
Background: Branching-Time Temporal Logic

In this chapter, we introduce the concept of branching-time structures and review some of the logics describing them. This will give us a background for constructing our own temporal structure underlying the semantic of CML.

Branching-time temporal logics constitute a research field within temporal logic. Temporal logic aims at modelling processes of reasoning specific to the temporal components of statements. The motivation for creating branching-time temporal logics is the view that the past of any event is determined, whereas its future is indeterminate, i.e. the assumption that at any time point, the world has only one unique past history, but there may be many potential continuations into the future. As a consequence, statements referring to past time points have a fixed truth value, whereas statements referring to future time points have a fixed truth value only if the future continuation the world’s development will take is specified (Ockhamist view), or they are true or false in all potential future continuations (Peircean view).

In the following, we first introduce the concept of forwards-branching, backwards-linear, trees, the kind of structures underlying the semantics of branching-time temporal logics. Then, we describe three approaches to branching-time temporal logic. First, we introduce the modal approach to branching time temporal logic. This approach, associated with Prior’s tense logic, uses modal operators to express temporal aspects of statements. We will look at two modal systems: Ockhamist logic and Peircean logic.

---

108 For a general introduction to temporal logic, see, for instance, (Goldblatt 1992; Rescher and Urquhart 1971). For the history of temporal logic, see (Ohrstrom and Hasle 1995). A computer-science oriented introduction to temporal logic with a focus on state systems is (Kroeger and Merz 2008).

109 The distinction between the Peircean and Ockhamist view in branching-time temporal logic was introduced by Prior. See (Prior 1967).

110 For an overview of systems of tense logic, see (McArthur 1976).
Second, we discuss the Rt-approach to branching-time temporal logic. This approach is based on first-order logic rather than modal logic. As an example, we present Rescher’s system $K_b$. Finally, we introduce computational tree logic (CTL) and its extension full computational tree logic ($CTL^*$), the main system developed in the computing school of branching-time temporal logic.

Forwards-branching, backwards-linear, trees are structures of the form $(T, <)$ consisting of a set of states/time points $T$ and an irreflexive, transitive, left-linear binary relation $<$ on $T$, i.e. for all $t_0, t_1, t_2 \in T$, it is not the case that $t_0 < t_0$ (irreflexivity), if $t_0 < t_1$ and $t_1 < t_2$, then $t_0 < t_2$ (transitivity), and if $t_0 < t_2$ and $t_1 < t_2$, then either $t_0 < t_1$, $t_0 = t_1$, or $t_1 < t_0$ (left-linearity).\footnote{Goranko and Zanardo 2004, p. 2; Hodkinson and Reynolds 2007, p. 660; Gurevich and Saharon 1985, 668} Graphically, we can display forwards-branching, backwards-linear, trees as nodes connected by lines. The nodes represent the elements of $T$, and the lines the binary relation $<$. For a given forwards-branching, backwards-linear, tree $(T, <)$, a history $h$ is defined as a subset of $T$, which is linearly ordered by $<$ and maximal for inclusion. The set of all histories in a tree $(T, <)$ is denoted by $H(T)$. To complete the general semantic framework for branching-time temporal logics, we still have to define an evaluation function $v$. This is done by assigning truth values to propositional constants relative to

\[\text{Figure 4: Forwards-branching, Backwards-linear, Tree}\]
states/time points, i.e. \( v \) is a function \( v: \text{VAR} \times T \rightarrow \{0, 1\} \). A branching-time structure \( M \) can now be defined as a pair consisting of a forwards-branching, backwards-linear, tree \( (T, <) \) and an evaluation function \( v \), i.e. \( M = ((T, <), v) \).\(^{112}\)

Before we look at particular systems of branching-time temporal logic, we introduce the distinction between state-dependent and path-dependent formulae. A formula is path-dependent, or a path-formula, if its truth value can only be determined relative to a pair consisting of a state/time point \( t \) and a history \( h \). If the truth value of a formula only depends on the state/time point at which the formula is evaluated, it is a state-formula. The distinction between state-formulae and path-formulae applies only to logics in which formulae are evaluated relative to states and histories, otherwise, all formulae are automatically state-formulae.

In chapter seven of *Past, Present, Future*, Prior presents two systems of branching-time temporal logic, which he calls “Peircean logic” and “Ockhamist logic”\(^{113}\). Both systems are developed within the modal paradigm, i.e. temporal operators are understood as modal operators quantifying over possible worlds, which are called states or time points in the context of temporal logic. The difference between the two systems is that in Peircean logic, formulae are evaluated relative to states, whereas in Ockhamist logic formulae are evaluated with respect to states and histories.

Both systems have the same propositional language consisting of a set of propositional variables \( \text{VAR} = \{p, q, r, \ldots\} \), the usual Boolean connectives \( \neg, \lor, \land, \text{ and } \Rightarrow \), and two tense operators P and F. The two tense operators can be used to build complex formulae

\(^{112}\) See (Zanardo 1996, p. 4). An alternative, but essentially equivalent approach, to the semantics of branching-time temporal logic is to start with Kamp-frames, instead of trees, as primitive elements. In Kamp-frames, histories are introduced as primitive elements. See (Thomason 1984).

\(^{113}\) (Prior 1967, p. 113ff)
of the form $P\varphi$ and $F\varphi$ which are interpreted as “It has been the case that $\varphi$” and “It will be the case that $\varphi$”, respectively.

In Peircean logic, formulae are evaluated at states $t \in T$ of a branching-time structure $M = ((T, <), v)$ by extending the evaluation function $v$ to arbitrary formulae. The truth conditions for the Boolean connectives are just the standard ones. The following two conditions apply to the tense operators $P_{Peirce}$ and $F_{Peirce}$.

(1) \[ v_t(P_{Peirce} \varphi) = 1 \text{ iff for every history } h \in H(T), \text{ if } t \in h, \text{ then there is a } t' \in h \text{ such that } t' < t \text{ and } v_{t'}(\varphi) = 1; \]

(2) \[ v_t(F_{Peirce} \varphi) = 1 \text{ iff for every history } h \in H(T), \text{ if } t \in h, \text{ then there is a } t' \in h \text{ such that } t < t' \text{ and } v_{t'}(\varphi) = 1. \]

The two conditions express that a formula $\varphi$ “has been true” at a state $t$ iff in every history that goes through $t$, $\varphi$ is true at some past state $t'$ in that history, whereas a formula $\varphi$ “will be true” at $t$ iff in every history that goes through $t$, $\varphi$ is true at some future state $t'$ in that history. To give an example, we can look at the statement “It will rain”. In Peircean logic, the statement is true iff in every potential continuation of the world there is some point at which it will rain.

Validity for Peircean logic is defined in the usual way, i.e. a formula is valid iff it is true in every state $t \in T$ of every branching-time structure $M$. An adequate finite axiomatisation of Peircean logic has been provided by John Burgess. Alberto Zanardo has proved a completeness theorem for Peircean logic over an infinite axiom system.

---

114 (Burgess 1980)
115 (Zanardo 1990)
In Ockhamist logic, formulae are evaluated at pairs consisting of a state $t \in T$ and a history $h \in H(T)$ of a branching-time structure $M$. Again, the evaluation is done by extending the evaluation function $v$ to arbitrary formulae.

\begin{align}
(1) \quad v_{t,h}(P_{\text{Ockham}} \varphi) &= 1 \text{ iff there is a } t' \in h \text{ such that } t' < t \text{ and } v_{t'}(\varphi) = 1; \\
(2) \quad v_{t,h}(F_{\text{Ockham}} \varphi) &= 1 \text{ iff there is a } t' \in h \text{ such that } t < t' \text{ and } v_{t'}(\varphi) = 1.
\end{align}

The two truth conditions express that at a state $t$ a formula $\varphi$ “has been true” with respect to a history $h$ iff $\varphi$ is true at a past state $t'$ in $h$, whereas $\varphi$ “will be true” with respect to $h$ iff $\varphi$ is true at a future state $t'$ in $h$. According to Ockhamist logic, the statement “It will rain” is true relative to a specified history iff there is a future point in this specified history at which it will rain.

If a modal operator $\square$ of historical necessity is added to Ockhamist logic, it is possible to express the Peircean tense operators in terms of $\square$ and the Ockhamist operators. The truth condition for the historical necessity operator is given as follows.

\[ v_{t,h}(\square \varphi) = 1 \text{ iff for every history } h' \in H(T) \text{ if } t \in h', \text{ then } v_{t,h'}(\varphi) = 1. \]

Note that the truth value of $\square \varphi$ does not depend on the set of histories $H(T)$, i.e. $\square \varphi$ is a state formula. The Peircean tense operators, $P_{\text{Peirce}}$ and $F_{\text{Peirce}}$, can then be expressed as $\square P_{\text{Ockham}}$ and $\square F_{\text{Ockham}}$, respectively.

A formula is valid in Ockhamist logic iff it is true in every state history pair of every branching-time structure $M$. Adequate finite axiomatisations of Ockhamist logic are presented by Nicholas Rescher and Alastair Urquhart, as well as Robert McArthur.\textsuperscript{116}

\textsuperscript{116} (Rescher and Urquhart 1971; McArthur 1976)
In the Rt-approach to temporal logic, states or time points are explicitly represented by terms in the language. The basic operator of these systems is the temporal realisation operator R connecting terms for time points with formulae of the language. Formulae of the form $Rt\varphi$ are interpreted as “$\varphi$ is true at $t$”. If a binary relation $U$ for “earlier than” is added to the logic, we are able to directly express properties of the branching-time structure in terms of formulae of the language. Such an approach is chosen in Rescher and Urquhart’s R/U-calculus for branching-time temporal logic, which we present as an example for the Rt-approach.\(^{117}\)

The language of the R/U-calculus consists of a set of propositional variables $VAR = \{p, q, r, \ldots\}$, a set of temporal variables $TEM = \{x, y, z, \ldots\}$, the indexical symbol $n$ for “now”, the usual Boolean connectives, quantifiers, the identity symbol $=$, the temporal realisation operator $R$, and the temporal precedence symbol $U$. The set of formulae is given by the following BNF definition.

$$
\varphi ::= p \mid a = b \mid Uab \mid \forall x \psi \mid \neg \psi \mid \psi \land \chi \mid Ra\psi, \text{ where } a, b \in TEM \cup \{n\}.
$$

For a given branching-time structure $M$, formulae are evaluated at states/time points. However, as the language contains variables for states, the evaluation $v$ also depends on a term assignment function $\mu$ assigning elements of $T$ to the temporal variables $x, y, z$, etc.

We only provide the truth conditions for $R$ and $U$. Conditions for all the other symbols are standard.\(^{118}\)

\(1\) \quad \forall v, \mu \left( v_{\mu, \mu}(Rx\varphi) = 1 \text{ iff } v_{\mu(\mu), \mu}(\varphi) = 1; \right.

---

\(^{117}\) See (Rescher 1966; Rescher and Urquhart 1971). A more recent treatment of the approach can be found in (Åqvist 2005).

\(^{118}\) The conditions stated here are a slightly different version of the conditions truth conditions presented in (Rescher and Urquhart 1971, p. 45) adjusted to our notation.
Special truth conditions are required for the indexical symbol $n$, which is interpreted as the state $t$ at which the formula containing $n$ is evaluated. The conditions for $R$ express that a formula $R_x \varphi$ is true at a state $t$ if $\varphi$ is true at the $\mu$-value of $x$. A formula of the form $R_n \varphi$ is true at a state $t$ if $\varphi$ is true at $t$.

Rescher and Urquhart first provide an adequate axiomatisation for a basic system without the precedence relation $U$, which they call the basic system $R$ of temporal logic. Then, they add $U$ to the language of $R$ and formulate two conditions for $U$ expressing the transitivity and left-linearity of $<$. The obtained system of branching-time temporal logic is called $K_b$. In order to axiomatise $K_b$, they introduce the tense operators $P$ and $F$ into the system by means of the following definitions.

\begin{align*}
(1) \quad F \varphi & \text{ for } \exists x (U x \land R x \varphi); \\
(2) \quad P \varphi & \text{ for } \exists x (U x n \land R x \varphi).
\end{align*}

Using the tense operators $P$ and $F$, Rescher and Urquhart provide an axiomatisation of $K_b$ for which they prove completeness. Their system is equivalent to Nino Cocchiarella’s system $CR$ with an additional axiom expressing left-linearity.\footnote{Cocchiarella 1966; Prior 1967; McArthur 1976}

Computer scientists have developed branching-time temporal logics primarily to study the specification and verification of transition systems. The most elaborated such system is $CTL^*$, full computational tree logic. $CTL^*$ is an extension of the simple branching-time
temporal logic CTL, computational tree logic. CTL was developed by Edmund Clarke and Allen Emerson in the early 1980s\textsuperscript{120} and extended to CTL* by Emerson, Joseph Halpern, and Prasad Sistla.\textsuperscript{121} First, we describe the syntax and semantic of CTL*, then we single out CTL as a fragment of CTL*.

The language of CTL* consists of a set of propositional variables \( \text{VAR} = \{ p, q, r, \ldots \} \), the classical connectives \( \neg \) and \( \land \), and the temporal connectives \( X, U \) and \( E \). Formulae are defined by the following BNF.

\[
\varphi =:: p \mid \neg \varphi \mid \varphi \land \chi \mid X\varphi \mid \varphi U \chi \mid E\varphi.
\]

Further operators are defined as follows.

(1) \( \Diamond \varphi \) for \( (p \lor \neg p)U\varphi \);

(2) \( \Box \varphi \) for \( \neg ((p \lor \neg p)U\neg \varphi) \)

(3) \( A \varphi \) for \( \neg E \neg \varphi \)

Formulae of the form \( X\varphi \), \( \varphi U \chi \) and \( E\varphi \) are interpreted as “at the next state it is the case that \( \varphi \)”, “\( \varphi \) is the case until \( \chi \) is the case”, and “there is at least one path starting from the current state where it is the case that \( \varphi \)”, respectively.

The temporal operators used in CTL* are significantly different from the ones used in tense logic or Rescher’s Rt logic. The next operator \( X \) is particularly designed for discrete structures, such as transition systems. The until operator \( U \) is a binary operator, in contrast to the unary operators of tense logic or Rescher’s realisation operator \( R \).

\textsuperscript{120} (Clarke and Emerson 1981)
\textsuperscript{121} (Emerson and Sistla 1985)
Formulae of CTL* are evaluated in transitions systems. A transition system \( M \) is a structure of the form \((S, R, g)\) where \( S \) is a nonempty set of states, \( R \) is a binary, total relation on \( S \), i.e. for all \( s \in S \) there is an \( s' \) such that \( sRs' \), and \( g \) is a labelling of the states with sets of propositional variables, i.e. \( g \) is a function assigning subsets of \( \text{VAR} \) to the elements of \( S \).

Before we state the truth conditions for the temporal operators of CTL*, we still have to define the concept of a ‘fullpath’ in a transition system \( M \). If \( M \) is a transition system, by a fullpath \( b \) in \( M \) we mean an infinite sequence \((s_1, s_2, s_3, \ldots)\) of states of \( M \) such that for every \( s_i, s_{i+1} \in \mathbb{N} \) the relation \( s_iR s_{i+1} \) holds. For a given fullpath \( b = (s_1, s_2, s_3, \ldots) \) we write \( b_i \) to designate the state \( s_i \) in \( b \) and \( b_{\geq i} \) to designate the fullpath \((s_i, s_{i+1}, s_{i+2}, \ldots)\).

The truth value of formulae of CTL* is defined relative to a fullpath \( b \) of a transition system \( M \). We write \( M, b \models \varphi \) to express that \( \varphi \) is true at (the initial state \( b_1 \)) of the fullpath \( b \) in the transition system \( M \). The truth conditions for propositional variables and the three temporal operators \( X, U \) and \( E \) can now be given as follows.\(^{122}\)

\[
\begin{align*}
(1) \quad & M, b \models p \iff p \in g(b_1); \\
(2) \quad & M, b \models X\varphi \iff M, b_{\geq 2} \models \varphi; \\
(3) \quad & M, b \models \varphi U \psi \iff \text{there is some } j \geq 0 \text{ such that } M, b_{\geq j} \models \psi \text{ and for every } k, 1 \leq k < j, \text{ then } M, b_{\geq k} \not\models \varphi; \\
(4) \quad & M, b \models E\varphi \iff \text{there is some fullpath } b' \text{ such that } b_1 = b'_1 \text{ and } M, b' \not\models \varphi.
\end{align*}
\]

The truth conditions for the logical connectives are standard. A formula \( \varphi \) is valid in CTL* iff \( M, b \models \varphi \) for all fullpaths \( b \) in all transition systems \( M \).

\(^{122}\) We follow a formulation of the truth conditions for CTL* given in (Hodkinson and Reynolds 2007, p. 681).
Depending on the particular order of temporal operators occurring in a CTL*-formula, its truth value might not be dependent on a fullpath h, but only on a state $s \in S$. Consider, for instance, the formula $\text{EX} \varphi$. This formula is true at any state $s$ such that there is a state $s'$ with $sRs'$ and $\varphi$ is true at $s'$. Formulae of this kind are state formulae. CTL is defined as the fragment of CTL* which contains only state formulae. The semantics of CTL is identical with the semantics of CTL*. Its formulae are defined by the following BNF.

$$\varphi ::= p \mid \neg \varphi \mid \varphi \land \chi \mid \text{EX} \varphi \mid \text{E} (\varphi \cup \chi).$$

Both CTL and CTL* have been adequately axiomatised. An axiomatisation of CTL was provided by Emerson and Halpern.\textsuperscript{123} CTL* was first axiomatised by Mark Reynolds in 2001.\textsuperscript{124}

### 3.2 CML-structures

CML-structures constitute the central concept in the semantics of CML. They are algebraic structures constituted by a number of sets, relations, and functions.\textsuperscript{125} The sets, relations, and functions jointly represent various aspects of conflicts, such as their temporal development, the possible states they go through, the parties participating in them, and the parties’ beliefs, goals, values, and emotions.

---

\textsuperscript{123} (Emerson and Halpern 1982)

\textsuperscript{124} (Reynolds 2001)

\textsuperscript{125} Algebraic structures are n-tuples of the form $<S_1, \ldots, S_n, R_1, \ldots, R_m, F_1, \ldots, F_k, c_1, \ldots, c_l>$, where $S_1, \ldots, S_n$ are sets called the domains of the structure, $R_1, \ldots, R_m$ are relations (on $S_1, \ldots, S_n$), $F_1, \ldots, F_k$ are functions (on $S_1, \ldots, S_l$), and $c_1, \ldots, c_l$ are constants. For a definition of algebraic structures, see (Ebbinghaus et al. 1994, p. 26; Marker 2002, p. 8). In the case of CML-structures, we are dealing with many-sorted structures, as they have more than one domain. A discussion of many-sorted structures can be found in (Ebbinghaus et al. 1994, p. 45f).
A general definition of CML-structures is provided as follows.

**Definition 22 (CML-Structure)**

A CML-structure \( x \) is a structure of the form \(<T, \prec, W, H, v, A, b, g, n, e, \mu>\), where:

1. \( T \) is a set of time points;
2. \( \prec \) a binary relation on \( T \);
3. \( W \) is a set of possible conflict states;
4. \( H \) is a set of possible conflict histories;
5. \( v \) is a truth assignment function;
6. \( A \) is a set of agents;
7. \( b \) is a belief function;
8. \( g \) is a goal function;
9. \( n \) is a norm function;
10. \( e \) is an emotion function;
11. \( \mu \) is a term assignment function.

Note that the above definition is just a general characterisation of CML-structures. In order for it to work, we will have to introduce further conditions for the sets \( T, W, H, \) and \( A, \) the relation \( \prec, \) and the functions \( v, b, g, n, e, \) and \( \mu. \) This will be done in subsequent sections.

CML-structures are defined for two purposes. First, they allow us to evaluate CML-formulae with respect to their truth value relative to a given CML-structure \( x. \) This, in turn, leads to the usual definitions of logical validity and the semantic consequence relation and thereby makes it possible to classify formulae as tautologies, contingencies, and contradictions, and to determine whether a formula \( \varphi \) semantically follows from a set of formulae \( \Phi. \)

Besides this ‘logical purpose’, CML-structures also frame an abstract ontology for conflicts. This second ‘ontological purpose’ is achieved to the extent that potential conflicts can be reconstructed as CML-structures, i.e. the various aspects of a conflict, such as its temporal development, its parties and their beliefs, goals, norms, and emotions can be identified with the corresponding components of an appropriate CML-

---

126 Various levels of validity for CML-formulae, as well as the semantic consequence relation \( \models, \) are introduced later in the chapter.
structure. Every CML-structure can, therefore, be regarded as an abstract representation of a potential conflict, and vice versa.127

### 3.2.1 Time

In CML-structures, time is represented by a discrete, strict linear order \((T, \prec)\) without endpoints.128 Such an order consists of a set \(T = \{t_0, t_1, t_2, \ldots\}\) and a transitive, irreflexive, trichotomous, and discrete relation \(\prec\) with no endpoints.129 In the context of CML, we call elements of \(T\) time points. According to the discreteness condition, every time point has a unique immediate predecessor and a unique immediate successor.130 Consequently, \(T\) has no endpoints with respect to \(\prec\).

Formally, discrete, strict linear orders without endpoints can be defined as follows.

**Definition 23 (Discrete, Strict Linear Order without Endpoints)**

A discrete, strict linear order without endpoints is a structure of the form \((T, \prec)\), where:

1. \(T = \{t_0, t_1, t_2, \ldots\}\) is a set of time points;
2. \(\prec\) is a binary relation on \(T\) such that:
   - (i) \(\forall t_0, t_1, t_2 (t_0 < t_1 \land t_1 < t_2 \Rightarrow t_0 < t_2)\) (transitivity);
   - (ii) \(\forall t_0 \lnot (t_0 < t_0)\) (irreflexivity);
   - (iii) \(\forall t_0, t_1, t_2 (t_0 < t_1 \lor t_0 = t_1 \lor t_1 < t_0)\) (trichotomy);
   - (iv) \(\forall t_0 \exists t_1, t_2 (t_0 < t_1 \land t_2 < t_0 \land \lnot \exists t_3 ((t_0 < t_3 \land t_3 < t_0) \lor (t_2 < t_1 \land t_3 < t_0)))\) (discreteness/no endpoints).

---

127 CML-structures represent only the logical aspects of a conflict. They can be thought of as mental images obtained from looking at a conflict through glasses that filter all but the logical features of the conflict.

128 An introduction to the theory of linear orders can be found in (Rosenstein 1982). For a detailed discussion of various classes of flows of time, see (Hodkinson and Reynolds 2007, p. 658f).

129 Transitivity, irreflexivity, and trichotomy are the defining properties of a strict linear order. It follows that the ordering relation is also antisymmetric.

130 This definition of discreteness as the property of every element having an immediate successor and an immediate predecessor follows Ian Hodkinson and Mark Reynolds. See (Hodkinson and Reynolds 2007, p. 658f).
Graphically, discrete, strict linear orders without endpoints can be illustrated by an infinite line on which the time points are lined up one after another as shown in the following figure.

![Figure 5: Discrete, Strict Linear Order without Endpoints](image-url)

In contrast to the semantics of branching-time temporal logics, time points are explicitly and independently represented in CML-structures. In particular, they are different from states, the entities at which formulae are evaluated. This distinction between time points and states is a specific feature of our system. In the semantics of branching-time temporal logic, elements of the trees are interchangeably interpreted as time points or states. However, for our purpose it is necessary to distinguish between different states a conflict can be in at one single time point. Consequently, time points and states must be represented by different semantic entities.

Representing time by a discrete, strict linear order without endpoints provides a useful idealisation for the purpose of modelling conflicts. It allows one to break down a conflict into a manageable number of conflict states chronologically ordered by $\prec$.

Having described time points as elements of discrete, strict linear orders without endpoints, we can further characterise them as definite dates. These are expanded points or intervals in time, i.e. specific years, months, days, hours, etc. such as “January 16, 2001”, “The time of Laurant-Désiré Kabila’s assassination”, or “The time before Laurant-Désiré Kabila’s assassination”. In contrast to pseudo-dates, such as “now”, “yesterday” or “five weeks ago”, definite dates are chronologically stable. They refer to the same temporal entity, independent of the context in which they occur.
When we model a conflict, the choice of dates and their status as points or intervals in time depends on the level of granularity at which the conflict is analysed. For example, when analysing the Second Congo War, we could look at it annually from its outbreak in 1998 to its end in 2003, i.e. our dates of interest would be the years from 1998 to 2003. Alternatively, we could analyse it more finely and just look at its outbreak in August 1998. In this case, our dates of interest would be the days of August 1998. Dates are not required to have equal temporal expansion. It is possible to examine a conflict with dates of varying levels of granularity. For instance, we can break up the Second Congo War into dates like “The time before the RCD offensive”, “The time of the RCD offensive”, “The time of the allied counteroffensive”, “The time of internal fighting within the RCD”, etc.

3.2.2 Conflict States and Conflict Histories

At each time point, a conflict is in a certain state. In CML-structures, possible conflict states are represented by letters $w_0$, $w_1$, $w_2$, etc. The set of all conflict states, in which a conflict can possibly be, $W = \{w_0, w_1, w_2, \ldots\}$, is called the conflict universe.

**Definition 24 (Conflict Universe)**

The conflict universe $W = \{w_0, w_1, w_2, \ldots\}$ is a set of possible conflict states.

Distinguishing between different states, phases, or stages through which a conflict passes is common to many conflict models. In his classic work, Anatol Rapoport, for instance, identified three stages of a conflict. More recently, Hayward Alker, Ted Gurr and Kumar Rupesinghe have developed a typology that distinguishes six phases Eric

---

131 Abbreviating possible conflict states by the letter $w$ is inspired by the similarity between possible states and possible worlds as used in possible world semantics.
132 (Rapoport 1960)
133 (Alker et al. 2001)
Brahm describes the life cycle of a conflict by differentiating between seven phases.\footnote{Brahm 2003} Conflict scholars typically link the phases of a conflict to its level of violence. In early and late phases, the level of violence is low, whereas in middle phases it is high.\footnote{The relationship between the progress of time and the level of violence is often illustrated as a reversed U-shaped curve in a coordinate-system, in which the x-axis symbolises time and the y-axis stands for the level of violence. See, for instance, (Pruitt and Rubin 1986; Baros and Jaeger 2004).} For the current purpose, it is irrelevant, however, whether conflict states are defined relative to the level of violence prevalent in the conflict, or whether they are prototypical in the sense that every conflict is assumed to pass through the same sequence of states. We only presuppose that conflicts can be described in terms of sequences of distinguishable states.

Conflict states can be characterised by sets of propositions. Accordingly, a conflict state \( w \) can be specified by providing the set of propositions \( \{p_0, p_1, p_2, \ldots\} \) which are true at \( w \).

The truth of a proposition \( p \), at a certain state \( w \), is independent from the epistemic access an agent may or may not have to it as well as the agent’s goals, norms, or emotions. For instance, it is possible that a proposition \( p \) is true at a state \( w \) even if no agent believes that \( p \) is true, wants \( p \) to be true, considers it a norm that \( p \) should be true, or is emotionally excited about \( p \).

The link between propositions and possible conflict states is expressed by the truth assignment function \( v \) which assigns to every state \( w \) the set of basic propositions that are true at \( w \). As propositions are represented by members of the set \( p\text{-CON} \) of propositional constants,\footnote{The syntactical representations of propositions are propositional constants summarised in the set \( p\text{-CON} \). For details, see section 2.2.1.} \( v \) can be defined as a function from the conflict universe \( W \) into the power set of \( p\text{-CON} \), where \( p \in v(w) \) is interpreted as “\( p \) is true at \( w \)”, as follows.
**Definition 25 (Truth Assignment Function)**

\( v : W \rightarrow \varnothing(\text{p-CON}) \) is a truth assignment function.

The truth assignment function \( v \) is only defined for basic propositions. For compound propositions, expressed by more complex formulae than propositional constants, we will provide a recursive extension of \( v \) in terms of truth conditions which make it possible to calculate the truth value of complex formulae.

If a conflict is in a certain state \( w \) at a certain time point \( t_0 \), it can evolve in different ways. At the next time point \( t_1 \), the conflict is in one of a number of possible states \( w_0, w_1, w_2, \ldots \). In this sense, each conflict state has an ‘open future’ allowing agents involved in the conflict to determine how the conflict develops and external factors to influence the course of the conflict.

As a consequence of the fact that there are a number of different possible states into which a conflict can evolve from a given state, we can distinguish between different possible conflict histories. A conflict history \( h \) is a sequence of possible conflict states specifying one unique conflict state \( w \) for every time point \( t \). If \( h \) specifies the state \( w \) for the time point \( t \), this means that the conflict is in the state \( w \) at \( t \), given that it develops as prescribed by the history \( h \). Conflict histories are represented in CML-structures by functions from the set of time points \( T \) into the conflict universe \( W \). The set of all conflict histories is denoted \( H = \{ h_0, h_1, h_2, \ldots \} \). Their formal definition is given as follows.

**Definition 26 (Conflict Histories)**

\[ H = \{ h_0, h_1, h_2, \ldots \} \subseteq W^T \] is a set of conflict histories such that:

1. \( \forall h_0, h_1, t_0, t_1 (h_0(t_0) = h_1(t_1) \Rightarrow t_0 = t_1) \);
2. \( \forall w \exists h, t (h(t) = w) \);
3. \( \forall h_0, h_1, t_0, t_1 ((t_1 < t_0 \Rightarrow h_0(t_1) = h_1(t_1)) \land (t_0 < t_1 \Rightarrow h_0(t_1) \neq h_1(t_1))). \)
The identity $h(t) = w$ is interpreted as “according to the history $h$, the conflict is in the state $w$ at the time point $t$” or “$w$ lies on $h$ at $t$”.

There are three conditions for conflict histories. According to the first condition, two conflict histories can overlap only at identical time points, i.e. if a state $w$ lies on two different histories $h_0$ and $h_1$, then this state must be reached by both histories at the same time point. In the special the case of $h_0 = h_1$, the condition just expresses that every conflict history is injective. The second condition describes a global surjectiveness property. For every state $w$, there must be a history $h$ and a time point $t$ such that $w$ is the image of $t$ under $h$. The condition does not require every history $h$ to be surjective; it only makes sure that the union of all co-domains of the functions contained in $H$ is exhaustive with respect to $W$. As a consequence, every state can be reached by a possible conflict history. The third condition expresses two properties of $H$. First, for any two histories $h_0$ and $h_1$, there is a time point $t$ such that $h_0$ and $h_1$ are identical up to $t$. This implies that any two histories in $H$ are connected, and, hence, there are no completely parallel histories. Second, once two histories $h_0$ and $h_1$ have split, they do not intersect again. The motivation for this is that if two different states $w_0$ and $w_1$ can be reached at a time $t$, the two possible histories going through $w_0$ and $w_1$, respectively, can never be the same; they differ at least with respect to the propositions true in the conflict state at $t$.

The picture obtained from ordering the conflict states relative to their occurrence in possible conflict histories is shown in the following figure.
The possible conflict states together with the histories constitute a forwards-branching, backwards-linear, tree. The backwards-linearity corresponds to the determinacy of the past, whereas the forwards-branching corresponds to the indeterminacy of the future. As described in the background section on branching-time temporal logic, backwards-linear, forwards-branching, trees are used in the semantics of various logics of branching-time. Hence, with the introduction of conflict histories, we have linked our semantics to semantics of branching-time temporal logic. However, in our semantics conflict histories are defined as functions from the set of time points into the set of states, whereas in the semantics of most branching-time temporal logics they are defined as inclusion maximal subsets of a set of states. Our functional approach, which is due to our distinction between time points and states, is similar to an approach suggested by Brian Chellas in the context of a semantics for a logic of action.

---

137 A proof that the three conditions for possible conflict histories are sufficient to characterise $W$ and $H$ as a backwards-linear, forwards-branching, tree is given in section 3.3.4. The picture of forwards-branching, backwards-linear, trees is closely related to the extensive form of games. Indeed, the main reason for using these kind of structures as the underlying framework for modelling conflicts is the fact that they allow us to incorporate useful elements of game theory into CML without having to deal with more problematic notions of this approach, such as the concepts of utility and probability.


139 (Chellas 1969)
3.2.3 Agents and Attitudes

Most conflict models explicitly represent the parties involved in a conflict and to some extent the parties’ beliefs, goals, values, and emotions.\textsuperscript{140} Depending on the discipline, conflict parties are called actors, parties, stakeholders, or factions.\textsuperscript{141} We use the general term “agent” because it is neutral and can be used in different contexts.

In CML-structures, agents involved in a conflict are represented by a finite and nonempty set $A = \{a_0, a_1, \ldots, a_n\}$ of agents. Each $a \in A$ stands for a unique agent involved in the conflict. Agents are assumed to be social units that can be distinguished from each other and from their environment. They are located either at the individual level or at a higher social level. Agents on the individual level are individual persons. Micro level conflicts include small groups, such as businesses, pressure groups, or minority groups. Agents on the macro level are social institutions, such as armies, rebel groups or political parties, entire nations, or ethnic groups.\textsuperscript{142}

Agents in a conflict have beliefs about the issues relevant to the conflict, they pursue goals, hold moral or legal norms, and have emotions. In CML-structures, all these cognitive and emotional states in which agents can be are expressed by four propositional attitude functions $b$, $g$, $n$, and $e$. Beliefs are represented by the belief function $b$, goals by the goal function $g$, values by the norm function $n$, and emotions by the emotion function $e$. The underlying reason for modelling beliefs, goals, norms, and emotions by four independent functions is the assumption that these mental states do not impact the truth values of propositions. For instance, the fact that an agent believes that $p$ does not

\textsuperscript{140} See, for instance, (Deutsch 1991; Burton 1996; Rapoport 1960).

\textsuperscript{141} Political scientists tend to talk about actors or factions, whereas the term “party” is used more frequently in the psychological literature.

\textsuperscript{142} For an overview over the theory of social institutions, see (Knight and Sened 1998).
affect the truth value of \( p \). In this sense, attitudes can be characterised as subjective. The subjectivity of cognitive and emotional states is also mirrored by the limited epistemic access we have to other agents’ mental states. Only an agent himself can verify whether or not he has a certain belief, goal, norm, or emotion.

The belief function \( b \) assigns a set of CML-formulae to pairs consisting of an agent \( a \) and a state \( w \). This set, called the agent’s belief set at \( w \), represents the agent’s beliefs at the state \( w \). Every formula \( \varphi \in b(w, a) \) is believed to be true by the agent \( a \) at the state \( w \).

Formally, \( b \) is a function from \( W \times A \) into the power-set of the set of CML-formulae FOR.

**Definition 27 (Belief Function)**

\[ b: W \times A \rightarrow \wp(\text{FOR}) \] is a belief function.

The goal function \( g \) assigns a set of CML-formulae to pairs consisting of an agent \( a \) and a state \( w \). Here, the set represents the agent’s goals at the state \( w \) and is called the agent’s goal set at \( w \). If a formula \( \varphi \) is in an agent’s goal set at the state \( w \), i.e. \( \varphi \in g(w, a) \), \( a \) wants \( \varphi \) to be true at the state \( w \). Formally, \( g \) is a function from \( W \times A \) into the power-set of FOR.

**Definition 28 (Goal Function)**

\[ g: W \times A \rightarrow \wp(\text{FOR}) \] is a goal function.

The norm function \( n \) and the emotion function \( e \) follow the same pattern as the functions \( b \) and \( g \). They assign sets of CML-formulae to pairs of agents and states. In the case of \( n \), the set is called the agent’s norm set at \( w \) and is interpreted as the set of norms held by the agent at the state \( w \). These can be moral, legal or religious norms or values. Formally, \( n \) is a function from \( W \times A \) into \( \wp(\text{FOR}) \).
Definition 29 (Norm Function)
n: W × A → \varnothing (FOR) is a norm function.

In the case of e, e(a, w) is interpreted as the set of formulae which trigger a state of emotional excitement in the agent a at the state w. Every formula \( \varphi \in e(w, a) \), causes an emotional response in a at the state w. Formally, e is a function from W × A into \( \varnothing (FOR) \).

Definition 30 (Emotion Function)
e: W × A → \varnothing (FOR) is an emotion function.

A graphical illustration of the four propositional attitude functions b, g, n and e is shown in the following figure.

Figure 7: Propositional Attitude Functions

The figure illustrates the beliefs, goals, norms and emotions held by an agent a₂ at the possible conflict state w₀. As the figure shows, the agent a₂ believes p₀, wants p₀ and p₁, considers p₂ a norm, and is emotionally excited about \( \neg p₁ \) and \( \neg p₂ \).
A basic condition for a proposition to qualify as the content of a propositional attitude is that the proposition must trigger some kind of cognitive or emotional reaction when presented to the agent who holds the attitude. As a further distinction, we can look at consequences following from the fact that an agent holds a particular type of propositional attitude. If the agent believes p, he believes that the world is structured such that the state of affairs corresponding to p holds. If an agent wants p, he prefers a world in which p is true to one in which p is not true. This preference holds independent of whether p is actually true or not or whether the agent believes that p is true or not. If an agent considers p a legal or moral norm, he thinks that a world in which p is true is a better world in some moral, legal, or religious sense than one in which p is not true. If an agent is emotionally excited about p, he exhibits a certain type of bodily reaction when confronted with p.

3.2.4 The Term Assignment Function

In order to state truth conditions for quantifiers and to provide a notion of validity for open formulae, we need an interpretation for terms. This is done by means of the term assignment function μ, which assigns elements of T and A to terms of the language of CML, i.e. it is a function from the set TER of CML-terms into the union T ∪ A.

As introduced in chapter 2.2, there are two types of terms in ℒCML: temporal terms and agent terms. For constants, μ is fixed, i.e. the μ-value of a temporal constant t is a specific time point t, and the μ-value of an agent constant a is a specific agent a. We use

---

143 A number of different approaches to the notion of propositional attitudes are discussed in (Baeurle and Cresswell 1989). A purely modal approach to propositional attitudes is described in (Moss and Tiede 2006, p. 1038ff).

144 For a concise introduction to emotional psychology see, for instance, (Oatley and Jenkins 1996). A more neurologically oriented introduction to emotions can be found in (Simonov 1986).
The semantics of Conflict Modelling Logic

the same letter in bold for the element a constant refers to. With regard to variables, we have to make sure that \( \mu \) assigns time points to temporal variables and agents to agent variables. Altogether, we get the following definition.

**Definition 31 (Term Assignment Function)**

\( \mu: \text{TER} \rightarrow \mathbb{T} \cup \mathbb{A} \) is a term assignment function such that:

1. \( \mu(t) = t \), for all \( t \in \text{t-CON} \);
2. \( \mu(x) \in \mathbb{T} \), for all \( x \in \text{t-VAR} \);
3. \( \mu(a) = a \), for all \( a \in \text{a-CON} \);
4. \( \mu(y) \in \mathbb{A} \), for all \( y \in \text{a-VAR} \).

Fixed term assignments are defined as usual. As there are two types of variables, it is possible to either fix a temporal variable or an agent variable. In the temporal case, the expression \( \mu[t/x] \) denotes the term assignment function that assigns \( t \) to \( x \) and is equal to \( \mu \) for all terms different from \( x \). Similarly, \( \mu[a/y] \) assigns \( a \) to \( y \) and is equal to \( \mu \) for all terms different from \( y \). Formally, we get the following definition.

**Definition 32 (Fixed Term Assignment)**

Fixed term assignments are defined as follows:

1. \( \mu[t/x](\tau) = \mu(\tau) \), for all \( \tau \neq x \);
2. \( \mu[t/x](\tau) = t \), for \( \tau = x \);
3. \( \mu[a/y](\alpha) = \mu(\alpha) \), for all \( \alpha \neq y \);
4. \( \mu[a/y](\alpha) = a \), for \( \alpha = y \).

### 3.3 Further Definitions

In this section, four further concepts are introduced. First, we define the immediate predecessor \( t-1 \) of a time point \( t \). Second, we introduce the concept of a set of possible states at a fixed time point \( t \). Third, we introduce a time index function which assigns to every state \( w \) the corresponding time point \( t \) at which the state occurs in the conflict universe. Finally, we define a binary accessibility relation \( R \) between states. \( R \) links the concept of possible histories as used in our semantics to relational semantics in which a
set \( W \) of possible worlds together with a binary accessibility relation \( R \subseteq W^2 \) are used as the primitive constituents of the semantics.\(^{145}\) We use these concepts to prove that the set of all sets of possible conflict states, at fixed time points, provides a partition of the conflict universe \( W \) (Theorem 3) and that the structure generated by the conflict universe \( W \) and the set of possible conflict histories \( H \) is a forwards-branching, backwards-linear, tree without endpoints (Theorem 5).

### 3.3.1 Immediate Predecessor of a Time Point

The immediate predecessor \( t^{-1} \) of a time point \( t \) is defined as follows.

**Definition 33 (Immediate Predecessor of a Time Point)**

The immediate predecessor \( t^{-1} \) of a time point \( t \) is defined as follows:

\[
t^{-1} = t_0 \text{ iff } t_0 \prec t \land \exists t_1 (t_0 \prec t_1 \land t_1 \prec t).
\]

The definition is possible as, by definition of \( \prec \), every time point has exactly one unique immediate predecessor.

### 3.3.2 Conflict States of a Fixed Time Point

At every time point \( t \), a conflict is in one of a number of possible conflict states \( w_0, w_1, w_2, \) etc. It is useful to collect all these states in a set \( W_t \) defined as the set of states in which a conflict can be at the fixed time point \( t \).

\(^{145}\) Possible world semantics provide the standard semantic interpretation for a large number of logics, such as modal, deontic, epistemic and temporal logics. The basic ideas of possible world semantics were independently developed in the late 1950 and early 1960 by Saul Kripke, Jaakko Hintikka and Stig Kanger. See (Kripke 1959; Kripke 1963; Hintikka 1962; Kanger 1957).
**Definition 34 (States at a Fixed Time Point)**
The set of possible conflict states at the fixed time point \( t \) is defined as follows:
\[
W_t = \{ w \in W \mid \exists h \in H(h(t) = w) \}.
\]

As every state occurs only at one unique time point, the sets \( W_t \) partition the conflict universe \( W \), i.e. any two sets \( W_{t_1} \) and \( W_{t_2} \) are mutually disjoint and the union \( \bigcup_{t \in T} W_t \) is equal to \( W \).

**Theorem 3 (Partition of \( W \))**
\( \{ W_t \mid t \in T \} \) is a partition of \( W \).

**Proof**

**(Disjointness)** We prove by reductio. Let \( W_{t_1}, W_{t_2} \in \{ W_t \mid t \in T \} \) with \( t_1 \neq t_2 \). Assume now there is a \( w \in W \) with \( w \in W_{t_1} \cap W_{t_2} \). Then, there are \( h_1, h_2 \in H \) with \( h_1(t_1) = w \) and \( h_2(t_2) = w \) by Definition 34. Then, \( t_1 = t_2 \) by Definition 26. This contradicts \( t_1 \neq t_2 \). Hence, there is no \( w \in W_{t_1} \cap W_{t_2} = \emptyset \).

**(Exhaustiveness)** Let \( w \in W \). Then, by Definition 26, there is a \( t \in T \) and a \( h \in H \) such that \( h(t) = w \). Then, \( w \in W_t \) by Definition 34, and, hence, \( w \in \bigcup_{t \in T} W_t \). Now let \( w \in \bigcup_{t \in T} W_t \). Then, by Definition 34, \( w \in W_{t_b} \) for some \( t_b \in T \), and, hence, \( w \in W \) as \( W_{t_b} \subseteq W \). As \( w \) was chosen arbitrarily we get \( \bigcup_{t \in T} W_t = W \).

\[ QED \]

A graphical illustration of the partition \( \{ W_t \mid t \in T \} \) is shown in the following figure.

---

**Figure 8: Partition of the Conflict Universe**
3.3.3 The Time Index Function

As a consequence of Theorem 3, every conflict state \( w \) can be associated with a unique time point \( t \) at which it occurs in the conflict universe \( W \). In the following, we define a time index function \( |\cdot| \) assigning this unique time point to every state \( w \).

**Definition 35 (Time Index Function)**

\[
|w| : W \rightarrow T \text{ is a time index function defined as follows:}
\]

\[
|w| = t \text{ iff } w \in W_t.
\]

\(|w|\) is called the time index of \( w \). The function is well defined because \( \{W_t | t \in T\} \) is a partition of \( W \), and, hence, every \( w \) belongs to only one set \( W_t \), i.e. if \( w_0 = w_1 \), then \( |w_0| = |w_1| \) for all \( w_0, w_1 \in W \).

The relationship between the time index function and individual conflict histories is expressed in the following theorem.

**Theorem 4 (Time Index Function and Possible Conflict Histories)**

For all \( h \in H \) and \( t \in T \), \( |h(t)| = t \).

**Proof**

Let \( h \in H \) and \( t \in T \). As \( h \) is a function from \( T \) into \( W \), there is a \( w \in W \) such that \( h(t) = w \). Then, \( w \in W_t \) by Definition 34, and, hence, \( |w| = t \) by Definition 35. Altogether we get \( |h(t)| = |w| = t \).

QED

The theorem shows that if \( |\cdot| \) is restricted to the co-domain of a conflict history \( h \), it is the inverse function of \( h \).
3.3.4 The Accessibility Relation

In the semantics of CML, both the set of possible conflict states $W$ and the set of possible conflict histories $H$ are introduced as primitive concepts of CML-structures.\(^{146}\) Instead of starting with possible states and possible histories as the primitive concepts, we could have built our semantics around the concept of backwards-linear, forwards-branching, trees. Then, we could have defined the set of histories $H$ in terms of the binary relation constituting the trees.\(^{147}\)

Conflict states and histories together constitute a backwards-linear, forwards-branching, tree. To prove this, we define a binary relation $R$ on $W$, which we show to be irreflexive, transitive, and treelike. The accessibility relation $R$ is defined in terms of the time index function $|\ |$ and the set of possible conflict histories $H$.

**Definition 36 (Accessibility Relation)**

$R \subseteq W \times W$ is an accessibility relation defined as follows:

$w_0 R w_1$ iff $|w_0| < |w_1|$ and $\exists h (w_0, w_1 \in \text{Im}(h))$.

According to this definition, a state $w_1$ is accessible from a state $w_0$ if and only if $w_0$ temporally precedes $w_1$, i.e. the time index of $w_0$ is smaller than the time index of $w_1$, and the two states lie on a common history $h$, i.e. both $w_0$ and $w_1$ are elements of the image of $h$, $\text{Im}(h)$.

In the context of conflicts, the accessibility relation $R$ can be interpreted as a transition relation between conflict states. If a conflict is in a state $w_0$, and a state $w_1$ is accessible

---

\(^{146}\) The strategy of defining a set of states $W$ and a set of histories, understood as functions from a discrete set of time points into $W$, as the primitive constituents of the semantics, was also chosen by Chellas in his logical system of action statements. See (Chellas 1969).

\(^{147}\) This approach is taken in many branching-time temporal logics as discussed in the respective background section, but also in STIT theory, where treelike structures are introduced as a primitive component of the semantics. Histories are then defined as maximal sets of linearly ordered moments of the structure. See (Belnap 1991; Belnap et al. 2001).
from \(w_0\), i.e. \(w_0Rw_1\), it is possible for the conflict to develop from the state \(w_0\) into the state \(w_1\).

The following theorem expresses that the structure \((W, R)\), consisting of the conflict universe \(W\) and the accessibility relation \(R\), constitutes a backwards-linear, forwards-branching, tree without endpoints.

Theorem 5  (Tree)

\((W, R)\) is a tree without endpoints.

Proof

We have to show that \(R\) is irreflexive, transitive, treelike, and has no endpoints.

(Irreflexivity) Let \(w \in W\). Assume that \(wRw\). Then, \(|w| \prec |w|\) by Definition 36. Then, there is a \(t \in T\) such that \(wRt\) by Definition 35. This contradicts the irreflexivity of \(\prec\).

(Transitivity) Let \(w_0, w_1, w_2 \in W\). Assume that \(w_0Rw_1\) and \(w_1Rw_2\). Then, \(|w_0| \prec |w_1|\), \(|w_1| \prec |w_2|\) and there are \(h_0, h_1 \in H\) such that \(w_0, w_1 \in \text{Im}(h_0)\) and \(w_1, w_2 \in \text{Im}(h_1)\) by Definition 36. Because \(w_1 \in \text{Im}(h_0)\) and \(w_1 \in \text{Im}(h_1)\) we get \(h_0(|w_1|) = h_1(|w_1|)\), and, hence, \(h_1(|w_0|) = h_0(|w_0|)\) by Definition 26. Altogether we get \(h_1(|w_0|) = w_0, h_1(|w_2|) = w_2\) and \(|w_0| \prec |w_2|\) by the transitivity of \(\prec\), and, hence, \(w_0Rw_2\) by Definition 36.

(Left-Linearity) Let \(w_0, w_1, w_2 \in W\). Assume that \(w_1Rw_2\) and \(w_1Rw_2\). Then, \(|w_0| \prec |w_1|\), \(|w_1| \prec |w_2|\) and there are \(h_0, h_1 \in H\) such that \(w_0, w_2 \in \text{Im}(h_0)\) and \(w_1, w_2 \in \text{Im}(h_1)\) by Definition 36. Because \(w_2 \in \text{Im}(h_0)\) and \(w_2 \in \text{Im}(h_1)\) we get \(h_0(|w_2|) = h_1(|w_2|)\), and, hence, \(h_0(|w_0|) = h_1(|w_0|)\) and \(h_0(|w_1|) = h_1(|w_1|)\) by Definition 26. Altogether we have \(h_1(|w_0|) = w_0, h_1(|w_1|) = w_1, h_1(|w_2|) = w_2\) and \(|w_0| \prec |w_1|\), \(|w_0| = |w_1|\) or \(|w_1| \prec |w_0|\) by the trichotomy of \(\prec\). Hence, \(w_0Rw_1, w_0 = w_1\) or \(w_1Rw_0\) by Definition 36.

(No endpoints) Let \(w_0 \in W\). Then, there is a \(h \in H\) and \(t_0 \in T\) such that \(h(t_0) = w_0\) by Definition 26. Then, there are \(t_0, t_2 \in T\) such that \(t_1 \prec t_0\) and \(t_0 \prec t_2\) because \(\prec\) has no endpoints. Then, there are \(w_1, w_2 \in W\) such that \(h(t_0) = w_1\) and \(h(t_2) = w_2\) because \(h\) is function from \(T\) into \(W\). Altogether there is a \(h \in H\) such that \(h(|w_0|) = w_0, h(|w_1|) = w_1, h(|w_2|) = w_2\), and \(|w_1| \prec |w_0|\) and \(|w_0| \prec |w_2|\), and, hence, \(w_0Rw_1, w_0 = w_1\) or \(w_1Rw_2\) by Definition 36.

QED

The theorem shows that history functions can be used to define an accessibility relation \(R\) and a corresponding tree \((W, R)\), given that they satisfy the three conditions of Definition 26.
3.4 Truth Conditions

The meaning of a CML-formula \( \varphi \) is provided by specifying the conditions under which \( \varphi \) is true in a given CML-structure \( x \). These conditions are referred to as the truth conditions of \( \varphi \). In order to make sure that every formula of \( \mathcal{L}_{\text{CML}} \) can be evaluated, we follow the usual recursive strategy, i.e. we start with a truth condition for atomic formulae and continue with conditions for compound formulae assuming that their constituents can already be evaluated. In the atomic case the truth condition is already provided by the truth assignment function \( v \).\(^{148}\) In the following, we extend the truth assignment function to all other operators and relations of \( \mathcal{L}_{\text{CML}} \). When we do this, we have to include the term assignment function \( \mu \) because formulae with free occurrences of variables can only be evaluated once it is specified which objects the terms occurring in them refer to. As a result, \( v \) depends on the term assignment function \( \mu \). Furthermore, \( v \) depends on the conflict universe \( W \) because formulae are evaluated relative to possible conflict states.\(^{149}\)

3.4.1 Conditions for the Logical Connectives, Identity, and the Quantifier

The truth conditions for logical connectives, the identity relation, and the quantifier are just the usual conditions of classical propositional and first-order logic.\(^{150}\) The only difference is that CML-formulae are evaluated relative to a possible conflict state \( w \).

\(^{148}\) A propositional constant \( p \) is true at a possible conflict state \( w \) if and only if \( w \in v(p) \).

\(^{149}\) To evaluate formulae relative to points is a common practice in possible word semantics. The points are typically interpreted as worlds, states, time points or situations, depending on the nature of the intended application of the semantics.

\(^{150}\) For a typical list of truth conditions for first-order formulae, see, for instance, (Ebbinghaus et al. 1994, p. 31).
Chapter 3 Interpreting Conflicts

Definition 37 (Truth Conditions for Logical Connectives, Identity, and the Quantifier)

1. \( \neg \phi \in v_p(w) \) iff \( \phi \notin v_p(w) \);
2. \( \phi \land \psi \in v_p(w) \) iff \( \phi \in v_p(w) \) and \( \psi \in v_p(w) \);
3. \( \tau_0 = \tau_1 \in v_p(w) \) iff \( \mu(\tau_0) = \mu(\tau_1) \), where \( \tau_0, \tau_1 \in \text{TER} \);
4. \( \forall x\phi \in v_p(w) \) iff \( \phi \in v_{p[t/\mu]}(w) \) for all \( t \in T \);
5. \( \forall y\phi \in v_p(w) \) iff \( \phi \in v_{p[a/\mu]}(w) \) for all \( a \in A \).

The first two conditions deal with the negation and conjunction symbol, respectively. The negation of a formula is true at a possible conflict state \( w \) if and only if the negated formula is false at \( w \). A conjunction is true at \( w \) if and only if the constituent conjuncts are true at \( w \). The third condition deals with formulae built with the identity symbol. As usual, the identity relation holds between two terms if and only if their respective \( \mu \)-values are identical. The last two conditions deal with the universal quantifier. Two separate conditions are required because CML-structures are two-sorted structures distinguishing between temporal variables and agent variables. In the case of temporal variables, a universally quantified formula is true at a state \( w \) if and only if the formula itself is true at \( w \) for every \( t \)-alternative to \( \mu \). In the case of agent variables, a universally quantified formula is true at \( w \) if and only if the formula is true at \( w \) for every \( a \)-alternative to \( \mu \). Conditions for the defined connectives and the existential quantifier can be derived in the usual way.

3.4.2 Conditions for the Modal Operator

The truth condition for the modal operator \( \square \) is given as follows.

Definition 38 (Truth Condition for the Modal Operator)

\( \square \phi \in v_p(w) \) iff \( \phi \in v_p(h_1(|w|)) \) for all \( h_1 \in H \) such that \( h_1(|w|-1) = h_0(|w|-1) \) for some \( h_0 \in H \) such that \( h_0(|w|) = w \).
According to the definition, a formula of the form $\Box \varphi$ is true at a possible conflict state $w$, if $\varphi$ is true at all states which have the same time index as $w$, and lie on histories going through the immediate predecessor $h_0(|w|-1)$ of $w$. As the truth condition for $\Box$ does not depend on the set of conflict histories $H$, formulae with $\Box$ as their main operator are state formulae.\(^{151}\)

Graphically, the truth condition for $\Box$ can be illustrated as follows.

![Figure 9: Truth Condition of the Necessity Operator](image)

The figure shows that $\Box \varphi$ is true at $w_3$, because $\varphi$ is true at $w_4$ and $w_5$, the only two states having the common immediate predecessor $w_2$. $\varphi$ is not required to be true at $w_3$, as $w_3$ is not accessible from $w_2$.

The modal operator expresses a type of necessity which is characterised by the concept of historical inevitability. Historical inevitability means that a formula $\varphi$ is necessarily true at a state $w$ if the truth of the formula was inevitable at the state immediately preceding $w$. Following Rescher and Urquhart, we can describe historical necessity by characterising

\(^{151}\) Indeed, all CML-formulae are state formulae.
Chapter 3  Interpreting Conflicts

a formula \( \varphi \) as necessarily true at a state \( w \) if \( \varphi \) would have held at \( w \) “no matter what the course of history”.\(^{152}\)

The notion of historical necessity used in the semantics of CML is a special case of the more general notion of nodally relativised necessity.\(^{153}\) The underlying idea of nodally relativised necessity is to define necessity relative to nodes, worlds, or the like. A formula \( \varphi \) is necessarily true relative to a state \( w \) if and only if the truth of \( \varphi \) was inevitable at \( w \).

In the semantics of CML, the state which the necessity operator is relative to is fixed to be the immediate predecessor of the state at which the formula is evaluated.

3.4.3  Conditions for the Temporal Operator and other Temporal Symbols

In this section, we specify the truth conditions for the temporal realisation operator \( R \), the temporal indexical symbol \( n \), and the temporal precedence relation \( < \). The conditions are defined as follows.

**Definition 39 (Truth Conditions for the Temporal Operator and other Temporal Symbols)**

1. \( R_\tau \varphi \in v_\mu(w) \) iff \( \varphi \in h(\mu(\tau)) \) for all \( h \in H \) such that \( h(|w|) = w \), where \( \tau \in t\text{-TER} \);
2. \( Rn\varphi \in v_\mu(w) \) iff \( \varphi \in v_\mu(w) \);
3. \( \tau = n \in v_\mu(w) \) iff \( \mu(\tau) = |w| \), where \( \tau \in t\text{-TER} \);
4. \( \tau < n \in v_\mu(w) \) iff \( \mu(x) < |w| \), where \( \tau \in t\text{-TER} \);
5. \( \tau_0 < \tau_1 \in v_\mu(w) \) iff \( \mu(\tau_0) < \mu(\tau_1) \), where \( \tau_0, \tau_1 \in t\text{-TER} \).

The first condition describes the behaviour of the temporal realisation operator \( R \). A formula of the form \( R_\tau \varphi \) is true at a state \( w \) if and only if \( \varphi \) is true at the \( \mu \)-value of \( \tau \) of every history \( h \) going through \( w \). Note that the truth condition for \( R \) does not depend on

\(^{152}\) See (Rescher and Urquhart 1971, p. 133).

\(^{153}\) See (Kripke 1963; Rescher and Urquhart 1971, p. 133; Thomason 1984).
the set of conflict histories \( H \). Hence, formulae with \( R \) as their main operator are state formulae.\(^{154}\) If \( \mu(\tau) \) lies in the future of \( w \), i.e. \( |w| < \mu(\tau) \), \( \varphi \) must be true at all states occurring at the time point \( \mu(\tau) \) into which the conflict can possibly develop from \( w \). If \( \mu(\tau) \) lies in the past of \( w \) or is equal to \( w \), i.e. \( \mu(\tau) < |w| \) or \( \mu(\tau) = |w| \), \( \varphi \) must be true at the unique state occurring at the time point \( \mu(\tau) \) which lies on the unique past of \( w \).\(^{155}\)

The above truth conditions characterises \( R \) as a Peircean operator. The Peircean approach interprets a proposition of the form “it will be the case that \( \varphi \)” in the sense that \( \varphi \) is bound to happen in every possible future. This is reflected by the truth condition for \( R \).\(^{156}\)

Graphically, we can illustrate the truth condition for \( R \) as follows.

\[ T \quad \cdots \quad h_0 \quad w_0 \quad w_1 \quad w_2 \quad w_3 \quad \cdots \]

\[ h_1 \quad w_4 \quad \varphi \quad R \varphi \]

\[ h_2 \quad w_5 \quad \varphi \]

\[ \mu(\tau) \]

**Figure 10: Truth Condition of the Temporal Realisation Operator (Future)**

\(^{154}\) The distinction between state formulae and path formulae was introduced in the background section on branching-time temporal logic. For a further discussion of the topic, see (Goranko and Zanardo 2004).

\(^{155}\) Due to the past determinacy and future indeterminacy of our semantics, there is only one unique state \( w_0 \) that lies on a history through \( w \) if \( \mu(\tau) \) refers to a point in the past of \( w \), but there can be more than one possible states that lie on histories going through \( w \) and have an identical time index lying in the future of \( w \).

\(^{156}\) We have introduced the Peircean view and the Ockhamist view in branching-time temporal logic in the background section on branching-time temporal logic. For further details, see (Prior 1967; Venema and Muijdergracht 2001, p. 14; Thomason 1984).
The figure illustrates the states at which \( \varphi \) must be true if \( R_\tau \varphi \) is assumed to be true at \( w_2 \). It shows that \( \varphi \) must be true at \( w_4 \) and \( w_5 \) as those two states lie on histories going through \( w_2 \) and occur at the time point \( \mu(\tau) \). \( \varphi \) does not need to be true at \( w_3 \) as \( w_3 \) does not lie on a history going through \( w_2 \).

The following figure is similar to Figure 10. It also illustrates the conditions under which \( R_\tau \varphi \) is true at \( w_2 \). Here, however, the \( \mu \)-value of \( \tau \) lies in the past of \( w \). As the figure shows, for \( R_\tau \varphi \) to be true at \( w_2 \), \( \varphi \) has only to be true at \( w_0 \) since this is the only state with a time index equal to the \( \mu \)-value of \( \tau \) lying on a history going through \( w_2 \).

![Figure 11: Truth Condition of the Temporal Realisation Operator (Past)](image)

The second condition provides a semantic interpretation of the indexical symbol \( n \). According to the condition, a formula of the form \( R_n \varphi \) is true at a possible state \( w \) if and only if \( \varphi \) is true at \( w \). Hence, \( n \) can be used to syntactically refer to the possible state at which a formula is evaluated. In contrast to terms, \( n \) does not have a fixed \( \mu \)-value. Its interpretation depends on the possible state at which the formula containing \( n \) is

---

157 A detailed semantic analysis of the indexical “now” can be found in (Kamp 1971). Burgess gives a concise overview of Kamp’s analysis in (Burgess 1984, p. 30f).
evaluated. A similar interpretation of \( n \) is given in Rescher’s system \( \text{K}_b \) of temporal logic.\(^{158}\)

The third and fourth conditions deal with formulae in which the indexical symbol \( n \) occurs as part of an identity statement or in the context of the temporal precedence relation \(<\). In both conditions, \( n \) is interpreted as the time index \(|w|\) of the state \( w \) at which the formula is evaluated. Thus, \( \tau = n \) is true at a possible state \( w \) if and only if the \( \mu \)-value of \( \tau \) is equal to \(|w|\). Similarly, \( \tau < n \) is true at a possible state \( w \) if and only if the \( \mu \)-value of \( \tau \) temporally precedes \(|w|\).

The fifth condition provides an interpretation for formulae built with the temporal precedence symbol \(<\). Formulae of the form \( \tau_0 < \tau_1 \) are true at a possible state \( w \) if and only if the corresponding order relation \(<\) holds between the \( \mu \)-value of \( \tau_0 \) and the \( \mu \)-value of \( \tau_1 \).\(^{159}\) The condition shows that the truth value of a formula of the form \( \tau_0 < \tau_1 \) does not depend on the possible state \( w \) at which it is evaluated. As a consequence, such formulae are either true at every possible state or false at every possible state in the conflict universe.

### 3.4.4 Conditions for the Propositional Attitude Operators

The truth conditions for the propositional attitude operators all follow the same pattern. A formula of the form \( B_\alpha \varphi, G_\alpha \varphi, N_\alpha \varphi, \) or \( E_\alpha \varphi \) is true at a state \( w \) just in case \( \varphi \) is in the respective belief set, goal set, norm set, or emotion set at \( w \) of the agent designated by the \( \mu \)-value of \( \alpha \). As these sets are referred to by the images of the propositional attitude

\(^{158}\) Cf. (Rescher and Urquhart 1971, p. 51).

\(^{159}\) The condition makes sure that the temporal precedence symbol is interpreted as \(<\) on \( T \).
functions $b$, $g$, $n$, and $e$, respectively, we can state the truth conditions for the four attitude operators as follows.

**Definition 40 (Truth Conditions for the Propositional Attitude Operators)**

1. $B\alpha\varphi \in \psi_\mu(w)$ iff $\varphi \in b(w, \mu(\alpha))$;
2. $G\alpha\varphi \in \psi_\mu(w)$ iff $\varphi \in g(w, \mu(\alpha))$;
3. $N\alpha\varphi \in \psi_\mu(w)$ iff $\varphi \in n(w, \mu(\alpha))$;
4. $E\alpha\varphi \in \psi_\mu(w)$ iff $\varphi \in e(w, \mu(\alpha))$.

The truth conditions show that the truth values of formulae having a propositional attitude operator as their main operator depend neither on the structure of the conflict universe, i.e. on how the possible states in the conflict universe are connected with each other, nor on the initial truth assignment function $v$. The truth value of such formulae is entirely determined by the term assignment function $\mu$ and the respective propositional attitude function $b$, $g$, $n$ or $e$.\(^{160}\)

### 3.5 Validity

Validity is defined in the usual way, i.e. valid formulae are those formulae that are invariantly true. The invariance can occur on different levels. At the most local level, a formula $\varphi$ can be true at a state $w$ under every term assignment function $\mu$. In this case, we call the formula valid at $w$. One step higher, we can define a formula $\varphi$ to be valid in a CML-structure $x$ if it is valid at every state $w$ in the conflict universe $W$ of $x$ under every truth assignment function $v$. General CML-validity is the property of a formula $\varphi$ to be valid in every CML-structure.

\(^{160}\) The propositional attitude functions are primitive in the definition of CML-structures.
3.5.1 Validity at States

On the most local level, we can define a formula \( \varphi \) to be valid at a possible conflict state \( w \) if the truth value of \( \varphi \) is true independent of how the terms occurring in \( \varphi \) are interpreted. In this case the truth value of \( \varphi \) does not depend on the term assignment function \( \mu \). Symbolically, we write \( x, w \vDash \varphi \) to express that the formula \( \varphi \) is valid at the state \( w \) in the CML-structure \( x \).

**Definition 41 (Validity at States)**
A formula \( \varphi \) is valid at the possible conflict state \( w \) in the CML-structure \( x \), i.e. \( x, w \vDash \varphi \), iff \( \varphi \in v_{\mu}(w) \) for all term assignment functions \( \mu \).

The definition of validity at a state \( w \) is just an expansion of the notion of truth at that state. The difference between a formula being true at a state \( w \) and being valid is that the formula is true at \( w \) under a specific term assignment function \( \mu \), whereas it is valid at \( w \) under every term assignment function. As the truth value of closed formulae does not depend on how the terms occurring in them are interpreted, the notion of validity at a state \( w \) and the notion of truth at a state \( w \) is identical for closed formulae. Open formulae are valid at a state \( w \) just in case their universal closure is true at \( w \).

3.5.2 Validity in CML-Structures

Having defined validity at conflict states \( w \), we can now generalise the concept and define a formula \( \varphi \) to be valid in a CML-structure \( x \) if the formula is valid at every state \( w \) in the universe of \( x \) under every truth assignment function \( v \). Symbolically, we write \( x \vDash \varphi \) to express that the formula \( \varphi \) is valid in the CML-structure \( x \).

**Definition 42 (Validity in CML-Structures)**
A formula \( \varphi \) is valid in the CML-structure \( x \), i.e. \( x \vDash \varphi \), iff \( x, w \vDash \varphi \) for all truth assignment functions \( v \) and all possible conflict states \( w \in W \).
To say that a formula $\phi$ is valid in a CML-structure $\mathcal{X}$ means that $\phi$ is valid in every CML-structure that is identical with $\mathcal{X}$ except for its term assignment function $\mu$ and its truth assignment function $v$.

### 3.5.3 Validity in CML

On the most global level, we can define a formula $\phi$ to be CML-valid if $\phi$ is valid in every CML-structure. Symbolically, we write $\models_{CML} \phi$ to express that the formula $\phi$ is generally valid in CML.

**Definition 43 (CML-Validity)**

A formula $\phi$ is CML-valid, i.e. $\models_{CML} \phi$, iff $\models \phi$ for all CML-structures $\mathcal{X}$.

The definition of CML-validity follows the standard practice in possible-world semantics to define validity relative to a pre-defined class of algebraic structures. In our case, this class is the class of CML-structures. Thus, a formula is valid in the logical system CML if it is true at every point in every CML-structure.

### 3.6 Satisfiability

Using the notion of validity as defined in the previous sections, we can define two further concepts. Satisfiability is defined in the usual way, i.e. a formula $\phi$ is CML-satisfiable if there is a CML-structure $\mathcal{X}$ and a possible conflict state $w$ in the universe $W$ of $\mathcal{X}$, such that $\phi$ is true at $w$.

**Definition 44 (CML-Satisfiability)**

A formula $\phi$ is CML-satisfiable iff there is a CML-structure $\mathcal{X}$ and a state $w$ in the universe of $\mathcal{X}$ such that $\phi \in v_\mu(w)$, where $\mu$ is the term assignment function of $\mathcal{X}$. 

152
3.7 The Semantic Consequence Relation

The semantic consequence relation $\models$, which holds between sets of formulae and single formulae, is defined in the usual way: a formula $\varphi$ semantically follows from a set of formulae $\Phi$ if $\varphi$ is valid in every CML-structure $x$ in which all formulae contained in $\Phi$ are valid.

**Definition 45 (Semantic Consequence Relation)**

A formula $\varphi$ semantically follows from a set of formulae $\Phi$, i.e. $\Phi \models \varphi$, iff $x \models \varphi$ for every CML-structure $x$ such that $x \models \Phi$, where $x \models \varphi$ if $x \models \psi$ for all $\psi \in \Phi$.

**Example: Second Congo War**

In this section, we apply the semantics of CML to our example of the Second Congo War. In particular, we show how aspects of this conflict can be reconstructed as elements of a CML-structure.

The first component of a CML-structure is the set $T$ of time points. When we identify time points in our example, we thereby fix the level of granularity at which we analyse the Second Congo War. We identify six time points at which we look at the conflict: $t_0$, $t_1$, $t_2$, $t_3$, $t_4$, and $t_5$ standing for the years 1998, 1999, 2000, 2001, 2002, and 2003, respectively.\(^{161}\)

The six time points are ordered by the temporal precedence relation $\prec$ in the natural order.

At each of the six time points, the Second Congo War was in a corresponding state $w_{10}$, $w_1$, $w_2$, $w_3$, $w_4$, and $w_5$. Furthermore, we assume that in each of the years from 1999 to

\(^{161}\) Our data is based on reports by the Heidelberg Institute on International Conflict, see (HIIK, 1998; HIIK, 1999; HIIK, 2000; HIIK, 2001; HIIK, 2002; HIIK, 2003) and descriptions of the conflict by the International Crisis Group, see (ICG 2003b; ICG 2005; ICG 2008b; ICG 2008a; ICG 2003a).
2003 the Second Congo War could have been in a different state \( w_6, w_7, w_8, w_9 \), and \( w_{10} \), respectively.

At each time point other than \( t_0 \), the Second Congo War could have been in two different states. Hence, there are six different possible conflict histories \( h_0, h_1, h_2, h_3, h_4, \) and \( h_5 \).

Graphically, we can illustrate the CML-structure representing parts of the Second Congo War as follows.
Next, we identify propositions that are true or false at the eleven conflict states. We identify propositions in such a way that certain facts about the Second Congo War can be represented. The facts we look at are listed in the following box.

In 1998 an offensive was launched against the Kabila government.
In 1999 the rebel group RCD split into two factions, the RCD-Goma and the RCD-Kisangani.
In 2000 the peace keeping mission MONUC is launched in the DRC.
In 2001 Laurant-Désiré Kabila is assassinated and replaced by his son Joseph Kabila.
In 2002 a peace agreement in Sun City, South Africa, is signed.
In 2003 a transitional government is appointed for the DRC.

In order to express these facts about the Second Congo War within the semantics of CML, we now identify six corresponding propositions.

\[ p_0: \text{An offensive is launched against the Kabila government.} \]
\[ p_1: \text{The rebel group RCD splits into two factions, the RCD-Goma and the RCD-Kisangani.} \]
\[ p_2: \text{The peace keeping mission MONUC is launched in the DRC.} \]
\[ p_3: \text{Laurant-Désiré Kabila is assassinated and replaced by his son Joseph Kabila.} \]
\[ p_4: \text{A peace agreement in Sun City, South Africa, is signed.} \]
\[ p_5: \text{A transitional government is appointed for the DRC.} \]

As expressed by the facts, each proposition is true in at least one of the conflict states of each time point. We assume that the states \( w_0, w_1, w_2, w_3, w_4 \) and \( w_5 \) represent the states in which the six propositions are true, respectively. The states \( w_6, w_7, w_8, w_9, \) and \( w_{10} \) represent the states in which the Second Congo War would have been if the corresponding propositions had not been true.

The truth of the propositions is expressed by the value assignment function \( v \) as follows.

\[
\begin{align*}
    v(w_0) &= \{p_0\} \\
    v(w_1) &= \{p_1\} \\
    v(w_2) &= \{p_2\}
\end{align*}
\]
The reconstruction shows, for instance, that the Second Congo War could have been in the state $w_{10}$ instead of $w_5$ or that the proposition $p_5$ expressing that a peace agreement in Sun City, South Africa, is signed is not true at $w_{10}$. Hence, $w_{10}$ represents the possible state in which the Second Congo War would have been if the Sun City peace agreement had not been signed. Similar considerations can be made for the other states and propositions. Each state is characterised by the complete set of propositions that are true at that state.

The fifth component of CML-structures is the set of agents involved in the conflict. In our example we only look at three agents.

$a_0$: Rwanda-aligned forces  
$a_1$: UN  
$a_2$: Kabila-aligned forces  
$A = \{a_0, a_1, a_2\}$

Finally, we have to specify the four propositional attitude functions representing the agents’ beliefs, goals, norms, and emotions in the context of the Second Congo War. We assume that the functions are given as follows.

$b(w_2, a_0) = b(w_2, a_1) = b(w_2, a_2) = b(w_3, a_0) = b(w_3, a_1) = b(w_3, a_2) = b(w_4, a_0) = b(w_4, a_1) = b(w_4, a_2) = \{p_2\}$  
$g(a_1, w_1) = g(a_1, w_2) = \{p_2\}$  
$e(a_0, w_3) = \{p_1\}$
According to the above specification of the propositional attitude functions, every agent believes that the propositions $p_2$ and $p_3$ are true, i.e. every agent believes from 2000 on that the peace keeping mission MONUC has been launched in the DRC, and every agent believes from 2001 on that Laurant-Désiré Kabila has been assassinated and replaced by his son Joseph Kabila.

The goal function shows that the goal of launching a peacekeeping mission in DRC was pursued by the UN in 1999 and then again in 2000, the year in which the mission was actually launched.\footnote{MONUC 2008}

The fact that the Rwanda-aligned forces were emotionally excited by Laurant-Désiré Kabila’s assassination and his replacement by his son Joseph Kabila is expressed by the emotion function $e$.

The CML-structure reconstructed so far expresses various aspects of the Second Congo War. Furthermore, it allows one to determine the truth value of formulae expressing statements made about the conflict.

### 3.8 Summary

The aim of this chapter was to introduce the semantics of CML. This has been achieved to the extent that we have defined the general structures within which CML-formulae can be evaluated and have stated the truth conditions for the operators of the language of CML. All relevant aspects of a conflict are represented by the components of CML-structures. In particular, a CML-structure reflects the temporal and modal dimensions of...
a conflict, the agents involved in it, and the goals, beliefs, norms, and emotions that constitute the conflict.

Having defined the semantics of CML, we can now interpret statements about conflicts in a uniform way by looking at the conditions under which they are true. Furthermore, we can check whether a set of CML-formulae is satisfiable, and, hence, whether the statements expressed by the set can simultaneously be true. This gives us a first means of checking whether a set of CML-formulae represents a conflict or not. Having a uniform and formal way of interpreting statements about a conflict is the second step in the process of modelling and resolving it.

In the following chapter, we will further analyse the notion of inconsistency by axiomatising CML and showing that it is equivalent to the notion of unsatisifiability. Once we have precisely defined what it means for a set of CML-formulae to be inconsistent, we will continue with a general conflict definition.
CHAPTER 4

Axiomatising Conflicts

The Axiomatics of Conflict Modelling Logic

4.1 Introduction

The aim of chapter 4 is to state an axiom system for CML and prove that the system is sound and complete over the semantics of CML. In order to axiomatise CML, we single out the modal fragment of CML, the temporal fragment of CML, and the propositional attitude fragment of CML, and then treat them individually. For the first two fragments we provide axiom systems. For the propositional attitude fragment, we only provide a list of ‘conditions’ as it is not a logical system. Finally, we show that the axioms of the temporal fragment also provide an axiomatisation of CML as a whole.

As a result we will be able to answer to the following questions.

- What are the axioms characterising the modal fragment of CML?
- What are the axioms characterising the temporal fragment of CML?
- Which conditions does an agent need to satisfy in order to qualify as a consistent believer?
Chapter 4 Axiomatizing Conflicts

- Why do the axioms of the temporal fragment provide an axiomatisation of CML as a whole?

An outline of the chapter is as follows: In section 4.2, we present an axiomatisation of the modal fragment. Before we state its axioms (4.2.2), we provide a semantic characterisation of the fragment and prove its equivalence to the modal system S5 (4.2.1). Once the equivalence between the two systems is proven, it is easy to prove soundness (4.2.3) and completeness (4.2.4) as we can assume the corresponding theorems for S5. In section 4.3, we axiomatise the temporal fragment. Again, we first provide a semantic characterisation of the fragment (4.3.1) followed by its axiom system (4.3.2) which we prove to be sound (4.3.3) and complete (4.3.4). Section 4.4 deals with the propositional attitude fragment of CML. As this fragment is not a logical system, we just state a number of ‘conditions’ for one of the attitude functions. The conditions, which we state exemplarily for the belief function b in (4.4.1), characterise different types of consistent believers. In section 4.5, we provide an axiomatisation of CML as a whole by showing that the axiomatisation of the temporal fragment, indeed, provides an axiomatisation of the whole system.

4.2 The Modal Fragment

The modal fragment of CML, □-CML, is constituted by the syntax and semantics of those CML-formulae which can be used to express modal properties of a conflict. Formulae contained in the modal fragment of CML are referred to as □-formulae. The set of all □-formulae is called □-FOR. Obviously, □-FOR is a subset of the set of all CML-formulae, i.e. □-FOR ⊆ FOR.
□-formulae are formed by propositional constants, the logical connectives, and the necessity operator □.

**Definition 46 (Modal Fragment)**
□-FOR is the set of all strings of elements of SYM satisfying the following BNF:
φ = :: p | ¬φ | (φ ∧ χ) | □φ, where p ∈ p-CON.

As the formation rules for □-formulae are identical with the formation rules for formulae of propositional modal logic, the modal fragment of CML is syntactically equivalent with the set of all formulae of propositional modal systems, such as S5. This syntactical identity is expressed in the following theorem.

**Theorem 6 (□-CML and S5)**
□-FOR = S5-FOR, where S5-FOR is the set of formulae of S5.

**Proof**
The formation rules for □-formulae are identical with the formation rules for S5-formulae. Hence, every □-formula can also be formed by the formation rules for S5-formulae and every S5-formula can be formed by the formation rules for □-formulae. And so, every □-formula is also an S5-formula and vice versa.

QED

Every CML-formula φ can be reduced to a □-formula φ* which we call the modal core of φ. The modal core does not contain operators other than the logical connectives and the necessity operator □. We obtain the modal core φ* of a CML-formula φ by consecutively translating sub-formulae of φ according to the principles stated in the following definition.

**Definition 47 (Modal Core)**
The modal core φ* of a CML-formula φ is obtained by the following rules:
(1) (p)* = p;
(2) (¬φ)* = ¬(φ)*;
(3) (φ ∧ ψ)* = (φ)* ∧ (ψ)*;
(4) (□φ)* = □(φ)*;
(5) (∀φ)* = (Rτ_0φ)* = (τ_0 < τ_1)* = (τ_2 = τ_3)* = (Oαφ)* = p_n, where p_n is a fresh propositional constant and O ∈ {B, G, N, E}.
According to condition five, all non-logical and non-modal operators are eliminated. Every sub-formula whose main operator is not a logical or modal operator is replaced by a new propositional constant $p_n$. As a result, the modal core $\varphi^*$ of a CML-formula $\varphi$ may contain different, and possibly more, propositional constants than the original formula $\varphi$.

**Theorem 7 (Modal Core and □-FOR)**
The modal core $\varphi^*$ of every CML-formula $\varphi$ is a □-formula.

**Proof**
The proof is a simple induction on the formation of CML-formulae.

QED

### 4.2.1 Semantic Characterisation of □-CML

In the previous section, we singled out a certain set of CML-formulae as the modal fragment of CML and showed how arbitrary CML-formulae can be reduced to their modal core. Now, we look at the semantic side of □-CML. We identify the parts of CML-structures necessary for providing an interpretation of □-formulae. Then, we prove that the set of CML-valid □-formulae is identical with the set of valid S5-formulae. This result, expressed in Theorem 13, enables us to use the proof theory of S5 for the modal fragment of CML.

If we want to evaluate a □-formula $\varphi$ within a CML-structure $x$, we only need particular information about $x$. Information about the agents’ attitudes, encoded by the four propositional attitude functions, for instance, is irrelevant as there are no agent symbols in $\varphi$. In order to interpret □-formulae, we can, therefore, reduce CML-structures to

---

163 Consider, for example, the formula $B\alpha p \land \neg \square q$ and its modal core $(B\alpha p \land \neg \square q)^* = r \land \neg \square q$. 

substructures containing only those components necessary for the evaluation of □-formulae. We call such substructures □-structures.

As the truth condition for □ depends only on the set of possible conflict states W, the truth assignment function v, and the set of histories H which, in turn, depends on the set of time points T and the binary relation <, we can define □-structures as substructures of CML-structures containing only the four components T, <, W, v, and H.

**Definition 48 (□-Structure)**

A □-structure x is a structure of the form <T, <, W, H, v>, where:

1. T is a set of time points;
2. < is a binary relation on T as defined in Definition 23;
3. W is a set of possible conflict states;
4. v: W \rightarrow ϕ (p-CON) is a truth assignment function;
5. H is a set of possible conflict histories as defined in Definition 26.

For a given CML-structure x, we can obtain a corresponding □-structure x* by removing the modally irrelevant components A, b, g, n, e, and μ. If a formula is valid in every □-structure, we say that the formula is valid in □-CML. Due to the substructure lemma, every □-formula φ is valid in CML, i.e. \( \models_{\text{CML}} \varphi \), iff φ is also valid in □-CML, i.e. \( \models_{\square} \varphi \).

**Theorem 8 (CML-Validity and □-Validity)**

\( \models_{\text{CML}} \varphi \) iff \( \models_{\square} \varphi \) for all □-formulae φ.

**Proof**

Let φ be a □-formula. Assume \( \not\models_{\square} \varphi \). Then, there is a □-structure x and a state w such that φ \( \not\in v(w) \). Expand x to a CML-structure x* by adding an arbitrary, nonempty set A of agents, arbitrary propositional attitude functions b, g, n, e and an arbitrary term assignment function μ. Then, the extended structure x* is a CML-structure and φ \( \not\in v(w) \). Hence, \( \not\models_{\text{CML}} \varphi \).

Assume now \( \not\models_{\text{CML}} \varphi \). Then, there is a CML-structure x and a state w of x such that φ \( \not\in v_j(w) \). Reduce x to a □-structure x* by eliminating the set of agents A, the propositional attitude functions b, g, n and e, and the term assignment function μ. Then, x* is a □-structure and φ \( \not\in v(w) \). Hence, \( \not\models_{\square} \varphi \).

Altogether, we get \( \not\models_{\square} \varphi \) iff \( \not\models_{\text{CML}} \varphi \) and by contraposition \( \models_{\square} \varphi \) iff \( \models_{\text{CML}} \varphi \).

QED
The main theorem of this section states that every □-formula that is valid in □-CML is also valid in S5 and every S5-valid formula is valid in □-CML. Hence, we can characterise the modal fragment of CML as the modal system S5. To prove this claim, we define two functions \( f_1 \) and \( f_2 \) which link □-structures and finite S5-structures. \( f_1 \) maps □-structures to S5-structures and \( f_2 \) assigns □-structures to finite S5-structures. \( f_1 \) is defined in such a way that the \( f_1 \)-image of a □-structure is invariant with respect to the truth values of □-formulae, i.e. if a □-formula \( \varphi \) is true/false at a state \( w \) of a □-structure \( x \), then there is a possible world \( w^* \) in the universe of the S5-structure \( f_1(x) \) at which \( \varphi \) is true/false. The same is true in the opposite direction for \( f_2 \), i.e. if an S5-formula \( \varphi \) is true/false at \( w \) in a finite S5-model \( x \), then there is a possible conflict state \( w^* \) in the universe of the □-structure \( f_2(x) \) at which \( \varphi \) is true/false.

First, we define \( f_1 \), show that the \( f_1 \)-image of a □-structure is an S5-structure, and prove the truth value preservation property of \( f_1 \). \( f_1 \) is defined as follows.

**Definition 49 (f_1)**

\( f_1 \) is a function from the set of □-structures to the set of finite S5-structures defined by

\[
 f_1(<T, \prec, W, H, v>) = <W^*, R^*, v^*>,
\]

where:

1. \( W^* = W \);  
2. \( w, R^* w \) iff there are \( h_0, h_1 \in H \), such that \( h_0(|w_0|) = w_0, h_1(|w_1|) = w_1 \), and \( h_0(|w_0| - 1) = h_1(|w_0| - 1) \);  
3. \( v^*(w) = v(w) \).

The following theorem expresses that the \( f_1 \)-image of a □-structure is an S5-structure. In particular, \( R^* \) is an equivalence relation on \( W^* \).

**Theorem 9 (f_1-Images and S5-Structures)**

\( f_1(x) \) is an S5-structure for every □-structure \( x \).

**Proof**

Let \( x = <T, \prec, W, H, v> \) be a □-structure and \( <W^*, R^*, v^*> \) the \( f_1 \)-image of \( x \), i.e. \( f_1(x) = <W^*, R^*, v^*> \). To show that \( <W^*, R^*, v^*> \) is an S5 structure, we have to prove that \( W^* \) is a nonempty set, \( R^* \) is an equivalence relation on \( W^* \), and \( v^* \) is a function from \( W^* \) into \( \wp(p\text{-CON}) \).
W* is a nonempty set as W* = W and W is a nonempty set. To prove that R* is an equivalence relation on W* we have to show that R* is reflexive, transitive and symmetrical.

(Reflexivity) Let w ∈ W*. Then, w ∈ W, as W* = W, so there is an h₀ ∈ H such that h₀(|w|) = w by Definition 26. Let now h₁ = h₀. Therefore, we have h₀, h₁ ∈ H such that h₀(|w|) = w₁, h₁(|w|) = w₁ and h₁(|w| |-1) = h₁(|w| |-1). Hence, wR* w₁.

(Transitivity) Let w₀, w₁, w₂ ∈ W*. Then, w₀, w₁, w₂ ∈ W, as W* = W and there are h₀, h₁, h₂ ∈ H such that h₀(|w₀|) = w₀, h₁(|w₀| |-1) = h₁(|w₀| |-1), h₀(|w₁| |-1) = w₁, h₂(|w₁| |-1) = h₂(|w₁| |-1). As h₁(|w₀|) = h₂(|w₁|), we get h₁(|w₀| |-1) = h₂(|w₀| |-1) by condition iii) of Definition 26. Then, there are h₀, h₁, h₁ ∈ H such that h₀(|w₀|) = w₀, h₁(|w₀| |-1) = w₁ and h₀(|w₀| |-1) = h₂(|w₀| |-1). Hence, we have w₀R* w₂.

(Symmetry) Let w₀, w₁ ∈ W* Assume w₀R* w₁. Then, there are h₀, h₁ ∈ H such that h₀(|w₀|) = w₀, h₀(|w₀| |-1) = w₁ and h₀(|w₀| |-1) = h₀(|w₀| |-1). Hence, we have w₁R* w₀.

v* is a function from W* into ϕ (p-CON), as v* = v and W* = W and v is a function from W into ϕ (p-CON).

QED

Having proven that f_i maps □-structures to S5-structure, we now show that f_i is truth value preserving with respect to □-formulae.

Theorem 10 (Truth Value Preservation Property of f_i)

Let χ = <T, <, W, H, v> be a □-structure and f_i(χ) = <W*, R* v*> be the f_i-image of χ. Then, ψ ∈ v(w) iff ψ ∈ v*(w) for every □-formula ψ and every w ∈ W.

Proof

Let χ = <T, <, W, H, v> be a □-structure, w ∈ W, f_i(χ) = <W*, R* v*>, and ψ an arbitrary □-formula. The proof now continues as an induction on the formation of ψ.

(p) Assume p ∈ v(w). Then, w ∈ W*, as W = W* and p ∈ v*(w), as v(w) = v*(w).

Assume now p ∈ v*(w). Then, w ∈ W, as W = W* and p ∈ v(w), as v(w) = v*(w). Hence, p ∈ v(w) iff p ∈ v*(w).

(Induction Hypothesis) We now assume ψ ∈ v(w) iff ψ ∈ v*(w) for every sub-formulae ψ of ϕ.

(¬ψ) Assume ¬ψ ∈ v(w). Then, ψ ∈ v(w). Hence, ψ ∈ v*(w) by the induction hypothesis, and so ¬¬ψ ∈ v*(w) by the truth conditions of S5.

Assume now ¬ψ ∈ v*(w). Then, ψ ∈ v*(w) by the truth conditions of S5. Hence, ψ ∈ v(w) by the induction hypothesis, and so ¬ψ ∈ v(w). Hence, ¬ψ ∈ v(w) iff ¬ψ ∈ v*(w).
(ψ ∧ χ) Assume ψ ∧ χ ∈ v(w). Then, ψ ∈ v(w) and χ ∈ v(w). Therefore, ψ ∈ v*(w) and χ ∈ v*(w) by the induction hypothesis, and hence, ψ ∧ χ ∈ v*(w) by the truth conditions of S5.

Assume now ψ ∧ χ ∈ v*(w). Then, ψ ∈ v*(w) and χ ∈ v*(w) by the truth conditions of S5. Therefore, ψ ∈ v*(w) and χ ∈ v*(w) by the induction hypothesis, and hence, ψ ∧ χ ∈ v*(w). Hence, ψ ∧ χ ∈ v(w) iff ψ ∧ χ ∈ v*(w).

(□ψ) Assume □ψ ∈ v(w0). Then, ψ ∈ v(wi) for all wi such that there are ho, hi ∈ H with h0(|w0|) = w0, hi(|w0| − 1) = w1 and h0(|w0| − 1) = h1(|w0| − 1). Thus, ψ ∈ v*(w0) for all wi such that there are ho, hi ∈ H with h0(|w0|) = w0, hi(|w0|) = |w0| − 1 and h0(|w0| − 1) = h1(|w0| − 1) by the induction hypothesis. Hence, ψ ∈ v*(w0) for all wi such that w0R*wi, and so □ψ ∈ v*(w0). Therefore, □ψ ∈ v(w0) iff □ψ ∈ v*(w0).

QED

Now we turn to the definition of f2.

Definition 50 (f2)

f2 is a function from the set of finite S5-structures to the set of □-structures defined by f2(<W*, R*, v*>) = <T, <, W, H, v>, where:

1. T = {ti | i ∈ Z} is a set of time points, one for each integer i ∈ Z;
2. ti < tj iff i < j;
3. W = K0 ∪ K0 ∪ K1 ∪ K1 where:
   i. K0 = {w0} | i ∈ Z is a set of conflict states, one for each negative integer i ∈ Z;
   ii. K0 = {w0} | [w*] ∈ W*/R* is a set of conflict states, one for each equivalence class [w*] ∈ W*/R*;
   iii. K1 = W*;
   iv. K1 = {w1} | w* ∈ W* and i ∈ Z∗ \{1} is a set of conflict states, one for each pair consisting of an element w* of W* and a positive integer i ∈ Z∗ \{1};
4. H = {hwi} is a set of functions from T into W, one for each w* ∈ W* such that:
   i. hwi(ti) = w0 if i ∈ Z;
   ii. h0(ti) = w0 if i = 0;
   iii. h0(ti) = w* if i = 1;
   iv. h1(si) = w1 if i ∈ Z∗ \{1};
5. v(w) = v*(w) if w ∈ W*, and ∅ if w ∉ W*.
The definition of $f_2$ encodes a rule telling us how to construct a $\Box$-structure on the basis of a finite S5-structure. This construction works only for finite S5-structures as it assumes that there is only a finite number of equivalence classes $[w^*]$ in the quotient set $W^* \setminus R^*$. The $\Box$-structure obtained from a finite S5-structure by applying $f_2$ can be described as follows: the elements of the original set $W^* = K_1$ are stacked on top of each other at the time point $t_1$. Elements of an equivalence class $[w^*] \in W^* \setminus R^*$ are all linked by a common, unique, immediate predecessor $w_{[w^*]} \in K_0$, occurring at the time point $t_0$.

At time points earlier than $t_0$, the structure consists of a linear order constituted by elements of $K_1$. At time points later than $t_0$, the structure consists of parallel, linear orders constituted by elements of $K_1$.

Graphically, we can illustrate the construction as follows.

![Diagram of transformation from finite S5-structure to $\Box$-structure]

Figure 12: Transformation of a Finite S5-Structure into a $\Box$-Structure

---

164 In fact, the construction also works for infinite S5-structures with the property that their quotient set $W \setminus R$ is finite.
The following theorem expresses that \( f_2 \)-image of a finite S5-structure \( \langle W^*, R^*, v^* \rangle \) is a \( \Box \)-structure.

**Theorem 11** (\( f_2 \)-Images and \( \Box \)-Structures)

\( f_2(\chi) \) is a \( \Box \)-structure for every finite S5-structure \( \chi \).

**Proof**

Let \( \chi = \langle W^*, R^*, v^* \rangle \) be a finite S5-structure and \( \langle T, \prec, W, H, v \rangle \) be the \( f_2 \)-image of \( \chi \), i.e. \( f_2(\chi) = \langle T, \prec, W, H, v \rangle \). To show that \( \langle T, \prec, W, H, v \rangle \) is a \( \Box \)-structure, we have to prove that \( (T, \prec) \) is a discrete, strict linear order without endpoints, \( W \) is a set, \( H \) is a set of functions as defined in Definition 26, and \( v \) is a function from \( W \) into \( \wp(p \cdot \text{CON}) \).

\( (T, \prec) \) is a discrete, strict linear order without endpoints as it is by definition isomorphic to \( (\mathbb{Z}, \prec) \) and \( (\mathbb{Z}, \prec) \) is a discrete, strict linear order without endpoints.

\( W \) is a nonempty set since \( W^* \subseteq W \), by definition, and \( W^* \) is a nonempty set by definition of an S5-structure.

By definition \( H \) is a set of functions from \( T \) into \( W \). To prove that \( H \) is a set of functions according to Definition 26 we have to show that elements of \( H \) satisfy the three conditions stated in the definition.

(i) Let \( h_{w^*_i}, h_{w^*_1} \in H, t_n, t_m \in T \) and \( h_{w^*_i}(t_n) = h_{w^*_i}(t_m) = w \). Then, \( w \in W \) as \( h_{w^*_i}, h_{w^*_1} \) are functions from \( T \) into \( W \). Then, \( w \in K_0 \cup K_0 \cup K_1 \cup K_\Box \) by definition of \( W \). Consider now the following four cases:

- **(w \in K_*)** If \( w \in K_* \), then \( h_{w^*_i}(t_n) = h_{w^*_i}(t_m) = w^*_i \) for some \( i \in \mathbb{Z}^* \) by definition of \( K_* \). By definition of \( H \) we get \( n = m = i \), and, hence, \( t_n = t_m \).

- **(w \in K_0)** If \( w \in K_0 \), then \( n = m = 0 \) by definition of \( H \). Hence, \( t_n = t_m \).

- **(w \in K_1)** If \( w \in K_1 \), then \( n = m = 1 \) by definition of \( H \). Hence, \( t_n = t_m \).

- **(w \in K_\Box)** If \( w \in K_\Box \), then \( h_{w^*_i}(t_n) = h_{w^*_i}(t_m) = w^*_i \) for some \( i \in \mathbb{Z}^* \setminus \{1\} \) by definition of \( K_\Box \). By definition of \( H \) we get \( n = m = i \), and, hence, \( t_n = t_m \).

(ii) Let \( w \in W \). Then, \( w \in K_0 \cup K_0 \cup K_1 \cup K_\Box \) by definition of \( W \). Consider now the following four cases:

- **(w \in K_*)** If \( w \in K_* \), then \( w = w^*_i \) for some \( i \in \mathbb{Z}^* \) by definition of \( K_* \). Then, \( h_{w^*_i}(t) = w \) for all \( h_{w^*_i} \in H \) by definition of \( H \). Hence, there is an \( h_{w^*_i} \in H \) and \( t_i \in T \) such that \( h_{w^*_i}(t_i) = w \).

- **(w \in K_0)** If \( w \in K_0 \), then \( w = w_{i+1} \) for some \( w^* \in W^* \) by definition of \( K_0 \). Then, \( h_{w^*_i}(t_0) = w \) by definition of \( H \). Hence, there is an \( h_{w^*_i} \in H \) and \( t_0 \in T \) such that \( h_{w^*_i}(t_0) = w \).

- **(w \in K_1)** If \( w \in K_1 \), then \( w = w^*_i \) for some \( w^* \in W^* \) by definition of \( K_1 \). Then, \( h_{w^*_i}(t) = w \) by definition of \( H \). Hence, there is an \( h_{w^*_i} \in H \) and \( t_i \in T \) such that \( h_{w^*_i}(t_i) = w \).

- **(w \in K_\Box)** If \( w \in K_\Box \), then \( w = w_{i+1} \) for some \( w^* \in W^* \) and \( i \in \mathbb{Z}^* \setminus \{1\} \) by definition of \( K_\Box \). Then, \( h_{w^*_i}(t) = w \) by definition of \( H \). Hence, there is an \( h_{w^*_i} \in H \) and \( t_i \in T \) such that \( h_{w^*_i}(t_i) = w \).

(iii) Let \( h_{w^*_i}, h_{w^*_1} \in H \) with \( h_{w^*_i} \neq h_{w^*_1} \). Consider now the following two cases:
(h_{w_i^*}(t_0) = h_{w_i^*}(t_0)) If h_{w_i^*}(t_0) = h_{w_i^*}(t_0), then h_{w_i^*}(t_0) = h_{w_i^*}(t_0) = w_0 for all i ∈ Z^+ by definition of H. Hence, h_{w_i^*}(t_0) = h_{w_i^*}(t_0) for all t_0 < t_0 by definition of T and <. By definition of H we also get h_{w_i^*}(t_0) = w_0^* ≠ w_i^* = h_{w_i^*}(t_0) and h_{w_i^*}(t_i) = w_{(w_i^*, 0)} ≠ w_{(w_i^*, 0)} = h_{w_i^*}(t_i) for all i ∈ Z^+ \ {1}. Hence, h_{w_i^*}(t_i) ≠ h_{w_i^*}(t_i) for all t_i < t_i by definition of T and <. Altogether, there is a t_0 ∈ H such that h_{w_i^*}(t_0) = h_{w_i^*}(t_0) for all t_0 < t_0 and h_{w_i^*}(t_i) ≠ h_{w_i^*}(t_i) for all t_i < t_i.

(v is a function from W into ψ(p-CON) as v is identical with v* for all w ∈ W* and v* is a function from W into ψ(p-CON) by definition of S5-structures and the v-image of every w ∈ W \ W* is the empty set Ø which is trivially an element of ψ(p-CON).

QED

Having proven that the f_2-image of a finite S5-structure x is indeed a □-structure, we now show that f_2 is truth value preserving, i.e. the truth value of an S5-formula p at w in a given S5-structure x is equal to the truth value of p at w in the □-structure f_2(x). Note that every element w of an S5-structure x also appears in the □-structure f_2(x) because W* = K_1 ⊆ W by definition of f_2.

Theorem 12 (Truth Value Preservation Property of f_2)

Let x = <W*, R* v*> be a finite S5-structure and f_2(x) = <T, <, W, H, v> the f_2-image of x. Then, for every S5-formula p and every w ∈ W*, we have p ∈ v*(w) iff p ∈ v(w).

Proof

Let x = <W*, R* v*> be an S5-structure, w ∈ W*, f_2(x) = <T, <, W, H, v> the f_2-image of x and p an arbitrary S5-formula. The proof now continues as an induction on the formation of p.

(p) Assume p ∈ v*(w). Then, w ∈ W, as W* = K_1 ⊆ W. Then, p ∈ v(w), as v(w) = v*(w) for all w ∈ W*. Assume now p ∈ v(w). Then, p ∈ v*(w), as v(w) = v*(w) for all w ∈ W*. Hence, p ∈ v*(w) iff p ∈ v(w).

(Induction Hypothesis) Assume ψ ∈ v*(w) iff ψ ∈ v(w) for all subformulae ψ of p.

(¬p) Assume ¬ψ ∈ v*(w). Then, ψ ∉ v*(w) by the truth conditions of S5. Then, ψ ∉ v(w) by the induction hypothesis. Then, ¬ψ ∈ v(w). Assume now ¬ψ
We can now prove the semantic equivalence between $\Box$-CML and S5.

**Theorem 13**  (Semantic Equivalence between $\Box$-CML and S5)

$\vdash_\Box \varphi \iff \vdash_{S5} \varphi$ for all $\Box$-formulae $\varphi$.

**Proof**

Let $\varphi$ be a $\Box$-formula. We prove the by contraposition. Assume $\varphi$ is not valid in S5. Then, there is a finite S5-structures $\mathcal{X} = <W^*, R^*, v^*>$ and a $w \in W^*$, such that $\varphi \not\in v^*(w)$, as S5 has the finite model property. Then, $f_\varphi(\mathcal{X}) = <T, \prec, W, H, v>$ is a $\Box$-structure by Theorem 11 and $\varphi \not\in v(w)$ by Theorem 12. Hence, $\varphi$ is not valid in $\Box$-CML.

Assume now $\varphi$ is not valid in $\Box$-CML. Then, there is a $\Box$-structure $\mathcal{X} = <T, \prec, W, H, v>$ and a $w \in W$, such that $\varphi \not\in v(w)$. Then, $f_\varphi(\mathcal{X}) = <W^*, R^*, v^*>$ is an S5-structure by Theorem 9 and $\varphi \not\in v(w)$ by Theorem 10. Hence, $\varphi$ is not valid in S5.

Altogether we get $\varphi$ is not valid in S5 iff $\varphi$ is not valid in $\Box$-CML and by contraposition $\varphi$ is valid in S5 iff $\varphi$ is valid in $\Box$-CML.

QED
4.2.2 Axioms for □-CML

Having shown that □-CML is semantically equivalent to S5 we can exploit the fact that S5 is adequately characterised by a simple axiom system. We can use the S5 axioms as an axiomatic characterisation of □-CML. Following our notational convention, we call the axioms characterising □-CML □-axioms.

**Definition 51 (□-Axioms)**

- (PC) Axioms of propositional logic;
- (K) □(φ ⊃ ψ) ⊃ (□φ ⊃ □ψ);   
- (T) □φ ⊃ φ;
- (E) ◊φ ⊃ □◊φ;
- (MP) ⊢ φ ⊃ ψ, φ ⊢ ψ;
- (N) ⊢ φ ⊃ □ φ.

Provability is defined in the usual way.

**Definition 52 (□-Provability)**

Let φ be a □-formula. Then, φ is □-provable, i.e. ⊢ □ φ, iff there is a finite sequence of □-formulae such that the last formula of the sequence is φ and every other formula in the sequence is either a □-axiom or the result of applying a rule of inference to formulae occurring earlier in the sequence.

Having defined the concept of □-provability, we can define □-inconsistency and the syntactic consequence relation. First, we define what it means for a □-formula φ to follow syntactically from a set of □-formulae Φ.

**Definition 53 (Syntactic Consequence Relation for □-CML)**

Let Φ be a set of □-formulae and φ be a t-formula. Then, the syntactic consequence relation holds between Φ and φ, i.e. Φ ⊩ □ φ, if and only if there is a finite sequence of □-formulae such that the last formula of the sequence is φ and every other formula in the sequence is either a □-axiom, an element of Φ, or the result of applying a rule of inference to formulae occurring earlier in the sequence.

Now we define the notion of □-inconsistency and, its complement, the notion of □-consistency.
**Definition 54 (□-Inconsistency/□-Consistency)**

Let $\Phi$ be a set of □-formulae and $\bot$ any □-formula of the form $\varphi \land \lnot \varphi$. Then, $\Phi$ is □-inconsistent if and only if $\Phi \vdash \Box \bot$ and $\Phi$ is □-consistent if and only if $\Phi$ is not □-inconsistent.

Before we prove soundness and completeness for □-CML, we provide a theorem expressing the proof theoretical equivalence between □-CML and S5. Its proof is trivial, as both systems have the same axiomatic proof theory, and, hence, every formula provable in S5 is also provable in □-CML and vice versa.

**Theorem 14 (Proof Theoretical Equivalence between □-CML and S5)**

□-CML is proof theoretically equivalent to S5 in the sense that for every □-formula/S5-formula $\varphi$: $\vdash \Box \varphi$ iff $\vdash_{S5} \varphi$.

**Proof**

Let $\varphi$ be a □-formula such that $\vdash \Box \varphi$. Then, there is a finite sequence of □-formulae such that the last formula of the sequence is $\varphi$ and every other formula in the sequence is either a □-axiom or the result of applying a rule of inference to formulae occurring earlier in the sequence. Then, there is a finite sequence of □-formulae such that the last formula of the sequence is $\varphi$ and every other formula in the sequence is either a S5-axiom or the result of applying a rule of inference to formulae occurring earlier in the sequence, because the □-axioms and □-rules of inference are identical with the S5-axioms and S5-rules of inference. Hence, $\vdash_{S5} \varphi$.

The reverse direction is proven analogously.

QED

**4.2.3 Soundness of □-CML**

Soundness of □-CML follows directly from the semantic and syntactic equivalence between □-CML and S5, and the soundness theorem for S5. To show that the axiom system of □-CML is sound, we have to prove that every □-formula provable in □-CML is □-valid. In the proof, we assume that S5 is sound and complete.\(^{165}\)

\(^{165}\) For a completeness proof of S5, see, for instance, (Hughes and Cresswell 1996, p. 119).
Theorem 15  (Soundness of □-CML)
If ⊢_□ φ, then ⊨_□ φ for all □-formulae φ.

Proof
Let φ be a □-formula such that ⊢_□ φ. Then, ⊩_S5 φ by Theorem 14. Then, ⊩_S5 φ by the soundness of S5. Then, ⊢_□ φ by Theorem 13.

QED

4.2.4 Completeness of □-CML

Completeness of □-CML follows from the semantic and syntactic equivalence between □-CML and S5, and the completeness theorem for S5. To show that the axiom system of □-CML is complete, we have to prove that every □-formula valid in □-CML is □-provable.

Theorem 16  (Completeness of □-CML)
If ⊩_□ φ, then ⊢_□ φ for all □-formulae φ.

Proof
Let φ be a □-formula such that ⊩_□ φ. Then, ⊩_S5 φ by Theorem 13. Then, ⊢_S5 φ by the completeness of S5. Then, ⊢_□ φ by Theorem 14.

QED

4.3 The Temporal Fragment

The temporal fragment of CML, t-CML, consists of those CML-formulae which can be used to express temporal properties of a conflict. We call such formulae t-formulae and refer to the set of all t-formulae by t-FOR.

t-formulae are formed by using only propositional constants, temporal constants and variables, the indexical symbol n, logical connectives and quantifiers, the identity symbol,
the temporal precedence symbol, and the temporal realisation operator as primitive
symbols.

**Definition 55 (Temporal Fragment)**
t-FOR is the set of all strings of elements of t-SYM satisfying the following BNF:
\[ \varphi = ::= p \mid \neg \varphi \mid \varphi \land \chi \mid \tau_0 < \tau_1 \mid \tau_0 = \tau_1 \mid R_{\tau_0} \varphi \mid \forall x \varphi, \text{where } p \in p-\text{CON}, \tau_0, \tau_1 \in t-\text{TER} \cup \{ n \}, x \in t-\text{VAR}. \]

Similar to the modal core, we can reconstruct the temporal core of CML-formulae by replacing sub-formulae whose main operator is a non-logical or non-temporal operator by new propositional constants. The temporal core \( \varphi^* \) represents just the temporal properties of \( \varphi \) and is obtained from \( \varphi \) by applying the translation principles stated in the following definition.

**Definition 56 (Temporal Core)**
The temporal core \( \varphi^* \) of a CML-formula \( \varphi \) is obtained by the following rules:

1. \( (p)^* = p \);
2. \( (\tau_0 < \tau_1)^* = \tau_0 < \tau_1 \);
3. \( (\tau_0 = \tau_1)^* = \tau_0 = \tau_1 \);
4. \( (\neg \varphi)^* = \neg (\varphi)^* \);
5. \( (\varphi \land \psi)^* = (\varphi)^* \land (\psi)^* \);
6. \( (\forall x \varphi)^* = \forall x (\varphi)^* \);
7. \( (R_{\tau_0} \varphi)^* = R_{\tau_0} (\varphi)^* \);
8. \( (\square \varphi)^* = (\forall y \varphi)^* = \alpha_0 = \alpha_1 = (O_{\alpha_0} \varphi)^* = p_n \) where \( p_n \) is a fresh propositional constant and \( O \in \{ B, G, N, E \} \).

The definition allows us to eliminate all the non-logical and non-temporal operators within a formula \( \varphi \). As a result, the temporal core \( \varphi^* \) of a CML-formula \( \varphi \) is always a t-formula.

**Theorem 17 (Temporal Core and t-CML)**
The temporal core \( \varphi^* \) of every CML-formula \( \varphi \) is a t-formula.

**Proof**
Induction on the formation of formulae.

QED
Before we continue with the semantic characterisation of t-CML, we restate the definitions for free and bound occurrences of t-variables and the indexical symbol n. We call the occurrence of a t-variable \( x \) bound if and only if \( x \) lies within the scope of a quantification of the form \( \forall x \) or \( \exists x \). The indexical symbol \( n \) is bound by the temporal realisation operator \( R \), i.e. the occurrence of \( n \) is bound if and only if \( n \) lies within the scope of a temporal realisation of the form \( R\tau \), where \( \tau \in t\text{-TER} \). Free occurrences of t-variables or \( n \) are occurrences that are not bound. As usual, in the case of a term-replacement \( \phi[\tau/x] \) or \( \phi[\tau/n] \), we replace \( x \) or \( n \) only at their free occurrences.

### 4.3.1 Semantic Characterisation of t-CML

In order to evaluate a t-formula \( \varphi \), we have to look only at those components of CML-structures which are relevant for the interpretation of t-symbols. We define t-structures as substructures of CML-structures containing only the components necessary for the interpretation of t-formulae.

**Definition 57 (t-Structure)**

A t-structure \( \mathcal{X} \) is a structure of the form \( <T, \prec, W, H, v, \mu> \), where:

1. \( T \) is a set of time points;
2. \( \prec \) is a binary relation on \( T \) as defined in Definition 23;
3. \( W \) is a set of possible conflict states;
4. \( H \) is a set of possible conflict histories as defined in Definition 26;
5. \( v: W \to \wp(p\text{-CON}) \) is a truth assignment function;
6. \( \mu: t\text{-TER} \to T \) is a term assignment function as defined in Definition 31.

As the definition shows, t-structures are substructures of CML-structures. They are CML-structures from which the set of agents \( A \), and the propositional attitude functions \( b, g, n \) and \( e \), have been removed.
With the truth conditions for the temporal operators, as stated in Definition 39, we can evaluate every t-formula $\varphi$ in a t-structure $x$. We use the expression $\models_t \varphi$, to express that a t-formula $\varphi$ is t-valid, i.e. $\varphi$ is true at all states $w$ of all t-structures $x$.

As t-structures are substructures of CML-structures, t-formulae are t-valid if and only if they are CML-valid.

**Theorem 18  (CML-Validity and t-Validity)**

$\models_t \varphi$ iff $\models_{CML} \varphi$ for all t-formulae $\varphi$.

**Proof**

We prove by reduction. Let $\varphi$ be a t-formula. Assume $\not\models_t \varphi$. Then, there is a t-structure $x$ and a state $w$ such that $\varphi \not\in v_\mu(w)$. Expand $x$ to a CML-structure $x^*$ by adding an arbitrary, nonempty set $A$ of agents and arbitrary propositional attitude functions $b, g, n, e$ and extending the domain of $\mu$ to a-TER assigning arbitrary members of $A$ to a-terms. Then, the extended structure $x^*$ is a CML-structure and $\varphi \not\in v_\mu(w)$. Hence, $\not\models_{CML} \varphi$.

Assume now $\not\models_{CML} \varphi$. Then, there is a CML-structure $x$ and a state $w$ of $x$ such that $\varphi \not\in v_\mu(w)$. Reduce $x$ to a t-structure $x^*$ by eliminating the set of Agents $A$ and the propositional attitude functions $b, g, n$ and $e$, and restricting the domain of $\mu$ to t-TER. Then, $x^*$ is a t-structure and $\varphi \not\in v_\mu(w)$. Hence, $\not\models_t \varphi$.

Altogether, we get $\not\models_t \varphi$ iff $\not\models_{CML} \varphi$ and by contraposition $\models_t \varphi$ iff $\models_{CML} \varphi$.

QED

In the case of the modal fragment $\Box$-CML, we were able to show that $\Box$-CML is equivalent to S5. For t-CML, the situation is different as there is no simple temporal system that t-CML is equivalent to. Thus, we provide our own axiom system for t-CML and show that the axiom system is sound and complete over the semantic of t-CML.

### 4.3.2 Axioms for t-CML

We call axioms characterising the temporal fragment of CML t-axioms. t-axioms fall into three categories. Being based on first-order logic with identity, the first group of t-axioms contains axioms of propositional logic, first-order logic and identity theory. Here, the
choice of the particular axioms and rules is irrelevant. The second group of t-axioms includes five axioms dealing with the temporal precedence relation <. The five order axioms characterise < as a discrete, strict linear order relation without endpoints. The third group contains t-axioms specific to t-CML. They reflect the forwards-branching, backwards-linear character of t-structures by providing axiomatic conditions for the logical connectives and quantifiers, the temporal realisation operator, and the indexical symbol n.

The list of t-axioms is given as follows.

**Definition 58 (t-Axioms)**

(FOL=) Axioms of first-order logic with identity;

(O1) \( \forall x \neg(x < x) \);

(O2) \( \forall x_0, x_1, x_2(x_0 < x_1 \land x_1 < x_2 \supset x_0 < x_2) \);

(O3) \( \forall x_0, x_i(x_0 < x_i \lor x_0 = x_i \lor x_1 < x_0) \);

(O4) \( \forall x_0 \exists x_1 \forall x_2(x_0 < x_1 \land \neg(x_0 < x_2 \land x_2 < x_1)) \);

(O5) \( \forall x_0 \exists x_1 \forall x_2(x_1 < x_0 \land \neg(x_1 < x_2 \land x_2 < x_0)) \);

(T1) \( \forall x(x \leq n \supset (Rx \neg \varphi \equiv \neg Rx \varphi)) \);

(T2) \( \forall x(n < x \supset (Rx \neg \varphi \supset \neg Rx \varphi)) \);

(T3) \( \forall x((Rx \varphi \land Rx \psi) \equiv Rx(\varphi \land \psi)) \);

(T4) \( R \tau \forall x \varphi \equiv \forall x R \tau \varphi \), where \( \tau \in t\text{-TER} \cup \{n\} \) and \( \tau \neq x \);

(T5) \( \forall x_0, x_i(n \leq x_0 \supset (Rx_i \varphi \supset R x_0 \varphi)) \);

(T6) \( \forall x_0, x_i(x_0 \leq n \supset (Rx_0 Rx_i \varphi \supset Rx_i \varphi)) \);

(T7) \( Rn \varphi \equiv \varphi \);

(T8) \( \exists x(x = n) \);

(T9) \( R \varphi \equiv R \varphi[\tau/n] \), where \( \tau \in t\text{-TER} \cup \{n\} \);

(∀) \( \vdash \varphi \supset \psi \Rightarrow \vdash \varphi \supset \forall x \psi \), where \( x \) is not free in \( \varphi \);

(MP) \( \vdash \varphi \supset \psi, \varphi \Rightarrow \vdash \psi \);

(∀R) \( \vdash \varphi \Rightarrow \vdash \forall x Rx \varphi \), where \( x \) and \( n \) are not free in \( \varphi \).

t-provability, the syntactic consequence relation, and t-consistency are defined as usual.

**Definition 59 (t-Provability)**

Let \( \varphi \) be a t-formula. Then, \( \varphi \) is t-provable, i.e. \( \vdash \varphi \), if and only if there is a finite sequence of t-formulae such that the last formula of the sequence is \( \varphi \) and every other formula in the sequence is either a t-axiom or the result of applying a rule of inference to formulae occurring earlier in the sequence.
Definition 60 (Syntactic Consequence Relation for t-CML)
Let \( \Phi \) be a set of t-formulae and \( \varphi \) be a t-formula. Then, the syntactic consequence relation holds between \( \Phi \) and \( \varphi \), i.e. \( \Phi \vdash \varphi \), if and only if there is a finite sequence of t-formulae such that the last formula of the sequence is \( \varphi \) and every other formula in the sequence is either a t-axiom, and element of \( \Phi \) or the result of applying a rule of inference to formulae occurring earlier in the sequence.

Definition 61 (t-Inconsistency/t-Consistency)
Let \( \Phi \) be a set of t-formulae and \( \bot \) any t-formula of the form \( \varphi \land \neg \varphi \). Then, \( \Phi \) is t-inconsistent if and only if \( \Phi \vdash \bot \) and \( \Phi \) is t-consistent if and only if \( \Phi \) is not t-inconsistent.

4.3.3 Soundness of t-CML

To show that the axiom system for t-CML is sound, i.e. \( \models \), \( \varphi \) if \( \vdash \varphi \), we have to prove that every t-axiom is t-valid and every rule of inference preserves t-validity. We omit the proofs for (FOL\(^-\)), (MP) and (U) as they are standard.

Theorem 19 (Validity of t-axioms)
If \( \varphi \) is one of the axioms O1, O2, O3, O4, O5 or T1 to T8, then \( \models \varphi \).

Proof
We prove by reductio.

(O1) Assume \( \not\models \varphi \), \( \forall x \neg(x < x) \). Then, there is a t-structure \( \mathcal{A} \) and a state \( w \) such that \( x, w \models \neg \forall x \neg(x < x) \). Then, there is a \( t \in T \) such that \( t \not< t \). This is a contradiction to the irreflexivity of \( \not< \). Hence, \( \models \forall x \neg(x < x) \).

(O2) Assume \( \not\models \varphi \), \( \forall x_0 x_1 x_2 (x_0 < x_1 \land x_1 < x_2 \supset x_0 < x_2) \). Then, there is a t-structure \( \mathcal{A} \) and a state \( w \) such that \( x, w \models \neg \forall x_0 x_1 x_2 (x_0 < x_1 \land x_1 < x_2 \supset x_0 < x_2) \). Then, there are \( t_0, t_1, t_2 \in T \) such that \( t_0 < t_1 \) and \( t_1 < t_2 \) and \( t_0 \not< t_2 \). This is a contradiction to the transitivity of \( \not< \). Hence, \( \models \forall x_0 x_1 x_2 (x_0 < x_1 \land x_1 < x_2 \supset x_0 < x_2) \).

(O3) Assume \( \not\models \varphi \), \( \forall x_0 x_1 (x_0 < x_1 \lor x_0 = x_1 \lor x_1 < x_0) \). Then, there is a t-structure \( \mathcal{A} \) and a state \( w \) such that \( x, w \models \neg \forall x_0 x_1 (x_0 < x_1 \lor x_0 = x_1 \lor x_1 < x_0) \). Then, there are \( t_0, t_1 \in T \) such that \( t_0 \not< t_1 \) and \( t_0 \not= t_1 \) and \( t_1 \not< t_0 \). This is a contradiction to the trichotomy of \( \not< \). Hence, \( \models \forall x_0 x_1 (x_0 < x_1 \lor x_0 = x_1 \lor x_1 < x_0) \).

(O4) Assume \( \not\models \varphi \), \( \forall x_0 \exists x_1 \forall x_2 (x_0 < x_1 \land \neg(x_0 < x_2 \land x_2 < x_1)) \). Then, there is a t-structure \( \mathcal{A} \) and a state \( w \) such that \( x, w \models \neg \forall x_0 \exists x_1 \forall x_2 (x_0 < x_1 \land \neg(x_0 < x_2 \land x_2 < x_1)) \). Then, there is a \( t_0 \in T \) such that there is no \( t_1 \in T \) such that \( t_0 < t_1 \) and there is no \( t_2 \in T \) such that \( t_0 < t_2 \) and \( t_2 < t_1 \). This is a contradiction to the
fact that every $t \in T$ has an immediate successor. Hence, $\models_{\mathcal{I}} \forall x_0 \exists x_1 \forall x_2(x_0 < x_1 \land \neg((x_0 < x_1 \land x_1 < x_2))$).

**O5** Assume $\mathcal{H}_I \models \forall x_0 \exists x_1 \forall x_2(x_1 < x_0 \land \neg((x_0 < x_1 \land x_1 < x_2 < x_0)))$. Then, there is a t-structure $x$ and a state $w$ such that $x, w \models \neg\forall x_0 \exists x_1 \forall x_2(x_1 < x_0 \land \neg((x_0 < x_1 \land x_1 < x_2 < x_0)))$. Then, there is a $t_0 \in T$ such that there is no $t_t \in T$ such that $t_t < t_0$ and there is no $t_2 \in T$ such that $t_t < t_2$ and $t_2 < t_0$. This is a contradiction to the fact that every $t \in T$ has an immediate predecessor. Hence, $\models_{\mathcal{I}} \forall x_0 \exists x_1 \forall x_2(x_1 < x_0 \land \neg((x_0 < x_1 \land x_1 < x_2 < x_0)))$.

**T1** Assume $\mathcal{H}_I \models \forall x(x \leq n \models (Rx \neg\varphi \equiv \neg Rx \varphi))$. Then, there is a t-structure $x$ and a state $w$ such that $x, w \models \neg\forall x(x \leq n \models (Rx \neg\varphi \equiv \neg Rx \varphi))$. Then, there is a $t \in T$ and a truth assignment function $v_\mu$ such that $t \not< |w|$ and one of the following two cases holds:

**i** $\neg\varphi \in v_\mu(h(t))$ for all $h \in H$ such that $h(|w|) = w$ and $\varphi \in v_\mu(h(t))$ for all $h \in H$ such that $h(|w|) = w$ by the truth condition for $R$. Then, $\varphi \not\in v_\mu(h(t))$ for all $h \in H$ by the truth condition for $\neg$. This is a contradiction to $\varphi \in v_\mu(h(t))$.

**ii** There is an $h_0 \in H$ such that $h_0(|w|) = w$ and $\neg\varphi \not\in v_\mu(h_0(t))$ and there is an $h \in H$ such that $h(|w|) = w$ and $\varphi \not\in v_\mu(h(t))$ by the truth condition for $R$. Then, $h_0(t) = h(t)$ by the definition of $H$ and, hence, by identity we get $\neg\varphi \not\in v_\mu(h_0(t))$. This is a contradiction to $\neg\varphi \not\in v_\mu(h_0(t))$.

As we get contradictions in both cases, $\models_{\mathcal{I}} \forall x(x \leq n \models (Rx \neg\varphi \equiv \neg Rx \varphi))$.

**T2** Assume $\mathcal{H}_I \models \forall x(n < x \models (Rx \neg\varphi \equiv \neg Rx \varphi))$. Then, there is a t-structure $x$ and a state $w$ such that $x, w \models \neg\forall x(n < x \models (Rx \neg\varphi \equiv \neg Rx \varphi))$. Then, there is a $t \in T$ and a truth assignment function $v_\mu$ such that $|w| < t$ and $\neg\varphi \in v_\mu(h(t))$ for all $h \in H$ such that $h(|w|) = w$ and $\varphi \in v_\mu(h(t))$ for all $h \in H$ such that $h(|w|) = w$ by the truth condition for $R$. Then, $\neg\varphi \not\in v_\mu(h(t))$ for all $h \in H$ by the truth condition for $\neg$. This is a contradiction to $\neg\varphi \in v_\mu(h(t))$. Hence, $\models_{\mathcal{I}} \forall x(n < x \models (Rx \neg\varphi \equiv \neg Rx \varphi))$.

**T3** Assume $\mathcal{H}_I \models \forall x(Rx \varphi \land Rx \psi \equiv Rx(\varphi \land \psi))$. Then, there is a t-structure $x$ and a state $w$ such that $x, w \models \neg\forall x(Rx \varphi \land Rx \psi \equiv Rx(\varphi \land \psi))$. Then, one of the following two cases holds:

**i** There is a $t \in T$ and a truth assignment function $v_\mu$ such that $\varphi \in v_\mu(h(t))$ for all $h \in H$ and $\psi \in v_\mu(h(t))$ for all $h \in H$ such that $h(|w|) = w$ and there is an $h \in H$ such that $h(|w|) = w$ and $\varphi \land \psi \not\in v_\mu(h(t))$ by the truth condition for $R$. Then, $\varphi \not\in v_\mu(h(t))$ or $\psi \not\in v_\mu(h(t))$, by the truth condition for $\land$, and, we have a contradiction.

**ii** There is a $t \in T$ and a truth assignment function $v_\mu$ such that $\varphi \land \psi \in v_\mu(h(t))$ for all $h \in H$ such that $h(|w|) = w$ and either there is an $h \in H$ such that $h(|w|) = w$ and $\varphi \not\in v_\mu(h(t))$ or there is an $h \in H$ such that $h(|w|) = w$ and $\psi \not\in v_\mu(h(t))$ by the truth condition for $R$. Then, $\varphi \in v_\mu(h(t))$ and $\psi \in v_\mu(h(t))$, by the truth condition for $\land$, we have a contradiction.

As we get contradictions in both cases, $\models_{\mathcal{I}} \forall x(Rx \varphi \land Rx \psi \equiv Rx(\varphi \land \psi))$.

**T4** Assume $\mathcal{H}_I, R \tau \forall x \varphi \equiv \forall x R \tau \varphi$, where $\tau \in t-TER \cup \{n\}$ and $t \not= x$. Then, there is a t-structure $x$ and a state $w$ such that $x, w \models \neg(R \tau \forall x \varphi \equiv \forall x R \tau \varphi)$. Then, one of the two following cases holds:

**i** There is a $t_0 \in T$ and a truth assignment function $v_\mu$ such that $\mu(t) = t_0$ and $\forall x \varphi \in v_\mu(h(t_0))$ for all $h \in H$ such that $h(|w|) = w$ by the truth condition
for \( R \) and \( \forall xR\tau \not\equiv v_\mu(w) \). Then, there is a \( t_1 \in T \) such that \( R\tau \not\equiv v_{\mu[t_1]}(w) \) by the truth condition for \( \forall \). Then, there is an \( h \in H \) such that \( \varphi \not\equiv v_{\mu[|w|]}(h(t)) \) by the truth condition for \( R \) and the fact that \( \mu(t) = t_0 \). On the other hand we have \( \varphi \in v_{\mu[|w|]}(h(t_0)) \) for all \( t_1 \in T \) by the truth condition for \( \forall \) and the fact that \( \forall x\varphi \in v_\mu(h(t_0)) \) and, thus, \( \varphi \in v_{\mu[|w|]}(h(t_0)) \). This is a contradiction.

(ii) There is a \( t_0 \in T \) and a truth assignment function \( v_\mu \) such that \( \mu(t) = t_0 \) and there is an \( h \in H \) such that \( h(|w|) = w \) and \( \forall x\varphi \not\equiv v_\mu(h(t_0)) \) by the truth condition for \( R \) and \( R\tau \in v_{\mu[|w|]}(w) \) for all \( t_1 \in T \) by the truth condition for \( \forall \). Then, \( \varphi \in v_{\mu[|w|]}(h(t_0)) \) for all \( t_1 \in T \) for all \( h \in H \) such that \( h(|w|) = w \) by the truth condition for \( R \) and there is a \( t_2 \in T \) such that \( \varphi \not\equiv v_{\mu[|w|]}(h(t_0)) \) by the truth condition for \( \forall \). Then, \( \varphi \in v_{\mu[|w|]}(h(t_0)) \) and, hence, we have a contradiction.

As we get contradictions in both cases, \( \not\models R\varphi \equiv \forall xR\varphi \).

(\( T5 \)) Assume \( \models \), \( \forall x_0x_i(n \leq x_0 \Rightarrow (Rx_0\varphi \supset Rx_0x_i\varphi)) \). Then, there is a t-structure \( x \) and a state \( w \) such that \( x_i \supset \forall x_0x_i(n \leq x_0 \Rightarrow (Rx_0\varphi \supset Rx_0x_i\varphi)) \). Then, there are \( t_0, t_1 \in T \) and a truth assignment function \( v_\mu \) such that \( |w| \leq t_0 \) and \( \varphi \in v_\mu(h_0(t_0)) \) for all \( h_0 \in H \) such that \( h_0(|w|) = w \) and there is an \( h_1 \in H \) such that \( h_1(|w|) = w \) and \( Rx_0\varphi \not\equiv v_\mu(h_1(t_0)) \) by the truth condition for \( R \). Then, there is an \( h_2 \in H \) such that \( h_2(|h_1(t_0)|) = h_2(t_0) \) and \( \varphi \not\equiv v_\mu(h_2(t_0)) \) by the truth condition for \( R \). Then, \( h_2(|w|) = h_2(|w|) \) by the definition of \( H \) and, hence, \( \varphi \in v_\mu(h_2(t_0)) \). This is a contradiction. Hence, \( \not\models \forall x_0x_i(n \leq x_0 \Rightarrow (Rx_0\varphi \supset Rx_0x_i\varphi)) \).

(\( T6 \)) Assume \( \models \), \( \forall x_0x_i(x_0 \leq n \Rightarrow (Rx_0x_i\varphi \supset Rx_i\varphi)) \). Then, there is a t-structure \( x \) and a state \( w \) such that \( x_i \supset \forall x_0x_i(x_0 \leq n \Rightarrow (Rx_0x_i\varphi \supset Rx_i\varphi)) \). Then, there are \( t_0, t_1 \in T \) and a truth assignment function \( v_\mu \) such that \( |w| \leq t_0 \) and \( Rx_0\varphi \in v_\mu(h_0(t_0)) \) for all \( h_0 \in H \) such that \( h_0(|w|) = w \) and there is an \( h_1 \in H \) such that \( h_1(|w|) = w \) and \( Rx_0\varphi \not\equiv v_\mu(h_1(t_0)) \) by the truth condition for \( R \). Then, \( Rx_0\varphi \in v_\mu(h_0(t_0)) \) and, hence, \( \varphi \in v_\mu(h_2(t_0)) \) for all \( h_2 \in H \) such that \( h_2(|h_1(t_0)|) = h_2(t_0) \) by the truth condition for \( R \). Then, \( \varphi \in v_\mu(h_2(t_0)) \). This is a contradiction. Hence, \( \not\models \forall x_0x_i(x_0 \leq n \Rightarrow (Rx_0x_i\varphi \supset Rx_i\varphi)) \).

(\( T7 \)) Assume \( \models \), \( R\varphi \equiv \varphi \). Then, there is a t-structure \( x \) and a state \( w \) such that \( x_i \supset \neg(R\varphi \equiv \varphi) \). Then, one of the two following cases holds:

(i) There is a truth assignment function \( v_\mu \) such that \( R\varphi \not\equiv v_\mu(w) \) and \( \varphi \not\equiv v_\mu(w) \). Then, \( \varphi \in v_\mu(w) \) by the truth condition for \( R \). Hence, we have a contradiction.

(ii) There is a truth assignment function \( v_\mu \) such that \( R\varphi \not\equiv v_\mu(w) \) and \( \varphi \in v_\mu(w) \). Then, \( \varphi \not\equiv v_\mu(w) \) by the truth condition for \( R \). Hence, we have a contradiction.

As we get contradictions in both cases, \( \models R\varphi \equiv \varphi \).

(\( T8 \)) Assume \( \models \), \( \exists x(x = n) \). Then, there is a t-structure \( x \) and a state \( w \) such that \( x_i \supset \exists x(x = n) \). Then, \( (x = n) \not\in v_{p[t]}(w) \) for all \( t \in T \) by the truth condition for \( \forall \). On the other hand we have \( (x = n) \in v_{\mu[|w|]}(w) \) for \( t = |w| \) by the definition of \( |w| \). This is a contradiction and, hence, \( \models \exists x(x = n) \).
Next, we show that the rule (TR) preserves t-validity.

**Theorem 20 (Correctness of TR)**
Let \( \varphi \) be a t-formula such that \( \models_t \varphi \), where \( x \) and \( n \) are not free in \( \varphi \). Then, \( \models_t \forall x \text{Rx}\varphi \).

**Proof**
Let \( \varphi \) be a t-formula such that \( \models_t \varphi \), where \( x \) and \( n \) are not free in \( \varphi \). Assume now that \( \not\models_t \forall x \text{Rx}\varphi \). Then, there is a t-structure \( \pi \) and a state \( w \) such that \( \pi, w \models \neg \forall x \text{Rx}\varphi \). Then, there is a \( t \in T \) and a truth assignment function \( v_t \) such that \( \text{Rx}\varphi \not\in v_{\pi[t/s]}(w) \). Then, there is an \( h \in H \) such that \( h(|w|) = w \) and \( \varphi \not\in v_{\pi[t/s]}(h(t)) \). Hence, there is a t-structure \( \pi \) and a state \( h(t) \in W \) such that \( \pi, h(t) \not\models \varphi \). This is a contradiction to the assumptions that \( \models_t \varphi \). Hence, \( \models_t \forall x \text{Rx}\varphi \).

QED

Having proven that all t-axioms are t-valid and that TR preserves t-validity, it follows that the axiom system is sound over the semantics of t-CML.

**Theorem 21 (Soundness of t-CML)**
If \( \varphi \) then \( \models_t \varphi \) for all t-formulae \( \varphi \).

**Proof**
Let \( \varphi \) be a t-formula such that \( \models_t \varphi \). Then, every formula in the proof sequence for \( \varphi \) is either a t-axiom, and, hence, t-valid by Theorem 19, or the result of applying TR to a t-valid formula, and, hence, t-valid by Theorem 20. Hence, every formula in the proof sequence for \( \varphi \) is t-valid and, thus, \( \varphi \), being the last element of the sequence, is t-valid.

QED

### 4.3.4 Completeness of t-CML

To show that the axiom system for t-CML is complete over the semantic of t-CML, we prove that every t-consistent set of t-sentences is t-satisfiable. The proof involves two steps. Starting with an arbitrary t-consistent set of t-formulae \( \Phi \), we first show, by means of a Henkin construction, that there is a superset \( \Psi \) of \( \Phi \) which is maximal consistent and has the witness property. Then, we define the canonical model \( M^\Psi \) for \( \Psi \) and show
that $M^\Psi$ satisfies all formulae contained in $\Psi$. Once we have proven that every t-consistent set of t-sentences has a model, it is only a small step to show that $\vdash_t \varphi$ implies $\models_t \varphi$.

We start with some standard definitions and a proof that the set of all t-formulae is enumerable. In the first definition we define t-sentences as t-formulae with no free occurrences of variables.

**Definition 62 (t-Sentences)**
A t-sentence $\varphi$ is a t-formula in which all variables occurring in $\varphi$ are bound. The set of all t-sentences is abbreviated by $t$-FOR$_0$.

Next, we define the notion of maximal t-consistency.

**Definition 63 (Maximal t-Consistency)**
A set $\Psi$ of t-sentences is maximal t-consistent iff:
1. $\Psi$ is t-consistent;
2. Either $\varphi \in \Psi$ or $\neg \varphi \in \Psi$ for every t-sentence $\varphi$.

Finally, we define what it means for a set of t-sentences to have the witness property.

**Definition 64 (Witness Property)**
A set $\Psi$ of t-sentences has the witness property iff:
If $\exists x \psi \in \Psi$, there is a t-constant $t^*$ such that $\psi[t^*/x] \in \Psi$ for every t-sentence of the form $\exists x \psi$.

The last requirement before we can define the Henkin construction is the enumerability of $t$-FOR$_\varphi$.

**Theorem 22 (Enumerability of t-FOR$_\varphi$)**
The $t$-FOR$_\varphi$ is enumerable.
Proof

t-FOR₀ is a subset of the set of all finite sequences consisting of elements of t-SYM. t-SYM is countable and, hence, enumerable as it is the union of the countable sets p-CON, t-CON, t-VAR, {n} and {¬, ∧, ∨, (, )}, =, < R}. Hence, t-FOR₀ is enumerable as it is a subset of the set of all finite sequences of t-SYM.

QED

Now we turn to the construction of a maximal t-consistent superset Ψ of a t-consistent set Φ of t-sentences that has the witness property. Starting with a t-consistent set Φ of t-sentences we follow the usual procedure of adding new t-sentences to Φ and witnessing all t-sentences of the form $\exists x \phi$ contained in the construction. This method was introduced by Leon Henkin in his 1947 doctoral thesis to prove the completeness of first-order predicate logic.

Definition 65 (Henkin Construction)

Let $\mathcal{L}_t$ be the language of t-CML and let t-CON* = \{t₁*, t₂*, ...\} be a countably infinite set of new constant symbols. Define $\mathcal{L}_t^* = \mathcal{L}_t \cup$ t-CON*. Let t-FOR₀* be the set of t*-sentences, i.e. sentences of $\mathcal{L}_t^*$.

Let now Φ be a t-consistent set of t-sentences and $\varphi_1, \varphi_2, ...$ be an enumeration of t-FOR₀*.

For every $n \in \mathbb{N}$ we define inductively a superset $\Phi_n$ of Φ as follows:

\[
\Phi_0 = \Phi; \\
\Phi_{n+1} = \begin{cases} 
\Phi_n \cup \{\varphi_{n+1}\} & \text{if } \Phi_n \cup \{\varphi_{n+1}\} \text{ is t-consistent and } \varphi_{n+1} \text{ is not of the form } \exists x \phi; \\
\Phi_n \cup \{\varphi_{n+1}\} \cup \{\varphi[t_n*/x]\} & \text{if } \Phi_n \cup \{\varphi_{n+1}\} \text{ is t-consistent and } \varphi_{n+1} \text{ is of the form } \exists x \phi, \text{ where } t_n* \in \text{t-CON}* \text{ not occurring in } \Phi_n \cup \{\varphi_{n+1}\}; \\
\Phi_n \cup \{\neg \varphi_{n+1}\} & \text{if } \Phi_n \cup \{\varphi_{n+1}\} \text{ is t-inconsistent.}
\end{cases}
\]

Define now $\Psi = \bigcup_{n \in \mathbb{N}} \Phi_n$.

By definition, each $\Phi_n$ is a superset of Φ and, hence, Ψ, being the union of all $\Phi_n$, is a superset of Φ. We still need to show that Ψ is maximal t-consistent and has the witness property. To do this, we first prove that every $\Phi_n$ is t-consistent. Then, we show that Ψ is t-consistent, maximal, and has the witness property.
Chapter 4  Axiomatising Conflicts

Theorem 23  (t-Consistency of $\Phi_n$)
$\Phi_n$ is t-consistent for every $n \in \mathbb{N}$.

Proof
We prove by induction on $n$.

(n = 0) $\Phi_0 = \Phi$ is t-consistent, as the original set $\Phi$ was assumed to be t-consistent.

(Induction Hypothesis) We assume that $\Phi_k$ is t-consistent.

(n = $k+1$) By definition of $\Phi_{k+1}$ we have to consider three cases:

(i) If $\Phi_{k+1} = \Phi_k \cup \{\varphi_{k+1}\}$, where $\Phi_k \cup \{\varphi_{k+1}\}$ is t-consistent and $\varphi_{k+1}$ is not of the form $\exists x \psi$, then $\Phi_{k+1}$ is t-consistent by definition.

(ii) If $\Phi_{k+1} = \Phi_k \cup \{\varphi_{k+1}\} \cup \{\psi[t^*/x]\}$, where $\Phi_k \cup \{\varphi_{k+1}\}$ is t-consistent and $\varphi_{k+1}$ is of the form $\exists x \psi$, then $\Phi_{k+1}$ is t-consistent. Assuming $\Phi_{k+1}$ to be t-inconsistent implies that $\Phi_k \cup \{\varphi_{k+1}\}$ is t-inconsistent as $\Phi_k \cup \{\exists x \psi\} \vdash \psi[t^*/x]$ by FOL and the fact that $t^* \in t$-CON not occurring in $\Phi_n \cup \{\varphi_{k+1}\}$. This is a contradiction to the assumption.

(iii) If $\Phi_{k+1} = \Phi_k \cup \{-\varphi_{k+1}\}$, where $\Phi_k \cup \{\varphi_{k+1}\}$ is t-inconsistent, then $\Phi_{k+1}$ is t-consistent. Assuming $\Phi_{k+1}$ to be t-inconsistent implies that $\Phi_k$ is t-inconsistent as $\Phi_k \vdash -\varphi_{k+1}$ by FOL and the fact that $\Phi_k \cup \{\varphi_{k+1}\}$ is t-inconsistent. This is a contradiction to the assumption.

QED

Now, we show that $\Psi$ is t-consistent.

Theorem 24  (t-Consistency of $\Psi$)
$\Psi$ is t-consistent.

Proof
We prove by reductio. Assume $\Psi$ to be t-inconsistent. Then, $\Psi \vdash \bot$.

Then, there is a finite subset $\Psi_0 \subseteq \Psi$ such that $\Psi_0 \vdash \bot$ as we have a finitary proof theory. Then, there is a $\Phi_n \subseteq \Psi$ such that $\Psi_0 \subseteq \Phi_n$. Then, $\Phi_n \vdash \bot$ and, hence, $\Phi_n$ is t-inconsistent. This is a contradiction to Theorem 23.

QED

Next, we prove that $\Psi$ is maximal t-consistent.

Theorem 25  (Maximal t-Consistency of $\Psi$)
$\Psi$ is maximal t-consistent.

Proof
$\Psi$ is t-consistent by Theorem 24. We must show that $\Psi$ is maximal. Let $\varphi \in t$-SEN* such that $\varphi \not\in \Psi$. As a $t^*$-sentence, $\varphi$ occurs somewhere in the enumeration $\varphi_1, \varphi_2, \ldots$ of all t-FOR*$_n$. Let $n$ be the position at which $\varphi$ occurs in
\( \varphi_1, \varphi_2, \ldots \). Then, \( \Phi_{n-1} \cup \{ \varphi \} \) must be inconsistent as otherwise \( \varphi \in \Phi_n \subseteq \Psi \). Hence, \( \neg \varphi \in \Phi_n \subseteq \Psi \) by construction of \( \Phi_n \).

QED

Finally, we show that \( \Psi \) has the witness property.

**Theorem 26  (Witness Property of \( \Psi \))**

\( \Psi \) has the witness property.

**Proof**

Let \( \varphi \in \text{t-FOR}_0^* \) be of the form \( \exists \psi \varphi \) with \( \varphi \in \Psi \). As a \( \text{t}^* \)-sentence, \( \varphi \) occurs somewhere in the enumeration \( \varphi_1, \varphi_2, \ldots \) of all \( \text{t-FOR}_0^* \). Let \( n \) be the position at which \( \varphi \) occurs in \( \varphi_1, \varphi_2, \ldots \). Then, \( \Phi_n = \Phi_{n-1} \cup \{ \exists \psi \varphi \} \cup \{ \psi[\text{t}*/x] \} \) by construction of \( \Phi_n \). Hence, there is a constant \( \text{t}^* \) such that \( \psi[\text{t}*/x] \in \Phi_n \subseteq \Psi \).

QED

Altogether we get the desired result.

**Theorem 27  (Existence of \( \Psi \))**

If \( \Phi \) is a \( \text{t} \)-consistent set of \( \text{t} \)-sentences, then there is a maximal \( \text{t} \)-consistent set \( \Psi \) that has the witness property such that \( \Phi \subseteq \Psi \).

**Proof**

Starting with a \( \text{t} \)-consistent set \( \Phi \) of \( \text{t} \)-sentences we construct \( \Psi \) as described in Definition 65. Then, \( \Psi \) is maximal \( \text{t} \)-consistent by Theorem 25 and has the witness property by Theorem 26.

QED

Having shown that every \( \text{t} \)-consistent set \( \Phi \subseteq \text{t-FOR}_0 \) can be expanded to a maximal \( \text{t} \)-consistent set \( \Psi \) that has the witness property, we turn now to the construction of the canonical model \( M^\psi \) for \( \Psi \). \( M^\psi \) is a \( \text{t} \)-structure of the form \( <T^\psi, <^\psi, W^\psi, H^\psi, v^\psi, \mu^\psi> \) which has the property that there is a \( w \in W^\psi \) such that \( \varphi \in v^\psi_\mu(w) \) for all \( \varphi \in \Psi \), i.e. \( M^\psi \) satisfies all \( \varphi \in \Psi \).
Before we define $M^w$, some prerequisite concepts are needed. For this purpose, we assume from now on that $w$ is a maximal $t$-consistent set of $t$-sentences that has the witness property.

First, we define an equivalence relation $\sim^w$ on $t$-CON on the basis of $w$.

**Definition 66 ($\sim^w$)**

$t_i \sim^w t_j$ iff $w \vdash t_i = t_j$.

The following theorem expresses that $\sim^w$ is an equivalence relation.

**Theorem 28 (Equivalence Relation)**

$\sim^w$ is an equivalence relation.

**Proof**

We have to prove that $\sim^w$ is reflexive, symmetric and transitive.

(Reflexive) Let $t \in t$-CON. Then, $w \vdash t = t$ by FOL$^w$ and the maximal $t$-consistency of $w$. Hence, $t \sim^w t$ by definition of $t \sim^w t$.

(Symmetric) Let $t_i, t_j \in t$-CON. Then, $w \vdash t_i \supset t_j = t_i$ by FOL$^w$ and the maximal $t$-consistency of $w$. Assume now $t_i \sim^w t_j$. Then, $w \vdash t_i = t_i$ by definition of $\sim^w$ and, hence, $w \vdash t_j = t_i$ by FOL$^w$. Hence, $t_j \sim^w t_i$ by definition of $\sim^w$.

(Transitive) Let $t_i, t_j, t_k \in t$-CON. Then, $w \vdash t_i = t_i \wedge t_j = t_k \supset t_i \sim^w t_k$. Then, $w \vdash t_i = t_i$ and $w \vdash t_i = t_k$ by definition of $\sim^w$. Then, $w \vdash t_i = t_k$ by FOL$^w$. Hence, $t_i \sim^w t_k$ by definition of $\sim^w$.

QED

Now we define $T^w$ as the set of all equivalence classes generated by $\sim^w$ on $t$-CON.

**Definition 67 ($T^w$)**

$T^w = t$-CON/$\sim^w$.

As a convention, we refer to elements of $T^w$ by letters $t, t_0, t_1$ etc. such that $t = [t], t_0 = [t_0], t_1 = [t_1]$ etc. Furthermore, if $w$ is a maximal $t$-consistent set of $t$-sentences that has the witness property, we use the $t$-constant $t_w$ to refer to an arbitrary constant such that $w \vdash t_w = n$. Formally, we can define $t_w$ and $t_w$ as follows.
Definition 68 \((t_w)\)
\[
t_w = \{ t \in t\text{-CON} \mid w \vdash t = n \};
\]
\(t_w\) is an arbitrary element of \(t_w\).

Note that \(t_w\) is a constant symbol, whereas \(t_w\) is an equivalence class of constant symbols.

Note also that \(t_w\) is well defined as for all \(t_i, t_j \in t\text{-CON}\) such that \(w \vdash t_i = n\) and \(w \vdash t_j = n\) we have \(t_i \sim t_j\). It remains to show that there is always a \(t_w\) such that \(w \vdash t_w = n\).

Theorem 29  \((Existence of t_w)\)
Let \(w\) be a maximal \(t\)-consistent set of \(t\)-sentences that has the witness property. Then, there is a \(t\)-constant \(t_w\) such that \(w \vdash n = t_w\).

Proof
\[
\begin{align*}
\text{w \vdash } & \exists x(x = n) \text{ as } \exists x(x = n) \text{ is the } t\text{-axiom (T8) and w is maximal } t\text{-consistent. Then, w \vdash } t_w = n \text{ for some constant } t_w \in t\text{-CON as w has the witness property. Hence, there is a } t\text{-constant } t_w \text{ such that w } \vdash t_w = n. \\
\end{align*}
\]
QED

In the next two definitions, we define two functions called \(Past\) and \(Future\). They assign to every maximal \(t\)-consistent set \(w\) of \(t\)-sentences that has the witness property, sets \(Past(w)\) and \(Future(w)\) of maximal \(t\)-consistent sets of \(t\)-sentences.

First, we define \(Past(w)\), the past of \(w\). The idea behind this definition is that for every \(t \in T^w\) such that \(w \vdash t \leq n\), \(w\) uniquely determines a set of \(t\)-sentences. This set can be interpreted as a complete description of the state in which the conflict was at the time point \(t\) according to \(w\). The \(Past\) function selects all these sets.

Definition 69 \((Past(w))\)
\[
Past(w) = \bigcup_{t \in T^w} p_w(t) \text{ where:}
\]
\[
p_w(t) = \begin{cases} 
\{ \{ \varphi \mid w \vdash R\varphi \} \} & \text{if } w \vdash t \leq n; \\
\emptyset & \text{if } w \vdash n < t.
\end{cases}
\]
Note that $Past(w) \subseteq \emptyset(t\text{-FOR}_0)$, i.e. $Past(w)$ is a set of subsets of $t\text{-FOR}_0$. For every $t \in T^w$ such that $w \vdash t \leq n$, $Past(w)$ contains exactly one set of $t$-sentences. As we will show in Theorem 30, this set is maximal $t$-consistent and has the witness property. For all $t \in T^w$ such that $w \vdash n < t$, $p_w(t)$ is the empty set and, hence, nothing is added to $Past(w)$.

Similar to the past of $w$, we can define its future $Future(w)$. Here the situation is more complicated. First, if $t$ lies in the future of $w$, i.e. $w \vdash n < t$, then $w$ determines, in general, more than one set of $t$-sentences because of the forwards-branching property of $t$-structures. $f_w(t)$, the future equivalent to $p_w(t)$, is, therefore, not a singleton, as in the case of $p_w(t)$, but a set of possibly infinite cardinality. Second, $f_w(t)$ is generally not uniquely determined by $w$. In order to make the definition unique, we define $f_w(t)$ as the set of all sets of $t$-sentences that are compatible with the description of $t$ according to $w$. $f_w(t)$ can be interpreted as an exhaustive description of the set of all states at the time point $t$, into which the conflict can possibly develop, according to $w$.

**Definition 70 (Future($w$))**

$$Future(w) = \bigcup_{t \in T^w} f_w(t)$$

$$f_w(t) = \begin{cases} \{w_0 \subseteq t\text{-FOR}_0 \mid w_0 \text{ is maximal } t\text{-consistent, } w_0 \text{ has the witness property, if } w \vdash Rt_\varphi \text{ and } w \vdash n < t, \text{ then } \varphi \in w_0\} ; \\
\emptyset \text{ if } w \vdash t \leq n. \end{cases}$$

Note that $Future(w) \subseteq \emptyset(t\text{-FOR}_0)$, i.e. $Future(w)$ is a set of subsets of $t\text{-FOR}_0$. For every $t \in T^w$ such that $w \vdash n < t$, $f_w(t)$ is uniquely defined as the set of all maximal $t$-consistent sets that have the witness property and contain all $t$-sentences $\varphi$ such that $Rt_\varphi \in w$. For all $t \in T^w$ such that $w \vdash n < t$, $f_w(t)$ is the empty set and, hence, nothing is added to $Future(w)$.
In the following, we prove three theorems expressing properties of $\text{Past}(w)$ and $\text{Future}(w)$.

The theorems will be used in the completeness proof for t-CML. The first theorem shows that elements of $\text{Past}(w)$ and $\text{Future}(w)$ are maximal t-consistent sets of t-sentences that have the witness property.

**Theorem 30 (Elements of $\text{Past}(w)$ and $\text{Future}(w)$)**  
If $w_0 \in \text{Past}(w) \cup \text{Future}(w)$, then $w_0$ is maximal t-consistent and has the witness property.

**Proof**

If $w_0 \in \text{Future}(w)$, then $w_0$ is maximal t-consistent and has the witness property by definition of $\text{Future}(w)$.

Let $w_0 \in \text{Past}(w)$.

(t-Consistency) We prove by reductio. Assume $w_0$ is t-inconsistent. Then, $w_0 \vdash \varphi \land \neg \varphi$ for some $\varphi \in \text{t-FOR}_0$. Then, $w \vdash \text{Rt}_{w_0}(\varphi \land \neg \varphi)$ and $w \vdash (t_{w_0} \leq n)$ by definition of $\text{Past}(w)$. Hence, $w \vdash \text{Rt}_{w_0}(\varphi \land \neg \varphi)$ by (FOL) and the t-axioms (T1) and (T3). Hence, $w$ is t-inconsistent. This is a contradiction to the assumption.

(Maximality) Assume $\varphi \notin w_0$. Then, $w \not\vdash \text{Rt}_{w_0}(\varphi)$ and $w \vdash (t_{w_0} \leq n)$ by definition of $\text{Past}(w)$. Hence, $w \vdash \neg \text{Rt}_{w_0}(\varphi)$ by the maximality of $w$. Then, $w \vdash \text{Rt}_{w_0}(\neg \varphi)$ by (FOL) and the t-axiom (T1). Hence, $\neg \varphi \in w_0$ by definition of $\text{Past}(w)$.

(Witness Property) Assume $\exists x \varphi \in w_0$. Then, $w \vdash \text{Rt}_{w_0}(\exists x \varphi)$ and $w \vdash (t_{w_0} \leq n)$ by definition of $\text{Past}(w)$. Hence, $w \vdash \exists x \text{Rt}_{w_0}(\varphi)$ by (FOL), the t-axioms (T1) and (T4), and the equivalence between $\neg \forall \neg$ and $\exists$. Then, $w \vdash (\text{Rt}_{w_0}(\varphi))[t^*/x]$ for some t-constant $t^*$ as $w$ has the witness property. Then, $w \vdash \text{Rt}_{w_0}(\varphi)[t^*/x]$ by definition of substitution. Then, $w \vdash \varphi[t^*/x]$ by definition of $\text{Past}(w)$. Hence, there is a constant $t^*$ such that $\varphi[x/t^*] \in w_0$.

QED

The second theorem links the two concepts $\text{Past}(w)$ and $\text{Future}(w)$ via their respective membership relation.

**Theorem 31 (Relationship between $\text{Past}(w)$ and $\text{Future}(w)$)**

$w_0 \in \text{Future}(w_1)$ iff $w_1 \in \text{Past}(w_0)$.

**Proof**

First we prove that $w_0 \in \text{Future}(w_1)$ implies $w_1 \in \text{Past}(w_0)$ by reductio. Assume $w_0 \in \text{Future}(w_1)$ and $w_1 \notin \text{Past}(w_0)$. Then, $w_1 \vdash \text{Rt}_{w_0}(\varphi)$ implies $\varphi \in w_0$ for all $\varphi \in \text{t-FOR}_0$, by the definition of $\text{Future}(w)$, and there is a $\psi \in w_1$ such that $w_0$
Axiomatising Conflicts

\[ \not \vdash R_{w_0} \psi \]

by definition of Past\((w)\). Then, \( w_1 \vdash n = t_{w_1} \) by Theorem 29. We also have \( w_1 \vdash t_{w_1} < t_{w_0} \) by definition of Future\((w)\) and, hence, \( w_1 \vdash n < t_{w_0} \) by (FOL\(=\)). Then, \( w_1 \vdash R_{w_0} R_{w_1} \psi \) by (FOL\(=\)) and the t-axiom (T5). Hence, \( w_0 \vdash R_{w_0} \psi \). This is a contradiction to \( w_0 \not \vdash R_{w_0} \psi \).

Now we show by reductio that \( w_1 \in \text{Past}(w_0) \) implies \( w_0 \in \text{Future}(w_1) \). Assume \( w_1 \in \text{Past}(w_0) \) and \( w_0 \not \in \text{Future}(w_1) \). Then, \( \varphi \in w_1 \) implies \( w_0 \vdash R_{w_1} \varphi \) for all \( \varphi \in \text{t-FOR}_0 \) by definition of Past\((w)\) and there is a \( \psi \not \in w_0 \) such that \( w_1 \vdash R_{w_0} \psi \) by definition of Future\((w)\). Then, \( w_0 \vdash R_{w_0} \psi \). We also have \( w_0 \vdash t_{w_1} < t_{w_0} \) by definition of Past\((w)\) and, hence, \( w_0 \vdash n < t_{w_0} \) by (FOL\(=\)) and because \( w_0 \vdash n = t_{w_0} \) by Theorem 29. Then, \( w_0 \vdash R_{w_0} \psi \) by (FOL\(=\)) and the t-axiom (T6). Then, \( w_0 \vdash \neg \psi \) by (FOL\(=\)) and, hence, \( w_0 \vdash \psi \) by the t-axiom (T7). This is a contradiction to \( \psi \not \in w_0 \).

QED

In the third theorem we prove that the past of \( w \) contains \( w \) itself.

**Theorem 32 (Reflexivity of Past\((w)\))**

\( w \in \text{Past}(w) \).

**Proof**

Assume \( \varphi \in w \). Then, \( w \vdash R_{w} \varphi \) and \( w \vdash (t_w \leq n) \) by (FOL\(=\)) and the t-axiom (T7). Hence, \( p_w(t_w) = \{ \{ \varphi \mid w \vdash R_w \varphi \} \} = \{w\} \). Thus, \( p_w(t_w) = \{w\} \). Hence, \( w \in \text{Past}(w) \).

QED

Now we can define the canonical model \( M^w \) for a maximal t-consistent set \( w \) of t-sentences that has the witness property.

**Definition 71 (Canonical Model)**

Let \( w \) be a maximal t-consistent set that has the witness property. The canonical model \( M^w = \langle T^w, <^w, W^w, H^w, v^w, \mu^w \rangle \) for \( w \) is defined by:

1. \( T^w = t\text{-CON}/\sim^w \) such that \( t_i \sim^w t_j \iff w \vdash t_i = t_j \);
2. \( <^w \subseteq T^w \times T^w \) such that \( t_i <^w t_j \iff w \vdash t_i < t_j \);
3. \( W^w = \bigcup_{w_0 \in \text{Past}(w)} \text{Future}(w_0) \);
4. \( H^w \) is the smallest set of functions \( h: T^w \to W^w \) such that:
   (i) \( \forall h: t; \ t = n \in h(t) \);
   (ii) \( \forall w \exists h, t: h(t) = w \);
   (iii) \( \forall w, h, t: w \in \text{Past}(h(t)) \iff w = h(t_w) \) and \( t_w <^w t \);
5. \( v^w: W^w \to \wp(\text{p-CON}) \) such that \( \forall w: v^w(w) = w \cap \text{p-CON} \).
(6) $\mu^w: \mathbf{t-TER} \rightarrow T^w$ such that $\mu^w(t) = [t] = t$ if $a \in t\text{-}CON$ and $\mu^w(x)$ is an arbitrary element of $T^w$ if $x \in t\text{-}VAR$.

The definition of $M^w$ is motivated by two goals. On the one hand, we have to make sure that $M^w$ turns out to be a $t$-structure. On the other hand we want $M^w$ to satisfy every element of $w$.

$T^w$ is the quotient set $t\text{-}CON$ modulo $\sim^w$, where $\sim^w$ is the equivalence relation defined in Definition 66. $\prec^w$ is a binary relation on $T^w$ such that $\prec^w$ holds between two elements $t_i$, $t_j$ of $T^w$ if and only if the $t$-sentence $t_i < t_j$ is derivable from $w$, where $t_i$ represents the equivalence class $t_i$, and $t_j$ represents the equivalence class $t_j$. The elements of $W^w$ are maximal $t$-consistent sets of $t$-sentences that have the witness property. $W^w$ is constructed as the union of all futures $\text{Future}(w_0)$ where $w_0$ is in the past of the original set $w$, $\text{Past}(w)$. This definition works, as the elements of $\text{Past}(w)$ are maximal $t$-consistent sets of $t$-sentences as shown in Theorem 30. The set of histories, $H^w$, resembles the definition of $H$ for CML-structures. Elements of $H^w$ are functions from $T^w$ into $W^w$ satisfying three conditions of which the last links them with the past function $\text{Past}(w)$ and the order relation $\prec^w$. The truth assignment function $v^w$ assigns to every element $w$ of $W^w$ the set of propositional constants that are contained in $w$. The term assignment function $\mu^w$ interprets $t$-constants by their corresponding equivalence classes and $t$-variables by arbitrary elements of $T^w$.

The following theorem expresses that if $w$ is a maximal $t$-consistent set of $t$-sentences that has the witness property, then $M^w$ is a $t$-structure. To prove this, we show in a number of lemmas that the components of $M^w$ have the required properties. It is always assumed that $w$ is a maximal $t$-consistent set of $t$-sentences that has the witness property and that $M^w = \langle T^w, \prec^w, W^w, H^w, v^w, \mu^w \rangle$ is the canonical model of $w$. 

191
Lemma 1
T* ≠ Ø.

Proof
Consider the following derivation.
1. \( w \vdash \exists x (x = n) \) (T8)
2. \( w \vdash t = n \) for some \( t \)
   Hence, \( t \in T* \neq \emptyset \) and, therefore, \([t] = t \in T* \neq \emptyset\).

QED

Lemma 2
\((T^*, \prec^*)\) is a discrete, strict linear order with no endpoints.

Proof
\( \prec^* \subseteq T^* \times T^* \) by definition. We have to show that \( \prec^*\) is irreflexive, transitive and trichotomous and every \( t \in T^* \) has a unique predecessor and a unique successor.

(Irreflexivity) Let \( t \in T^* \). Consider now the following derivation:
1. \( w \vdash \forall x \neg (x < x) \) (O1)
2. \( w \vdash \neg (t < t) \) (FOL\(^\preceq\)), 1
Hence, \( w \not\vdash t < t \) and, therefore, \( t \nprec^* t \).

(Transitivity) Let \( t_0, t_1, t_2 \in T^* \) such that \( t_0 \prec^* t_1 \) and \( t_1 \prec^* t_2 \). Consider the following derivation:
1. \( w \vdash t_0 < t_1 \) by definition of \( \prec^* \)
2. \( w \vdash t_1 < t_2 \) by definition of \( \prec^* \)
3. \( w \vdash \forall x, y, z \ (x < y \land y < z \Rightarrow x < z) \) (O2)
4. \( w \vdash t_0 < t_1 \land t_1 < t_2 \Rightarrow t_0 < t_2 \) (FOL\(^\preceq\)), 3
5. \( w \vdash t_0 < t_2 \) (FOL\(^\preceq\)), 1, 2, 4
Hence, \( t_0 \prec^* t_2 \).

(Trichotomy) Let \( t_0, t_1 \in T^* \). Then:
1. \( w \vdash \forall x, y, z \ (x < y \lor y = x \lor y < x) \) (O3)
2. \( w \vdash t_0 < t_1 \lor t_0 = t_1 \lor t_1 < t_0 \) (FOL\(^\preceq\)), 1
Hence, \( w \vdash t_0 < t_1 \) or \( w \vdash t_0 = t_1 \) or \( w \vdash t_1 < t_0 \) and, therefore, \( t_0 \nprec^* t_1 \) or \( t_0 = t_1 \) or \( t_1 \nprec^* t_0 \).

(Predecessor/Successor) Let \( t_0 \in T^* \). We have:
1. \( w \vdash \exists x, y \forall x_1 \forall x_2 \ (x_0 < x_1 \land \neg (x_0 < x_2 \land x_2 < x_1)) \) (O4)
2. \( w \vdash \forall x_2 (t_0 < t_1 \land \neg (t_0 < x_2 \land x_2 < t_1)) \) for some \( t_1 \) (FOL\(^\preceq\)), 1, witness property
3. \( w \vdash t_0 < t_1 \) (FOL\(^\preceq\)), 2
4. \( w \vdash \forall x_2 \ (t_0 < x_2 \land x_2 < t_1) \) (FOL\(^\preceq\)), 3
Hence, there is a \( t_1 \in T^* \) such that \( t_0 \nprec^* t_1 \) and for all \( t_2 \in T^* \) either \( t_0 \nprec^* t_2 \) or \( t_2 \nprec^* t_1 \). Thus, \( t_1 \) is the immediate successor of \( t_0 \). Furthermore, we have:
1. \( w \vdash \exists x, y \forall x_1 \forall x_2 (x_1 < x_0 \land \neg (x_1 < x_2 \land x_2 < x_0)) \) (O5)
2. \( w \vdash \forall x_2 (t_1 < t_0 \land \neg (t_1 < x_2 \land x_2 < t_0)) \) for some \( t_1 \) (FOL\(^\preceq\)), 1, witness property
3. \( w \vdash t_i < t_0 \) \hspace{1cm} (FOL\(^-\)), 2
4. \( w \vdash \forall x_1 \sim x_2 (t_1 < x_2 \land x_2 < t_0) \) \hspace{1cm} (FOL\(^-\)), 3

Hence, there is a \( t_i \in T^w \) such that \( t_i <^w t_0 \) and for all \( t_2 \in T^w \) either \( t_i <^w t_2 \) or \( t_2 <^w t_0 \). Thus, \( t_i \) is the immediate predecessor of \( t_0 \).

QED

**Lemma 3**

\( W^w \neq \emptyset \).

**Proof**

\( w \in \text{Past}(w) \) by Theorem 32. Then, \( w \in \text{Future}(w) \) by Theorem 31. Hence, 

\[
\bigcup_{w \in \text{Past}(w)} \text{Future}(w) = W^w \neq \emptyset .
\]

QED

**Lemma 4**

\( H^w \) is a set of functions from \( T^w \) into \( W^w \) such that

1. \( \forall h_0 , h_1, t_0, t_1 \ (h_0(t_0) = h_1(t_1) \supset t_0 = t_1) \)
2. \( \forall w \exists h, t \ (h(t) = w) \)
3. \( \forall h_0, h_1, \exists t_0, \forall t_1 (t_1 < t_0 \supset h_0(t_1) = h_1(t_1)) \)

**Proof**

\( H^w \) is a set of functions from \( T^w \) into \( W^w \) by definition.

(i) Let \( h_0, h_1 \in H^w \) and \( t_0, t_1 \in T^w \) such that \( h_0(t_0) = h_1(t_1) \). Then, there is a \( w_0 \in W^w \) such that \( w_0 \in \text{Past}(w) \) and \( h_0(t_0) \in \text{Future}(w_0) \) by definition of \( W^w \). Then, \( w \in \text{Future}(w_0) \) and \( w_0 \in \text{Past}(h_0(t_0)) \) by Theorem 31. Consider now the following derivation.

1. \( h_0(t_0) \vdash t_0 = n \) \hspace{1cm} by definition of \( H^w \)
2. \( h_0(t_1) \vdash t_1 = n \) \hspace{1cm} by definition of \( H^w \)
3. \( h_0(t_0) \vdash t_1 = n \) \hspace{1cm} as \( h_0(t_0) = h_1(t_1) \)
4. \( h_0(t_0) \vdash t_0 = t_1 \) \hspace{1cm} (FOL\(^-\)), 1, 3
5. \( h_0(t_0) \vdash \forall xRx(t_0 = t_1) \) \hspace{1cm} (\( \forall R \)), 4
6. \( h_0(t_0) \vdash Rt_{w_0}(t_0 = t_1) \) \hspace{1cm} (FOL\(^-\)), 5
7. \( w_0 \vdash t_0 = t_1 \) \hspace{1cm} as \( w_0 \in \text{Past}(h_0(t_0)) \)
8. \( w_0 \vdash \forall xRx(t_0 = t_1) \) \hspace{1cm} (\( \forall R \)), 7
9. \( w_0 \vdash Rt_w(t_0 = t_1) \) \hspace{1cm} (FOL\(^-\)), 8
10. \( w \vdash t_0 = t_1 \) \hspace{1cm} as \( w \in \text{Future}(w_0) \)

Hence, \( t_0 \sim^w t_1 \) and, therefore, \( t_0 = [t_0] = [t_1] = t_1 \).

(ii) \( \forall w \exists h, t \ (h(t) = w) \) by condition ii) of the definition of \( H^w \).

(iii) Let \( h_0, h_1 \in H^w \) with \( h_0 \neq h_1 \).

First, we show that there is a \( t \in T^w \) such that \( h_0(t) = h_1(t) \). As \( h_0 \neq h_1 \), there is a \( t \in T^w \) such that \( h(t) \neq h(t) \). Then, there are \( w_0, w_1 \in W^w \) such that \( w_0, w_1 \in \text{Past}(w), h_0(t) \in \text{Future}(w_0) \) and \( h_1(t) \in \text{Future}(w_1) \) by definition of \( W^w \). Then, \( w_0 \in \text{Past}(h_0(t)) \) and \( w_1 \in \text{Past}(h_1(t)) \) by Theorem 31. Then, \( h_0(t_{w_0}) = w_0 \) and \( h_1(t_{w_1}) = w_1 \) by definition of \( H^w \).
By definition of $H^w$ there is an $h_2 \in H^w$ such that $h_2(t_w) = w$, as $w \in W^w$ by Lemma 3. Then, $w_0, w_1 \in Past(h_2(t_w))$. Then, $h_2(t_w) = w_0$ and $h_2(t_w) = w_1$ by definition of $H^w$. By trichotomy of $<_w$ we get $t_{w_0} <_w t_{w_1}$ or $t_{w_0} = t_{w_1}$ or $t_{w_1} <_w t_{w_0}$.

If $t_{w_0} <_w t_{w_1}$, then $w_0 \in Past(h_2(t_{w_0})) = Past(w_1) = Past(h_1(t_{w_0}))$ by definition of $H^w$. Hence, $h_1(t_{w_0}) = w_0 = h_0(t_{w_0})$.

If $t_{w_0} = t_{w_1}$, then $h_1(t_{w_0}) = w_0 = h_2(t_{w_0}) = h_0(t_{w_0}) = h_2(t_{w_0}) = h_1(t_{w_0})$.

If $t_{w_1} <_w t_{w_0}$, then $w_1 \in Past(h_2(t_{w_0})) = Past(w_0) = Past(h_0(t_{w_0}))$ by definition of $H^w$. Hence, $h_0(t_{w_0}) = w_1 = h_1(t_{w_0})$.

As in all three cases there is a $t \in T^w$ such that $h_0(t) = h_1(t)$, the set $\{ t \mid h_0(t) = h_1(t) \}$ is nonempty. Define $t_{\text{max}}$ as the maximum of $\{ t \mid h_0(t) = h_1(t) \}$ with respect to $<_w$. Then, $h_0(t_{\text{max}}) = h_1(t_{\text{max}})$.

Let $t_0 \in T^w$ such that $t_0 <_w t_{\text{max}}$. Then, $h_0(t_0) \in Past(h_0(t_{\text{max}})) = Past(h_1(t_{\text{max}}))$ by definition of $H^w$ and, therefore, $h_1(t_0) = h_0(h_0(t_0)) = h_0(t_0)$.

Let $t_0 \in T^w$ such that $t_{\text{max}} <_w t_0$. Then, $t_0 \notin \{ t \mid h_0(t) = h_1(t) \}$. Hence, $h_0(t_0) \neq h_1(t_0)$.

QED

**Lemma 5**

$v^w$ is a function from $W^w$ into $\wp(\text{p-CON})$.

**Proof**

$v^w$ is a function from $W^w$ into $\wp(\text{p-CON})$ by definition of $M^w$.

QED

**Lemma 6**

$\mu$ is a function from $t$-TER into $T^w$.

**Proof**

$\mu$ is a function from $t$-TER into $T^w$ by definition of $M^w$.

QED

Now it follows readily from Lemma 1 to Lemma 6 that $M^w$ is a t-structure.

**Theorem 33  (M$^w$ and t-Structures)**

If $w$ is a maximal t-consistent set of $t$-sentences that has the witness property, then $M^w$ is a t-structure.

**Proof**

All components of $M^w$ satisfy the conditions of a t-structure by Lemma 1 to Lemma 6. Hence, $M^w$ is a t-structure.

QED
Having proven that $M^w$ is a t-structure we show now that for a given set $w$ of maximal t-consistent sentences, $M^w$ satisfies all elements of $w$. To do this, we have to show that there is a state $w_0 \in W^w$ such that $\varphi \in v^w_\mu(w_0)$ if and only if $\varphi \in w$ for all $\varphi \in w$. The state $w_0$ which has this property is $w$ itself. The choice of $w$ is possible as $w \subseteq W^w$ by Lemma 3.

**Theorem 34** *(Satisfiability of $w$)*

If $w$ is a maximal t-consistent set of t-sentences\(^\text{166}\) that has the witness property and $M^w = \langle T^w, w^w, H^w, v^w_\mu, \mu^w \rangle$ is the canonical model of $w$, then $\varphi \in v^w_\mu(w)$ iff $\varphi \in w$, for every $\varphi \in t$-FOR\(_w\).

**Proof**

We prove by induction on the formation of $\varphi$.

- **(p)** Assume $p \in w$. Then, $p \in w \cap p$-TCON = $v^w_\mu(w)$ by definition of $v^w_\mu$. Hence, $p \in v^w_\mu(w)$.

  Assume now $p \notin w$. Then, $p \notin w \cap p$-TCON = $v^w_\mu(w)$ by definition of $v^w_\mu$. Hence, $p \notin v^w_\mu(w)$.

  **(t$_0$ = t$_1$)** Assume $t_0 = t_1 \in w$. Then, $w \models t_0 = t_1$ and, hence, $t_0 \equiv w t_1$ by definition of $\equiv w$. Then, $\mu^w(t_0) = t_0 = t_1 = \mu^w(t_1)$ by definition of $\mu^w$. Hence, $t_0 = t_1 \in v^w_\mu(w)$.

  Assume now $t_0 = t_1 \notin w$. Then, $w \models t_0 = t_1$ as $w$ is maximal t-consistent and, hence, $t_0 \equiv_w t_1$ by definition of $\equiv_w$. Then, $\mu^w(t_0) = t_0 \neq t_1 = \mu^w(t_1)$ by definition of $\mu^w$. Hence, $t_0 = t_1 \notin v^w_\mu(w)$.

- **(n = t)** Assume $n = t \in w$. Then, $w \models n = t$ and $w \models t_w = n$ by definition of $t_w$. Then, $w \models t_w = t$ by (FOL\(_n\)). Then, $|w| = t_w = t = \mu^w(t)$ as in the previous case. Hence, $n = t \in v^w_\mu(w)$.

  Assume now $n = t \notin w$. Then, $w \models n = t$ as $w$ is maximal t-consistent and $w \models t_w = n$ by definition of $t_w$. Then, $w \models t_w = t$ by (FOL\(_n\)). Then, $|w| = t_w \neq t = \mu^w(t)$ as in the previous case. Hence, $n = t \notin v^w_\mu(w)$.

---

\(^{166}\) The result also holds for open t-formulae as we can transform every open t-formula $\varphi$ into a t-sentence $\varphi^\#$ such that $\varphi$ is satisfied if $\varphi^\#$ is satisfied by the following construction. For a given maximal t-consistent set $\Phi$ of (possibly open) t-formulae that has the witness property, we first ensure that all free variables occurring in $\Phi$ are distinct from each other by replacing some of the free variables as required. Second, we extend our language by a countably infinite set CON\(_n\) of new constants and define an injective function $f$ from the set of (distinct) free variables occurring in $\Phi$ into the set CON\(_n\). Next, we replace each open formula $\varphi \in \Phi$ by a formula $\varphi^\# = \varphi[x_1/f(x_1), \ldots, x_n/f(x_n)]$, where $x_1, \ldots, x_n$ are the free variables occurring in $\varphi$. This gives us a set $\Phi^\#$ of t-sentences, for which the satisfiability theorem holds, i.e. every formula $\varphi[x_1/f(x_1), \ldots, x_n/f(x_n)]$ is satisfied by the canonical model $M^\#$. Then, an induction on the formation of t-formulae shows that $M^\# \models \varphi[x_1/f(x_1), \ldots, x_n/f(x_n)]$ implies $M^\# \models f(x_1/\mu(x_1), \ldots, x_n/\mu(x_n))$ by $\varphi$. Note that we can choose $\mu(x_1/f(x_1), \ldots, x_n/f(x_n))$ as $\mu(x)$ can be an arbitrary element of $T^\#$ in the canonical model $M^\#$.  

195
(c, < c) Assume $t_0 < t_1 \in W$. Then, $w \vdash t_0 < t_1$, and, hence, $t_0 \triangleleft^w t_1$ by definition of $\triangleleft^w$. Then, $\mu^w(t_0) = t_0 \triangleleft^w t_1 = \mu^w(t_1)$ by definition of $\mu^w$. Hence, $t_0 < t_1 \in v^w_\mu(w)$.

Assume now $t_0 < t_1 \in W$. Then, $w \not\vdash t_0 < t_1$ as $w$ is maximal $t$-consistent and, hence, $t_0 \not\triangleleft^w t_1$ by definition of $\triangleleft^w$. Then, $\mu^w(t_0) = t_0 \not\triangleleft^w t_1 = \mu^w(t_1)$ by definition of $\mu^w$. Hence, $t_0 < t_1 \not\in v^w_\mu(w)$.

(n < t) Assume $n < t \in W$. Then, $w \vdash n < t$ and $w \vdash t_w = n$ by definition of $t_w$. Then, $w \vdash t_w < t$ by ($\text{FOL}^\gamma$). Then, $|w| = t_w \triangleleft^w t = \mu^w(t)$ as in the previous case. Hence, $n < t \not\in v^w_\mu(w)$.

Assume now $n < t \not\in W$. Then, $w \not\vdash n < t$ as $w$ is maximal $t$-consistent and $w \vdash t_w = n$ by definition of $t_w$. Then, $w \not\vdash t_w < t$ by ($\text{FOL}^\gamma$). Then, $|w| = t_w \not\triangleleft^w t = \mu^w(t)$. Hence, $n < t \not\in v^w_\mu(w)$.

(Induction Hypothesis) We assume $\psi, \chi \in v^w_\mu(w)$ iff $\psi, \chi \in W$ for all sub-formulae $\phi$ and $\chi$ of $\varphi$.

($\neg \psi$) Assume $\neg \psi \in W$. Then, $\psi \not\in W$ as $w$ is maximal $t$-consistent. Then, $\psi \not\in v^w_\mu(w)$ by the induction hypothesis and, hence, $\neg \psi \not\in v^w_\mu(w)$.

Assume now $\neg \psi \not\in W$. Then, $\psi \not\in W$ as $w$ is maximal $t$-consistent. Then, $\psi \in v^w_\mu(w)$ by the induction hypothesis and, hence, $\neg \psi \not\in v^w_\mu(w)$.

($\psi \land \chi$) Assume $\psi \land \chi \in W$. Then, $\psi \in W$ and $\chi \in W$ as $w$ is maximal $t$-consistent. Then, $\psi \in v^w_\mu(w)$ and $\chi \in v^w_\mu(w)$ by the induction hypothesis and, hence, $\psi \land \chi \in v^w_\mu(w)$.

Assume now $\psi \land \chi \not\in W$. Then, $\psi \not\in W$ or $\chi \not\in W$ as $w$ is maximal $t$-consistent. Then, $\psi \not\in v^w_\mu(w)$ or $\chi \not\in v^w_\mu(w)$ by the induction hypothesis and, hence, $\psi \land \chi \not\in v^w_\mu(w)$.

($\forall x \psi$) Assume $\forall x W$. Let $t \in W$. Then, $\phi[t/x] \in W$ by ($\text{FOL}^\gamma$) and the maximal $t$-consistency of $w$. Then, $\phi[t/x] \in v^w_\mu(w)$ by the induction hypothesis. Then, $\psi \in v^w_\mu(w)$. As $t$ was chosen arbitrarily, we get $\psi \in v^w_\mu(w)$ for all $t \in W$ and, hence, $\forall x \psi \in v^w_\mu(w)$.

Assume now $\forall x \psi \not\in W$. Then, $\neg \forall x \psi \in W$ as $w$ is maximal $t$-consistent. Then, $\exists x \neg \psi \in W$ by definition of $\exists$. Then, $\neg \psi[t/x] \in W$ for some $t$-constant $t$ as $w$ has the witness property. Then, $\neg \psi[t/x] \not\in W$ as $w$ is maximal $t$-consistent and, hence, $\psi[t/x] \not\in v^w_\mu(w)$ by the induction hypothesis. Then, $\psi \not\in v^w_\mu(w)$. Thus, there is a $t \in W$ such that $\psi \not\in v^w_\mu(w)$ and, hence, $\forall x \psi \not\in v^w_\mu(w)$.

($\text{Rt} \psi$) Assume $\text{Rt} \psi \in W$. Then, $w \vdash t < n \lor t = n \lor n < t$ by the $t$-axiom (O3) and ($\text{FOL}^\gamma$) Then, $t < n \in W$ or $t = n \in W$ or $n < t \in W$.

If $t < n \in W$ or $t = n \in W$, let $h \in H^w$ be such that $h(|w|) = w$. Define $w_i = \{ \psi \mid \text{Rt} \psi \in W \}$. Then, $\psi \in W$, and $\{w_i\} = p_w(t)$ by definition of $p_w$ Hence, $w_i \in \text{Past}(w) = \text{Past}(h(|w|)).$ Then, $w_i = h(t_{w_i}) = h(t)$ by definition of $H^w$, and, hence, $\psi \in h(t) = h(\mu^w(t))$ by definition of $\mu^w$. As $h$ was chosen arbitrarily, we get $\psi \in h(\mu^w(t))$ for all $h \in H^w$ such that $h(|w|) = w$. Hence, $\text{Rt} \psi \in v^w_\mu(w)$.

If $n < t \in W$, let $h \in H^w$ such that $h(|w|) = w$. Then, $\psi \in W$, for all $w_i \in f_w(t)$ by definition of $f_w$. Now we show that $h(t) \in f_w(t)$. We have $w \vdash t_w < t$ by ($\text{FOL}^\gamma$) and by definition of $t_w$ and, hence, $|w| = t_w \triangleleft^w t$ by definition of $\triangleleft^w$.

Then, $w \in \text{Past}(h(t))$ by definition of $H^w$. Then, $h(t) \in \text{Future}(w)$ by Theorem 31. Then, $h(t) \in f_w(t)$ by definition of $\text{Future}$ and, hence, $\psi \in h(t) = h(\mu^w(t))$ by
The Axiomatics of Conflict Modelling Logic

definition of $\mu^w$. As $h$ was chosen arbitrarily we get $\psi \in h(\mu(t))$ for all $h \in H^w$ such that $h(|w|) = w$. Hence, $Rt\psi \in v^w_p(w)$.

Assume now $Rt\psi \not\in w$. Then, $w \vdash t < n \lor t = n \lor n < t$ by the t-axiom (O3) and (FOL). Then, $t < n \in w$ or $t = n \in w$ or $n < t \in w$.

If $t < n \in w$ or $t = n \in w$, then there is an $h \in H^w$ such that $h(|w|) = w$. Hence, $\psi \vdash h(\mu^w(t))$ and, hence, is an $h \in H^w$ such that $h(|w|) = w$ and $\psi \not\in h(\mu^w(t))$. Hence, $Rt\psi \not\in v^w_p(w)$.

If $n < t \in w$, then there is a $w \vdash t < n \lor t = n \lor n < t$ by (FOL) and by definition of $t^w$ as $w$ is maximal t-consistent and, hence, $\psi \not\in w$ by the t-axiom (T7) and the maximal t-consistency of $w$. Hence, $\psi \vdash h(\mu^w(t))$ by definition of $\mu^w$. Then, $\psi \not\in h(\mu^w(t))$ and, hence, there is an $h \in H^w$ such that $h(|w|) = w$ and $\psi \not\in h(\mu^w(t))$. Hence, $Rt\psi \not\in v^w_p(w)$.

Let $w \in f_w(t)$ such that $\psi \not\in w$, by definition of $f_w$. We show that $h(\mu^w(t)) = w$. Define $w_1 = \{\psi \mid Rt\psi \in w\}$. Then, $\psi \in w$, and $w_1 = p_w(t)$ by definition of $p_w$. Then, $w_1 = h(\mu^w(t))$ by definition of $H^w$ and, hence, $\psi \not\in w$ by the t-axiom (T7) and the maximal t-consistency of $w$. Then, $\psi \not\in h(\mu^w(t))$ and, hence, there is an $h \in H^w$ such that $h(|w|) = w$ and $\psi \not\in h(\mu^w(t))$. Hence, $Rt\psi \not\in v^w_p(w)$.

We also have $w \in \text{Future}(w)$ as $w \in f_w(t)$. Then, $w \in \text{Past}(w) = \text{Past}(h(t))$ by Theorem 31 and, hence, $h(\mu^w(t)) = w$ by definition of $H^w$. Hence, there is an $h \in H^w$ such that $h(|w|) = w$ and $\psi \not\in h(t) = h(\mu^w(t))$ by definition of $\mu^w$. Hence, $Rt\psi \not\in v^w_p(w)$.

Assume $Rt\psi \not\in v^w_p(w)$. Then, $\psi \not\in w$ by the t-axiom (T7) and the maximal t-consistency of $w$. Then, $\psi \not\in v^w_p(w)$ by the induction hypothesis. Hence, $Rt\psi \not\in v^w_p(w)$.

QED

With the satisfiability theorem, we have shown that every maximal t-consistent set $w$ that has the witness property is satisfied by its canonical model $M^w$, i.e. $M^w, w \models \varphi$ for all $\varphi \in w$. We now show that every t-consistent set $\Phi$ is satisfiable. This is a simple consequence of our Henkin construction.

**Theorem 35 (Satisfiability of $\Phi$)**

Let $\Phi$ be a t-consistent set of t-sentences. Then, $\Phi$ is satisfiable.

**Proof**

Let $\Phi$ be a t-consistent set of t-sentences. Then, there is a maximal t-consistent superset $\Psi$ of $\Phi$ that has the witness property by Theorem 27. Then, $M^\Phi, \Psi \models \Phi$ by Theorem 34. Then, $M^{\Psi}, \Psi \models \Phi$ as $\Phi \subseteq \Psi$. Hence, $\Phi$ is satisfiable.

QED
Finally we prove the completeness theorem for t-CML, i.e. we show that $\vdash_t \varphi$ implies $\models_t \varphi$

for every t-sentence $\varphi$.

**Theorem 36  (Completeness of t-CML)**

If $\models_t \varphi$, then $\vdash_t \varphi$, for every t-sentence $\varphi$.

**Proof**

We prove by reduction. Let $\varphi$ be a t-sentence. Assume $\models_t \varphi$ and $\not\vdash_t \varphi$.

Then, $\{ \neg \varphi \}$ is t-consistent and, hence, satisfiable by Theorem 35. Hence, there is a t-structure $\mathfrak{x}$ and a state $w$ in $\mathfrak{x}$ such that $\mathfrak{x}, w \models_t \neg \varphi$ and, hence, $\mathfrak{x} \not\models_t \varphi$. This is a contradiction to $\models_t \varphi$.

QED

### 4.4 The Propositional Attitude Fragment

The propositional attitude fragment of CML, p-CML, is significantly different from the other CML-fragments in that it is not a logical system. The reason for this is that we have made no restrictions on the four propositional attitude functions $b$, $g$, $n$, and $e$. As a consequence, there are no theorems involving the propositional attitude operators other than substitution instances of theorems of the other fragments.

We have not restricted the propositional attitude functions because real agents’ propositional attitudes, i.e. their beliefs, goals, norms, and emotions, are generally not restricted. Agents hold arbitrary combinations of propositional attitudes. In particular, agents are often not logically omniscient, their beliefs are inconsistent, they pursue contradictory goals, and they feel uncommitted to the logical consequences of the norms they hold. Putting constraints on $b$, $g$, $n$, or $e$, or formulating axioms for the propositional attitude operators $B$, $G$, $N$ or $E$, would require agents to satisfy conditions of consistency, rationality, morality, or emotional coherence. Such conditions, which are typically imposed on belief operators of epistemic logic, goal operators of preference
logic, or norm operators of deontic logic, have the effect that only artificial agents’ propositional attitudes can be described by these logics.

Although agents are not required to satisfy certain conditions with regard to their propositional attitudes, they are not prohibited from it either. Hence, we can formulate conditions for b, g, and n which, if satisfied by an agent, qualify the agent as a consistent believer, a rational agent, or a morally coherent agent. The conditions allow one to check various levels of consistency of an agent’s propositional attitudes. In the following, we formulate some constraints on the belief function b as an example illustrating how attitude functions can be constrained.

### 4.4.1 Consistent Believers

The following conditions express different aspects of epistemic consistency. Each of them is assumed to hold for every formula φ and every state w.

**Definition 72 (Consistency Conditions)**

1. If $\varphi \in b(a, w)$, then $\neg \varphi \notin b(a, w)$;
2. If $\varphi \in b(a, w)$ and $\varphi \supseteq \psi \in b(a, w)$, then $\neg \psi \notin b(a, w)$;
3. If $\varphi \in b(a, w)$ and $\varphi \supseteq \psi \in b(a, w)$, then $\psi \in b(a, w)$;
4. If $\varphi \in b(a, w)$ and $\vdash_{\text{CML}} \varphi \supseteq \psi$, then $\neg \psi \notin b(a, w)$;
5. If $\varphi \in b(a, w)$ and $\vdash_{\text{CML}} \varphi \supseteq \psi$, then $\psi \in b(a, w)$;
6. If $\vdash_{\text{CML}} \varphi$, then $\varphi \in b(a, w)$;
7. If $\varphi \in b(a, w)$, then $Ba\varphi \in b(a, w)$.

An agent $a$ satisfying the first condition is excluded from believing that a statement is true at the same time as believing that the negation of the statement is true. In other words, the belief set $b(a, w)$ cannot contain a formula $\varphi$ and its negation $\neg \varphi$. The condition is rather weak. It neither prevents agents from believing contradictions, such as $\varphi \land \neg \varphi$, nor commits them to believing in the logical consequences of their beliefs. If the
condition is assumed to hold for all agents and all formulae, the scheme \( \forall x(Bx\varphi \supset \neg Bx\neg \varphi) \) becomes valid in CML.

The second condition can be interpreted as follows. If an agent believes \( \varphi \) as well as that \( \psi \) is a consequence of \( \varphi \), then the agent does not believe that the negation of \( \psi \) is true. The condition is a generalisation of the first consistency condition, although it is independent from it. The first condition follows from the second one only if the agent believes that \( \varphi \supset \psi \) is true. The corresponding scheme for the second condition is \( \forall x(Bx(\varphi \supset \psi) \supset (Bx\varphi \supset \neg Bx\neg \psi)) \), which is valid in CML for all formulae \( \varphi \) and \( \psi \) if the condition is assumed to hold.

We can strengthen the second condition by requiring agents to believe that \( \psi \) is true, when they believe that \( \varphi \) is true and that \( \psi \) is a consequence of \( \varphi \). This is expressed in the third consistency condition which is also independent from the first two conditions. The third and first condition, however, jointly imply the second one. The third condition can be characterised as an internal epistemic closure principle for agents. It requires agents to believe in the truth of statements they believe to be the consequences of other beliefs they hold. It does not require agents to be logically omniscient in the sense that they are committed to believe that all the logical consequences of their beliefs are true. Furthermore, it does not require agents to be consistent with regard to their beliefs. This can be seen by looking at an agent \( a \) who believes \( \varphi \) and \( \neg \varphi \), but does not believe \( \psi \lor \neg \psi \). The agent satisfies condition three, although his belief set is inconsistent and he does not believe \( \psi \lor \neg \psi \), even though this follows logically \( \varphi \). Hence, he is neither a consistent believer nor logically omniscient. The corresponding scheme, which is true if the third condition is assumed to hold, is \( \forall x(Bx(\varphi \supset \psi) \supset (Bx\varphi \supset Bx\psi)) \).
The fourth and fifth condition are similar to the previous two. The difference is that in their antecedents, agents are no longer required to believe that a certain conditional is true. Instead, the conditional in question must be a theorem of CML. In the case of condition five, for example, an agent who believes $\varphi$ is committed to believing that $\psi$ is true if $\varphi$ logically implies $\psi$, i.e. if $\varphi \supset \psi$ is a theorem of CML. The fourth condition can be characterised as an external consistency condition. Agents satisfying the condition are excluded from believing in statements that are logically inconsistent with other beliefs they hold. The fifth condition describes external, or logical, omniscience. Agents satisfying the condition are required to believe that a statement is true if it logically follows from a belief they hold. This condition implies that agents who have at least one belief also believe any theorem of CML is true since CML-theorems follow from arbitrary beliefs. Both conditions are independent from the other conditions.

The sixth consistency condition requires agents to believe that any theorem of CML is true. It is similar to condition five, but stronger in the sense that even agents who do not have a single belief (other than a theorem) are forced to believe at least all theorems of CML. Hence, the condition also expresses a form of logical omniscience which is even stronger than the logical omniscience expressed in condition five. If the condition is assumed to hold in the semantics of CML, the following rule of inference becomes correct: if $\vdash_{\text{CML}} \varphi$, then $\vdash_{\text{CML}} \forall x Bx \varphi$.

The last condition expresses the KK-principle of epistemic logic. This condition, which expresses positive introspection for agents, makes all formulae of the form $\forall x (Bx \varphi \supset BxBx \varphi)$ valid in CML. It makes sure that every agent who believes $\varphi$ also believes that he believes that $\varphi$. 


More consistency conditions can be formulated in addition to the above seven. In particular, conditions dealing with beliefs that involve other connectives, such as $\land$ or $\lor$, as well as conditions dealing with the belief operator itself, are relevant.

The logical relationships among the conditions are as follows. Condition one and three jointly imply condition two. Condition two and six jointly imply condition one and four. Condition three and six jointly imply condition five. Condition three and four jointly imply condition two. Condition four implies condition one.

### 4.5 Axioms for CML

In the following section, we show that the axioms provided for the temporal fragment of CML provide also an axiomatisation of CML as a whole. The reason for this is that the modal operator $\square$ can be defined in terms of temporal symbols and, as explained in section 4.4, the propositional attitude operators $B$, $G$, $N$, and $E$ do not require any axioms as there are no restrictions on the four propositional attitude functions $b$, $g$, $n$, and $e$. From this, it follows that all elements of $L_{CML}$, i.e. all connectives and operators, can be axiomatically characterised by the temporal axioms alone.

Note that, although we can define the modal operators in terms of temporal operators, the completeness result of section 4.2 for the modal fragment provides an independent construction making sure that the modal fragment can be used even if we do not want to rely on the temporal fragment.

Before we state the definitions for $\square$ and $\lozenge$ in terms of temporal symbols, we look at the semantic relationship between $\square$-CML, $t$-CML, and CML itself. Taking into account that $\square$-structures are substructures of $t$-structures which are, in turn, substructures of CML-
structures, we obtain the following theorem expressing the substructure relation \( \subseteq \) with respect to \( \Box \)-structures, \( t \)-structures, and CML-structures.

**Theorem 37  (Substructure Relation)**

If \( x \subseteq y \) expresses that \( x \) is a substructure of \( y \), then the following substructure relations hold: \( \Box \)-structure \( \subseteq \) \( t \)-structure \( \subseteq \) CML-structure.

**Proof**

The relations hold by definition of the respective structures.

QED

Now we turn to the axiomatisation of CML. To do this, we have to express the modal operators in terms of temporal symbols. First we repeat two abbreviations introduced in section 2.4.1. There we defined the next operator \( X \) expressing that a formula \( \varphi \) is true at the next time point and the yesterday operator \( Y \) expressing that \( \varphi \) is true at the immediately preceding time point.

**Definition 73  (Next Operator)**

The next operator, \( X \), is defined as follows:

\[
X\varphi = \forall x_0 (n < x_0 \land \exists x_1 (n < x_1 \land x_1 < x_0)) \Rightarrow Rx_0(\varphi).
\]

**Definition 74  (Yesterday Operator)**

The yesterday operator, \( Y \), is defined as follows:

\[
Y\varphi = \exists x_0 (x_0 < n \land \exists x_1 (x_0 < x_1 \land x_1 < n)) \Rightarrow Rx_0(\varphi).
\]

The antecedents of the two definitions express that \( x_0 \) is an immediate successor in the case of \( X \) and an immediate predecessor in the case of \( Y \). As \( \prec \) is a discrete, strict linear order relation, the successor and the predecessor of an arbitrary element of \( T \) are uniquely defined.

We now define the two modal operators, \( \Box \) and \( \Diamond \), in terms of \( X \) and \( Y \). The definitions show that the modal operators can be expressed entirely in the temporal language of \( t \)-CML. First, we define the necessity operator.
Definition 75 (Temporal Definition of the Necessity Operator)
The historical necessity operator $\Box$ is defined as follows:
\[ \Box \varphi = YX\varphi. \]

Now we define the possibility operator in the usual way.

Definition 76 (Temporal Definition of the Possibility Operator)
The historical possibility operator $\Diamond$ is defined as follows:
\[ \Diamond \varphi = \neg YX\neg \varphi. \]

In order to make sure that the above definitions are semantically well defined, we have to show that the respective formulae have the same truth conditions.

Theorem 38 ($\Box$ and $YX$)
If $\varphi$ is a CML-formula, then $YX\varphi \in v_\mu(w)$ iff $\Box \varphi \in v_\mu(w)$.

Proof
\[ YX\varphi \in v_\mu(w) \text{ iff } \varphi \in v_\mu(h_0(|w|)) \text{ for all } h_0 \in H \text{ such that there is a } h_1 \in H \]
with $h_0(|w|-1) = h_1(|w|-1)$ and $h_1(|w|) = w \text{ iff } \Box \varphi \in v_\mu(w). \]
QED

The theorem shows that $\Box$-CML can entirely be expressed in $t$-CML. An analogous theorem holds for $\Diamond$ and $\neg YX\neg \varphi$.

Theorem 39 ($\Diamond$ and $\neg YX\neg$)
If $\varphi$ is a CML-formula, then $\neg YX\neg \varphi \in v_\mu(w)$ iff $\Diamond \varphi \in v_\mu(w)$.

Proof
\[ \neg YX\neg \varphi \in v_\mu(w) \text{ iff } \neg \Box \varphi \in v_\mu(w) \text{ by Theorem 38 iff } \Diamond \varphi \in v_\mu(w) \text{ by definition of } \Diamond. \]
QED

As a consequence of the two theorems, we can axiomatise CML by using the definition of $\Box$ as an additional axiom to the axioms of $t$-CML.

Definition 77 (CML-Axioms)
(t-CML) Axioms of t-CML;
□φ = YXφ, where Y and X are defined as in Definition 73 and Definition 74
(∀) ⊢ φ ⊃ ψ ⇒ ⊢ φ ⊃ ∀xψ, where x is not free in ψ;
(MP) ⊢ φ ⊃ ψ, φ ⇒ ⊢ ψ;
(∀R) ⊢ φ ⇒ ⊢ ∀xRxp, where x and n are not free in ψ.

CML-consistency, CML-inconsistency, and CML-provability are defined in the usual way.

Definition 78 (CML-Provability)
Let φ be a CML-formula. Then, φ is CML-provable, i.e. ⊢CML φ, iff there is a finite sequence of CML-formulae such that the last formula of the sequence is φ and every other formula in the sequence is either a CML-axiom or the result of applying a rule of inference to formulae occurring earlier in the sequence.

Definition 79 (Syntactic Consequence Relation for CML)
Let Φ be a set of CML-formulae and φ be a t-formula. Then, the syntactic consequence relation holds between Φ and φ, i.e. Φ ⊢CML φ, if and only if there is a finite sequence of CML-formulae such that the last formula of the sequence is φ and every other formula in the sequence is either a CML-axiom, and element of Φ or the result of applying a rule of inference to formulae occurring earlier in the sequence.

Definition 80 (CML-Inconsistency/CML-Consistency)
Let Φ be a set of CML-formulae and ⊥ any CML-formula of the form φ ∧ ¬φ. Then, Φ is CML-inconsistent if and only if Φ ⊢CML ⊥ and Φ is CML-consistent if and only if Φ is not CML-inconsistent.

Soundness and completeness of CML follow directly from the respective theorems for t-CML.

4.6 Summary

The aim of chapter 4 was to axiomatise CML. This has been achieved to the extent that we have provided axiom systems for the modal fragment and the temporal fragment, and have shown why the axiom system for the temporal fragment also provides an axiomatisation for CML as a whole. An axiomatisation of the propositional attitude
fragment has been shown to be redundant as this fragment does not constitute a logical system. For this fragment we have exemplarily stated various conditions, which constrain agents’ beliefs.

The two completeness proofs are the most technical result of the thesis. In particular, the completeness proof for the temporal fragment has required a relatively large amount of abstract definitions and theorems. They were required, though, as only axiom systems make it possible to define the notion of inconsistency in a precise, proof theoretical, way.

Having provided a sound and complete axiomatisation of CML, we can check whether a set of CML-formulae is inconsistent or not. The axiomatisation of CML can be seen as the third step in the process of modelling and resolving a conflict as it brings us a step closer to a general definition of conflict and a criterion for what counts as a solution to a conflict.

In the following chapter we will use the notion of inconsistency to formulate a general conflict definition. Furthermore, we will use the notion in a number of specific definitions stressing different aspects of conflicts and making it possible to classify them.
CHAPTER 5

Defining Conflicts

The Classification Scheme of Conflict Modelling Logic

5.1 Introduction

The aim of chapter 5 is to provide a general conflict definition, and a classification scheme for conflicts, based on the formal framework of CML. For the general definition, we need to clarify the relationship between conflicts and the concept of inconsistency. The classification scheme differentiates between six parameters for the description of conflicts and allows us to identify the specific types of conflicts.

As a result we will be able to answer to the following questions.

- What is the defining property of conflicts?
- How is the concept of inconsistency related to the definition of conflicts?
- How can conflicts be classified with respect to their temporal structure, the type and number of agents occurring in them, their relationship to the outside world, the issues they are about, and their emotional load?
Chapter 5  Defining Conflicts

- How can descriptions of conflicts in terms of CML-formulae be transformed into a unified propositional form?
- How can sub-conflicts of the Second Congo War be identified and classified within the formal framework of CML?

An outline of the chapter is as follows: We start with a background section on conflict definitions. This is followed by a general characterisation of conflicts in section 5.2. In section 5.3, we look at the temporal dimension of conflicts, and contrast their synchronic description with the diachronic description of conflicts. In section 5.4, we define three social levels on which conflicts can occur, and, in section 5.5, we classify them, according to the number of agents occurring in them, as either intra-agent or inter-agent conflicts. Then, we define subjective and objective conflicts in section 5.6 and in section 5.7 we further distinguish them with regard to the type of issues they are about. In particular, we identify factual disputes (5.7.1), conflicts over goals (5.7.2), and conflicts over values (5.7.3). In section 5.8, we deal with the emotional load of conflicts and in section 5.9 we show how the various conflict definitions can be combined with each other. In section 5.10 we explain, how conflicts can be transformed into a unique form called the propositional form of a conflict. Conflicts have to be transformed into propositional form, in order to input them into the resolution algorithms which we will introduce in the last two chapters of the thesis. Finally, we apply the definitions to our example of the Second Congo War.

Background: Conflict Definitions

In the following section we present a number of conflict definitions as proposed by various conflict scholars or defined in dictionaries. We present the definitions without
further comment just as a background for our own conflict definition and classification scheme.

We start by looking at six definitions of conflict from dictionaries. According to the American Heritage Dictionary a conflict is

“1. [a] state of open, often prolonged fighting; a battle or war.

2. A state of disharmony between incompatible or antithetical persons, ideas, or interests; a clash.

3. Psychology A psychic struggle, often unconscious, resulting from the opposition or simultaneous functioning of mutually exclusive impulses, desires, or tendencies.

4. Opposition between characters or forces in a work of drama or fiction, especially opposition that motivates or shapes the action of the plot.”

In Roget’s II: The New Thesaurus a conflict is defined as

“[a] state of disagreement and disharmony.

Synonymys: clash, confrontation, contention, difference, difficulty, disaccord, discord, discordance, dissension, dissent, disentience, dissonance, faction, friction, inharmony, schism, strife, variance, war, warfare.”

The Concise Encyclopedia Britannica provides the following definition of conflict.

“In psychology, a struggle resulting from incompatible or opposing needs, drives, wishes, or demands. Interpersonal conflict represents such a struggle between two or more people, while internal conflict is a mental struggle. A child experiencing internal conflict, for example, may be dependent on his mother but fear her because she is rejecting and punitive. Conflicts that are not readily resolved may cause the person to suffer helplessness and anxiety.”

The Oxford Dictionary lists three conflict definitions.

“1. An overt struggle between individuals or groups. Conflict occurs whenever the action of one person or group prevents, obstructs, or interferes with the goal achievement or action of another person.

---

167 (Dictionary 2000)
168 (Roget 1996)
169 (Encyclopaedia Britannica 2002)
2. A group motive where the group functions together to overcome natural obstacles or the opposition. The group motive will be to beat the opposition, or to struggle against opposing forces, whether those forces are from the natural environment or other people.

3. The tension or stress involved when the satisfaction of specific needs is thwarted by equally attractive or unattractive desires.”

Now we look at conflict definitions suggested by a number of individual conflict scholars including psychologists, political scientists, and sociologists. The political scientists Thorsten Bonacker and Peter Imbusch define social conflicts as follows.

“Social facts which involve at least two parties and are based on differences in the social position and/or interests of the parties involved. [...] In general, social conflicts consist of incompatible expectations of at least two parties. The incompatibility must be perceived by the parties as such and can be reflected, for instance, in a clash of interests or diverging norms.”

The social psychologist Morton Deutsch provides the following definition.

“A broad definition of destructive conflict sees it as a social situation in which there are perceived incompatibilities in goals or values between two (or more) parties, attempts by the parties to control one another, and antagonistic feelings toward each other. [...] The definition stresses that incompatibilities by themselves do not constitute conflict, since the parties could live in peaceful coexistence. However, when there are attempts to control the other party in order to deal with the incompatibility, and when such interactions result in and are fuelled by antagonistic emotions, destructive conflict exists.”

The political scientist Jacob Bercovitch defines a conflict as a

“situation which generates incompatible goals or values among different parties.”

The sociologist Georg Simmel defines conflict as that which is

“designed to resolve divergent dualisms; it is a way of achieving some kind of unity, even if it be through the annihilation of one of the conflicting parties.”

---

170 (Simpson and Weiner 2004)
171 Author’s translation of (Bonacker and Imbusch 1999, p. 75)
172 (Deutsch et al. 2006, p. 178)
173 (Bercovitch 1984, p. 6)
174 (Simmel 1955, p. 13)
For the sociologist Lewis Coser a conflict is

“the clash of values and interests, the tension between what is and what some groups feel ought to be.”\textsuperscript{175}

The educational psychologist Joseph Folger defines conflict as

“the interaction of interdependent people who perceive incompatible goals and interference from each other in achieving those goals.”\textsuperscript{176}

The legal scholar Gregory Tillett provides the following definition.

“[…] conflict exists when two or more parties perceive that their values or needs are incompatible. Values are incompatible if each contradicts or opposes the other (for example, a belief that abortion is always murder and a belief that abortion is not murder). One need would be seen as incompatible with another if meeting that need is thought to prevent, obstruct, interfere with, or in some way make meeting the other need less likely or effective (for example, two job applicants competing for the same, single position.”\textsuperscript{177}

The sociologist Louis Kriesberg defines social conflicts as follows.

“A social conflict exists when two or more parties believe they have incompatible objectives.”\textsuperscript{178}

The sociologist James Schellenberg uses the definition of social conflict

“as the opposition between individuals and groups on the basis of competing interests, different identities, and/or differing attitudes.”\textsuperscript{179}

The political scientist and former diplomat John Burton defines conflict as

“a relationship in which each party perceives the other’s goals, values, interests or behaviour as antithetical to its own.”\textsuperscript{180}

\textsuperscript{175} (Coser 1957, p. 197)
\textsuperscript{176} (Folger et al. 1993, p. 4)
\textsuperscript{177} (Tillett 1999, p. 16)
\textsuperscript{178} (Kriesberg 1982, p. 17)
\textsuperscript{179} (Schellenberg 1996, p. 8)
\textsuperscript{180} (Burton 1993, p. 11)
5.2 General Conflict Definition

With $L_{CML}$, the language of CML, it is possible to describe social conflicts in terms of sets of CML-formulae. The semantics of CML provides an unambiguous interpretation for all the formulae contained in such a set.

However, $L_{CML}$ is a general language able to describe more than just conflicts. Containing operators for propositional attitudes and modalities, as well as their temporal occurrences, $L_{CML}$ can be used to describe all sorts of social situations in which agents hold certain attitudes. Indeed, there is nothing in $L_{CML}$ which makes it exclusive to the description of conflicts.

Although the generality of $L_{CML}$ corresponds with our view of conflicts as a special type of social situations, we have to address the question of what distinguishes a conflict, understood as a set of CML-formulae, from any other arbitrary set of CML-formulae. Suppose we have two sets of CML-formulae, $C_1$ and $C_2$, such that $C_1$ is a description of a social conflict, whereas $C_2$ represents the description of my last birthday party. How can we then formally determine that $C_1$ is the description of a conflict, whereas $C_2$ is not?

In order to distinguish sets of CML-formulae that represent conflicts from sets that do not, we need a precise conflict definition. We provide such a definition in two steps. First, we define a general condition that every set $C$ of CML-formulae has to satisfy in order to be classified as a conflict. In the second step, we define a number of further conditions for specific conflict sets, each of which constitutes a certain type of conflict. We refer to the first condition as the general conflict definition, because it applies to every conflict. The conditions defined in the second step provide the classification scheme of CML.
Looking at the conflict definitions proposed by conflict scholars, we can see that almost all of them include the notion of inconsistency or a related notion, such as incompatibility, clash, competition, disharmony, opposition, difference, etc. Hence, the general characteristic of a conflict seems to be the existence of some kind of inconsistency among its elements. With the axiom system for CML, we have pinned down the notion of consistency/inconsistency of sets of CML-formulae, and, as the axiom system is complete over the semantics of CML, this notion corresponds with the notion of satisfiability/unsatisfiability.

In a first approach, we might think that a set $C$ of CML-formulae must be inconsistent in order to be characterised as a conflict. Inconsistency, however, is too weak as a general condition for conflicts. There are sets that represent typical conflicts, but are not inconsistent within the framework of CML. For instance, the set $C = \{Ga_1p, Ga_2\neg p\}$ represents a conflict in which two agents pursue goals that are directly opposed to each other. However, the set is not inconsistent according to CML.

We have defined inconsistency in such a way that sets like $C$ do not fall under its scope because the fact that two agents pursue contradictory goals does not constitute an inconsistency itself. Indeed, there are many situations in which this is the case. The property characterising sets like $C$ as a conflict, is the fact that they describe inconsistent situations without being inconsistent themselves. For instance, the situation in which both goals contained in $C$ are satisfied at the same time is inconsistent; the fact that two agents pursue contradictory goals is not.

Generally, sets of propositional attitudes, such as beliefs, goals, or norms, describe certain ideal situations. A set of beliefs describes a situation in which all the beliefs are true, i.e. it describes how the world would be if all the beliefs contained in the set were
true. In the case of a set of goals, we get a description of a situation in which all the goals in the set are satisfied. From the point of view of the agents pursuing the goals, such a situation represents a perfect state as all their goals are satisfied. Going back to the example of the set $C = \{Ga_1 p, Ga_2 \neg p\}$, we can represent the situation described by $C$ by the set $C^* = \{p, \neg p\}$. $C^*$ represents a situation in which all the goals contained in $C$ are satisfied. Obviously, $C^*$ is inconsistent and, hence, $C$ is a conflict. The situation is analogous for beliefs and norms. For instance, if we have a set of norms $C = \{Na_1 \phi_1, Na_2 \phi_2, \ldots, Na_n \phi_n\}$, $C$ describes an ideal situation which is represented by the set $C^* = \{\phi_1, \phi_2, \ldots, \phi_n\}$, a situation in which all the norms contained in $C$ are realised. If $C^*$ is inconsistent, then $C$ is a conflict.

Before we state the general conflict definition, we still have to define what it means for a set $C$ to describe another set $C^*$. As illustrated by the example, the set $C^*$, described by a set $C$, represents a situation in which all beliefs contained in $C$ are true, all goals contained in $C$ are satisfied, and all norms contained in $C$ are realised. $C^*$ also contains the formulae that were already contained in $C$.

**Definition 81 (Described Set)**

If $C$ is a finite set of CML-formulae, then $C$ describes the set $C^*$, where $C^*$ is defined by $C^* = C \cup \{\phi | C \vdash_{\text{CML}} \exists y (By\phi \lor Gv\phi \lor Ny\phi)\}$.

The definition makes sure that the content of every belief, goal, and norm held by any agent occurring in $C$ is contained in $C^*$. Not only is the content of propositional attitudes directly contained in $C$ added to $C^*$, but also the content of propositional attitudes that logically follow from $C$. This closure condition is needed as otherwise there could be propositional attitudes contained in $C$ which are not satisfied in $C^*$. Consider, for example, the set $C = \{Ba_1 \phi \land Ba_2 \psi\}$. If $C^*$ contained only the content of propositional attitudes directly contained in $C$, then neither $\phi$ nor $\psi$ would be an element of $C^*$. 
However, if an agent $a_1$ believes $\varphi$ and an agent $a_2$ believes $\psi$, the set $C^*$, representing a situation in which all beliefs contained in $C$ are true, should contain both $\varphi$ and $\psi$.

Having defined $C^*$, we can now state the general conflict definition as follows.

**Definition 82 (General Conflict Definition)**

If a set $C$ of CML-formulae represents a conflict, then $C^*$ is CML-inconsistent.

Note that any inconsistent set $C$ is immediately a conflict as $C \subseteq C^*$.

The definition reflects the view that conflicts are sets of propositional attitudes, such as goals, beliefs, norms, or emotions, which have the property that it is not possible to simultaneously realise all the propositional attitudes contained in the set. The impossibility of realising the propositional attitudes is reflected, in turn, by the inconsistency of the situation resulting from simultaneously realising all propositional attitudes.

The conflict definition does not incorporate the behavioural aspect of conflicts, although actions, or intentions to act, have been proposed as a part of some conflict definitions. The reason why we have not included behavioural components in our model is that the formalisation of action statements requires a complex theoretical framework in itself, which goes beyond the scope of this thesis. However, incorporating a logic of action, such as STIT theory, into CML remains an important task that could be achieved in a further project.
5.3 Synchronic versus Diachronic Conflicts

If we describe a social situation with $L_{CML}$, we obtain a finite set $C$ of CML-formulae. Applying the general conflict definition to $C$, we can check whether $C$ represents a conflict by checking whether $C^*$, the set described by $C$, is inconsistent.

Most conflict definitions, including the ones we have listed in the background section on conflict definitions, make no explicit reference to the temporal dimension of conflicts. They are synchronic definitions, i.e. they state properties of a conflict looked at from one time point. However, the temporal development of conflicts is studied as its own subject in conflict studies, usually under the name “conflict dynamics”. The general approach is to divide up conflicts in stages, phases, or states, i.e. to adopt a diachronic view of conflicts and to look at them from different time points.

To analyse the temporal structure of conflicts, we define conflict sets $C_t$ for every $t$-constant $t$ occurring in $C$. $C_t$ represents the situation at the time point $t$ as described by the agents’ beliefs, goals, and norms, i.e. $C_t$ contains all the formulae that would be true at the time point $t$ if all beliefs, goals, and norms contained in $C$ were satisfied. Every $C_t$ provides a description of a snapshot of the conflict $C$ at a certain time point. Hence, every $C_t$ represents a synchronic description of $C$.

Formally we can define $C_t$ as follows.

**Definition 83 (Synchronic Conflict Description)**

Let $C$ be a conflict. Then, the synchronic conflict description at the time point $t$ is defined by the set $C_t = \{ \varphi | C^* \vdash Rt_{t}\varphi \}$.

To illustrate the definition, we look at the set $C = \{ \text{Ga}_1Rt_0\varphi, \ \text{Ga}_2Rt_0\neg\varphi, \ \text{Ga}_1Rt_1\varphi, \ \text{Ga}_2Rt_1\psi \}$. $C$ describes the set $C^* = C \cup \{ Rt_0\varphi, \ Rt_0\neg\varphi, \ Rt_1\varphi, \ Rt_1\psi \}$. Hence, we get $C_{t_0} = \{ \}$.
\{\varphi, \neg \varphi\} \text{ and } C_{t_0} = \{\varphi, \psi\}. \text{ This shows that the conflict between the agents } a_1 \text{ and } a_2 \text{ can be identified as a disagreement about what the world should be like at the time point } t_0.

As a special case we can build the set } C_n \text{ which represents the conflict at the current time point. } C_n \text{ contains all formulae that would be true at the current time point if all agents’ propositional attitudes were satisfied.}

Constructing the set } C_t \text{ for a number of } t \text{-constants } t_0, t_1, \text{ etc., we obtain a diachronic description of the conflict, i.e. we obtain a description of the development of the conflict through time, which allows us to track its history and its possible future development. The diachronic description is complete if we can construct } C_t \text{ for every } t \text{ occurring in } C. \text{ This leads to the following definition of diachronic conflict description.}

**Definition 84 (Diachronic Conflict Description)**

If } C \text{ is a conflict, then the diachronic conflict description of } C \text{ is defined by } C_{\text{dia}} = \{C_t \mid t \text{ occurs in } C\}.

The complete diachronic description of a conflict can only be constructed if formulae of the form } R_{t\varphi} \text{ can be deduced from } C^* \text{ for all relevant time points } t_0, t_1, \text{ etc. If this is not the case, we only get a partial description of the history of } C. \text{ If no such formulae can be deduced from } C^*, \text{ the best description of the conflict we can get is the set } C^* \text{ itself. } C^* \text{ provides a description of the conflict which makes no reference to its temporal structure.}

### 5.4 Conflicts on the Individual, Micro, and Macro Level

Apart from their temporal structure, we can classify conflicts according to the social level on which they occur, i.e. the social level on which the agents involved in the conflict are located.
The distinction between individual level, micro level, and macro level conflicts is also reflected by some established conflict definitions. Here, categories like interpersonal conflict, group conflict, social conflict, class conflict, etc. are often used to differentiate between the different social levels of conflicts.

As CML itself does not distinguish between different kinds of agents, we rely on other, external theories, capable of distinguishing between agents on the individual level, agents on the micro level, and agents on the macro level. Presupposing that there are such theories, the agents occurring in a conflict \( C \) can be classified according to these categories. Hence, we can talk about individual-agents, micro-agents, and macro-agents.

Typical examples of individual-agents are, for instance, individual persons, such as a family member in a family conflict or a staff member in a work related conflict. Examples of micro-agents are social groups, such as conflicting minority groups or the group of employers and the group of employees in a collective bargaining round. Examples of macro-agents are, for instance, social institutions, such as states, rebel organisations, or political movements involved in a political conflict.

Having identified the social level of the agents involved in a conflict, we can provide the following definition.

**Definition 85 (Individual, Micro, and Macro Level Conflicts)**

If \( C \) is a conflict, then:

1. \( C \) is an individual level conflict \( C_{\text{ind}} \) iff every agent \( a \) occurring in \( C \) is an individual-agent;
2. \( C \) is a micro level conflict \( C_{\text{mic}} \) iff every agent \( a \) occurring in \( C \) is a micro-agent;
3. \( C \) is a macro level conflict \( C_{\text{mac}} \) iff every agent \( a \) occurring in \( C \) is a macro-agent.
We say that an agent \(a\) occurs in a set \(C\) if there is a formula in \(C\), in which the agent constant \(a\) designating the agent \(a\) occurs. As mentioned above, the definition only works in combination with a theory of social strata.

### 5.5 Intra-Agent and Inter-Agent Conflicts

Next, we look at the number of agents involved in a conflict. With regard to this parameter, we establish a dichotomy between intra-agent conflicts on the one hand, and inter-agent conflicts on the other. This dichotomy expresses a common distinction made in many conflict definitions occurring in the literature.

Intra-agent conflicts are conflicts in which only one agent occurs, i.e. situations in which the propositional attitudes held by a single agent describe an inconsistent situation. Inter-agent conflicts involve at least two agents conflicting with each other.

**Definition 86 (Intra- and Inter-Agent Conflicts)**

If \(C\) is a conflict, then:

1. \(C\) is an intra-agent conflict \(C_{\text{intra}}\) iff exactly one agent \(a\) occurs in \(C\);
2. \(C\) is an inter-agent conflict \(C_{\text{inter}}\) among the agents \(a_0, a_1, \ldots, a_n\) iff the agents \(a_0, a_1, \ldots, a_n\) occur in \(C\) and \(1 < n\).

The definition shows that not only conflicts between different parties, but also intra-agent conflicts, can be modelled by CML. From a logical point of view, it does not matter whether the propositional attitudes describing an inconsistent set are held by different agents or only one agent. In both cases, it is the inconsistency of the described set \(C^*\) which qualifies \(C\) as a conflict.
5.6 Subjective versus Objective Conflicts

The distinction between subjective and objective definitions of conflict has been discussed by various conflict scholars. Scholars like Morton Deutsch emphasise the subjectivity of conflict, whereas scholars like John Galtung argue for an objective notion of conflict independent of the parties’ beliefs, goals, norms, or emotions.

The underlying question of this debate is whether the inconsistency among the propositional attitudes is merely an inconsistency perceived by the agents involved in the conflict or whether it is a real inconsistency. If the former is the case, we are dealing with a subjective conflict, if the latter is the case, we are looking at an objective conflict. In CML, we can grasp both the subjective and objective view by providing definitions for both of them.

A subjective conflict exists if at least one agent believes that the goals, beliefs, or norms held by the conflict parties are inconsistent, although they are, in fact, consistent. If the goals, beliefs, or norms held by the agents are inconsistent, then the agents are facing an objective conflict. Within our framework, we can characterise these two notions of conflicts by the following definition.

**Definition 87 (Subjective and Objective Conflicts)**
If $C$ is a set of CML-formulae, then:

1. $C$ is a subjective conflict $C_{\text{sub}}$ iff there is an agent $a$ in $C$ such that $\text{Ba}(\bigwedge_{\neg \in C^*} \bot)$ and $C^* \not\models \bot$, i.e. $C^*$ is consistent;
2. $C$ is an objective conflict $C_{\text{obj}}$ iff $C$ is a conflict.

Note that the definition of objective conflicts is identical with the general conflict definition stated in Definition 82, whereas the definition of subjective conflicts does not satisfy the general conflict definition. Thus, subjective conflicts can be detected by CML, although they do not fall under the general conflict definition.
The above definition establishes a dichotomy: no subjective conflict can be objective and vice versa. However, there are objective conflicts which are also recognised as a conflict by at least one of the conflict parties. Such recognised objective conflicts can be defined as follows.

**Definition 88 (Recognised Objective Conflicts)**

If $C$ is a conflict, then $C$ is a recognised objective conflict $C_{rec}$ iff there is an agent $a$ in $C$ such that $Ba(\bigwedge_{qoc^a} \supset \bot)$ and $C$ is a conflict.

Furthermore, the definition of subjectivity is relative to the agents involved in the conflict, i.e. we can identify conflicts which are recognised by some agents but not by others. For instance, a situation described by a set $C$ of CML-formulae can be perceived as a conflict by an agent $a_1$, i.e. $Ba_1(\bigwedge_{qoc^a} \supset \bot)$ holds, whereas another agent $a_2$ does not consider $C$ a conflict, i.e. $Ba_2(\bigwedge_{qoc^a} \supset \bot)$ does not hold.

### 5.7 Conflict Issues

In the previous three sections, we have focused on structural properties of $C$. First, we looked at the temporal structure of $C$, then we classified conflicts according to the social strata and the number of agents involved in them, and, finally, we assessed the real or perceived inconsistency of $C$. In the following sections, we analyse the nature of conflicts in terms of the type of propositional attitudes constituting them.

Conflicts can arise because agents have incompatible, i.e. inconsistent, beliefs about the world, because they pursue inconsistent goals, or because they hold norms that are mutually incompatible. Using the framework of CML, we can identify each such type of conflict by looking at the propositional attitude operators used to describe them. The
distinction between factual disputes, goal conflicts, value conflicts, etc. is also reflected in
the literature.

5.7.1 Factual Disputes

First, we define a factual dispute as a conflict C which is constituted by inconsistent
beliefs.

**Definition 89 (Factual Dispute)**
If C is a conflict, then \( C_b \) is a factual dispute iff \( C_b = \{ \varphi | C \vdash_{CML} \exists yBy\varphi \} \) is
inconsistent.

\( C_b \) represents a situation in which all beliefs contained in C are realised. It represents the
world as jointly believed to be the case by the agents occurring in C. If \( C_b \) is inconsistent,
the agents disagree about the way the world is structured and are, therefore, engaged in a
factual dispute. Note that \( C_b \subseteq C^* \) as \( C \vdash_{CML} \exists yBy\varphi \) implies \( C \vdash_{CML} \exists y(By\varphi \lor Gy\varphi \lor
Ny\varphi) \).

5.7.2 Conflicts over Goals

Conflicts over goals are characterised by the fact that the goals pursued by a group of
agents are inconsistent. This can be defined as follows.

**Definition 90 (Conflicts over Goals)**
If C is a conflict, then \( C_g \) is a conflict over goals iff \( C_g = \{ \varphi | C \vdash_{CML} \exists yGy\varphi \} \) is
inconsistent.

\( C_g \) represents a situation in which all goals contained in C are satisfied. It represents the
world as the agents in C jointly desired it to be. If \( C_g \) is inconsistent, the agents pursue
goals which are impossible to satisfy simultaneously and are, hence, engaged in a conflict over goals. Again we have \( C_g \subseteq C^* \).

5.7.3 Conflicts over Values

Finally, we define a conflict over values as a set of formulae which has the property that the norms contained in the set describe an inconsistent set.

**Definition 91 (Conflicts over Values)**
If \( C \) is a conflict, then \( C_n \) is a conflict over values iff \( C_n = \{ \phi | C \vdash_{CML} \exists y N y \phi \} \) is inconsistent.

\( C_n \) describes a situation in which all norms contained in \( C \) are abode by, i.e. a situation that is jointly considered to be morally or legally desirable by the agents occurring in \( C \). If \( C_n \) is inconsistent, it is impossible to abide by all norms held by the agents simultaneously and, hence, the agents are facing a conflict over values. Again, \( C_n \) is a subset of \( C^* \).

5.8 The Emotional Load of Conflicts

The emotion operator \( E \) does not constitute its own type of conflict as emotional excitement expressed by \( E \) is not directional. As a consequence, it is not necessarily the case that inconsistencies among the content of emotions held by a group of agents gives rise to a conflict among these agents. For example, the two formulae \( E_a \phi \) and \( E_a \neg \phi \) just express that the agent \( a_0 \) is emotionally excited because of \( \phi \) and the agent \( a_1 \) is excited because of \( \neg \phi \). As the type of excitement is not specified, we cannot conclude that \( a_0 \) and \( a_1 \) are facing a conflict with regard to their emotions.
However, the emotional load of a conflict \( C \), i.e. the level of emotional involvement of the agents in the conflict \( C \), can be expressed by the number of emotions contained in \( C \). The more emotions contained in \( C \), of the form \( E \alpha \varphi \), the higher the degree of emotional involvement in the conflict. This gives rise to the following definition of the emotional content \( C_e \) of a conflict and the emotional load \( emo(C) \) of a conflict \( C \).

**Definition 92 (Emotional Content and Degree of Emotional Involvement)**

If \( C \) is a conflict, then:

1. \( C_e = \{ \varphi \mid C \vdash_{CML} \exists \gamma E \gamma \varphi \} \) is the emotional content of \( C \);
2. \( emo(C) = |C_e| \), i.e. the size of \( C_e \) is the emotional load of \( C \).

## 5.9 Combining the Definitions

Starting with a general conflict description \( C \), i.e. a set of CML-formulae having the property that \( C^* \), the set described by \( C \), is inconsistent, we can simultaneously apply the conflict definitions introduced in the previous sections. This allows us to characterise the specific nature of the conflict and to identify and classify the sub-conflicts constituting it.

In order to identify and classify the sub-conflicts of a conflict \( C \), we have to combine the specific conflict definitions of CML. This is necessary as conflicts typically consist of a number of different sub-conflicts, such as goal conflicts, value conflicts, factual disputes, etc., which occur at different time points within the temporal development of the conflict, and which are sometimes recognised by some parties and sometimes not. Also, a conflict can occur among parties on different social levels and can contain intra-agent, as well as inter-agent, components.

For a given conflict \( C \), we can, for instance, identify two of its sub-conflicts \( C_{t_0} \) and \( C_{t_1} \) occurring at the time points \( t_0 \) and \( t_1 \), respectively. \( C_{t_0} \) could, for instance, be further classified as a subjective, inter-agent, micro level goal conflict, and \( C_{t_1} \) as a recognised
The Classification Scheme of Conflict Modelling Logic

objective, inter-agent, macro level conflict over goals and values. This example merely illustrates how the definitions can be combined. Depending on the available information about a particular conflict, we can identify and classify different sub-conflicts.

5.10 Conflicts in Propositional Form

In order to be able to input sub-conflicts described by CML-formulae into the resolution algorithms to be introduced in chapter 7, we have to transform them into a unique form called the propositional form \( C_p \). A sub-conflict \( C_p \) in propositional form is a description of the original sub-conflict \( C \) in the language of propositional logic, i.e. all non-propositional symbols are removed from \( C \).

Similar to the definition of the modal core or the temporal core of a formula, we can define the propositional core of a CML-formula by providing a set of translation rules which allow us to transform arbitrary CML-formulae into propositional formulae. We obtain the propositional core \( \varphi^* \) of a CML-formula \( \varphi \) by successively translating sub-formulae of \( \varphi \) according to the principles stated in the following definition.

**Definition 93 (Propositional Core)**
The modal core \( \varphi^* \) of a CML-formula \( \varphi \) is obtained by the following rules:

1. \( (p)^* = p \);
2. \( (\neg \varphi)^* = \neg (\varphi)^* \);
3. \( (\varphi \land \psi)^* = (\varphi)^* \land (\psi)^* \);
4. \( (\Box \varphi)^* = (\forall \varphi)^* = (R_{\tau_0} \varphi)^* = (\tau_0 < \tau_1)^* = (\tau_2 = \tau_3)^* = (O\varphi)^* = p_n \), where \( p_n \) is a fresh propositional constant and \( O \in \{B, G, N, E\} \).

Note that aspects of the logical system CML are still reflected insofar as the original description \( C \) of a conflict is a set of CML-formulae and we have to make use of the operators of CML in order to single out the sub-conflicts from \( C \). However, the development of resolution algorithms that allow one to directly input sets of CML-formulae would be a promising research project.
The definition lists four translation rules for CML-operators assuming that translations for sub-formulae are already available. The fourth condition ensures that all non-propositional operators are eliminated. Sub-formulae, whose main operators are not propositional connectives, are replaced by new propositional constants.

As all non-propositional operators are removed when translating a CML-formula \( \varphi \) into its propositional core \( \varphi^* \), \( \varphi^* \) is a formula of propositional logic.

**Theorem 40  (Propositional Core and Propositional Logic)**
The propositional core \( \varphi^* \) of every CML-formula \( \varphi \) is a formula of propositional logic.

**Proof**
The proof is an induction on the formation of formulae.

QED

Having defined the propositional core of arbitrary CML-formulae, we can now define the propositional form of a conflict \( C \) as follows.

**Definition 94 (Propositional Form)**
If \( C \) is a set of CML-formulae representing a conflict, then the propositional form of \( C \) is defined by \( C_p = \{ \varphi^* \mid \varphi \in C \} \).

**Example: Second Congo War**

In the following section, we apply some of the conflict definitions introduced in this chapter to our example of the Second Congo War. The classification scheme allows one to identify and categorise some sub-conflicts within this war.

We assume that \( C \) is a description of the Second Congo War in terms of CML-formulae. To illustrate the concept of \( C^* \), i.e. the set of formulae described by \( C \), we assume that the four formula \( Ga_1p, Ga_1\neg q, Ga_2p \), and \( Ga_2(p \supset q) \) are contained in \( C \). The
interpretation of the agent constants \( a_1 \) and \( a_2 \), and the propositional constants \( p \) and \( q \), is given as follows.\(^{182}\)

\[
\begin{array}{l}
a_1: \text{Rassemblement Congolais pour la Democratie (RCD)} \\
a_2: \text{Kabila government} \\
p: \text{The combatants of the RCD are integrated into the Congolese national army.} \\
q: \text{High ranking officers of the RCD are replaced.}
\end{array}
\]

Hence, the four formulae express the following claims.

\[
\begin{array}{l}
\text{Ga}_1 p: \text{The RCD wants its combatants to be integrated into the Congolese national army.} \\
\text{Ga}_1 \neg q: \text{The RCD does not want its high ranking officers being replaced.} \\
\text{Ga}_2 p: \text{The Kabila government wants the combatants of the RCD to be integrated into the Congolese national army.} \\
\text{Ga}_2 (p \supset q): \text{The Kabila government wants the combatants of the RCD only to be integrated into the Congolese national army if its high ranking officers are replaced.}
\end{array}
\]

\{Ga_1 p, Ga_1 \neg q, Ga_2 p, Ga_2 (p \supset q)\} \subseteq C

Looking at the set \( C^* \) described by \( C \), we get the following membership relations.

\[
\begin{array}{l}
\{p, \neg q, p \supset q\} \subseteq C^* \\
p \in C^*, \text{ because } C \vdash \exists y G y p \\
\neg q \in C^*, \text{ because } C \vdash \exists x G x \neg q \\
p \supset q \in C^*, \text{ because } C \vdash \exists x G x (p \supset q) \\
C^* \vdash \bot
\end{array}
\]

From this we can conclude that \( C \) represents a conflict as \( C^* \), the set described by \( C \), is inconsistent. Note that \( C \) itself is not inconsistent and that the two agents partly pursue the same goals. However, the situation resulting from all agents’ goals being realised

---

\(^{182}\) The integration of the combatants of the RCD was a major issue in the peace negotiations of the conflict. For details of the integration process, see chapter 10 and 11 of the Lusak Ceasefire Agreement. (UNSC 1999)
simultaneously, expressed by \( C^* \), is inconsistent as it would require the high ranking officers of the RCD to be replaced and not to be replaced at the same time.

Next, we illustrate the concept of a synchronic conflict description at a time point \( t \). Let us assume that \( C \) contains the formulae \( G_{a_i}R_{t_1}p \), \( G_{a_i}R_{t_1}\neg q \), and \( G_{a_2}R_{t_1}(p \supset q) \) as well as the formulae \( B_{a_1}R_{t_2}r \) and \( B_{a_2}R_{t_2}\neg r \), where the two temporal constants \( t_1 \) and \( t_2 \) and the propositional constant \( r \) have the following interpretations.

\[
\begin{align*}
  t_1 &\colon 2003 \\
  t_2 &\colon 1998 \\
  r &\colon \text{Banyamulenge are threatened to be attacked by Hutu-aligned forces in Kivu.}
\end{align*}
\]

\( \{G_{a_1}R_{t_1}p, G_{a_1}R_{t_1}\neg q, G_{a_2}R_{t_1}(p \supset q), B_{a_1}R_{t_2}r, B_{a_2}R_{t_2}\neg r\} \subseteq C \)

Then, \( C_{t_1} \), the synchronic conflict description of the Second Congo War in 2003, and \( C_{t_2} \), the synchronic conflict description of the Second Congo War in 1999, are given as follows.

\[
\begin{align*}
  \{p, \neg q, p \supset q\} &\subseteq C_{t_1} \\
  p &\in C_{t_1} \text{ because } C^* \vdash R_{t_1}p \\
  \neg q &\in C_{t_1} \text{ because } C^* \vdash R_{t_1}\neg q \\
  p \supset q &\in C_{t_1} \text{ because and } C^* \vdash R_{t_1}(p \supset q) \\
  C_{t_1} &\vdash \bot \\
  \{r, \neg r\} &\subseteq C_{t_2} \\
  r &\in C_{t_2} \text{ because } C^* \vdash R_{t_2}r \\
  \neg r &\in C_{t_1} \text{ because } C^* \vdash R_{t_2}\neg r \\
  C_{t_1} &\vdash \bot
\end{align*}
\]

Both \( C_{t_1} \) and \( C_{t_2} \) are inconsistent. Thus, the goals of the Kabila government and the RCD, with regard to the situation of the conflict in 2003, constitute a conflict as do the beliefs shared by the two parties with regard to the situation in 1999. The factual dispute
between the two parties with regard to the situation in 1999 is straightforward: The RCD believes that Banyamulenge are threatened to be attacked by Hutu-aligned forces in Kivu, whereas the Kabila government believes that this is not the case. $C_1$ and $C_2$ express how the situation would have looked in 2003 and 1999, respectively, if all goals of the two parties had been realised and all beliefs had been true.

The distinction between subjective, objective, and recognised objective conflicts can be used when analysing the Second Congo War as follows. Using the two agent constants $a_1$ and $a_2$ and the propositional constants $p$ and $q$ as above, we assume that $Ga_1 p \in C$, $Ga_1 \neg q \in C$, $Ga_2 p \in C$ and $Ga_2 (p \supset q) \in C$.

$$\{Ga_1 p, Ga_1 \neg q, Ga_2 p, Ga_2 (p \supset q)\} \subseteq C$$
$$\{p, \neg q, p \supset q\} \subseteq C^*$$

If $Ba_1 (p \land \neg q \land (p \supset q) \supset \bot)$ or $Ba_2 (p \land \neg q \land (p \supset q) \supset \bot)$ holds, then $C_{rec}$

If neither $Ba_1 (p \land \neg q \land (p \supset q) \supset \bot)$ nor $Ba_2 (p \land \neg q \land (p \supset q) \supset \bot)$ holds, then $C_{obj}$

The reason why $C$ can be classified as a recognised objective conflict in the first case, but not in the second case, is that in the first case either the RCD or the Kabila government is aware of the fact that their goals clash and, therefore, constitute a conflict, whereas in the second case, their goals clash but neither of the two parties recognise this clash.

A subjective conflict can be identified as follows. Assume that $Ga_1 s \in C$ and $Ga_2 t \in C$ where the interpretation of $s$ and $t$ are given as follows.

$$\{Ga_1 s, Ga_2 t\} \subseteq C$$
$$\{s, t\} \subseteq C^*$$

If $Ba_1 (s \land t \supset \bot)$ or $Ba_2 (s \land t \supset \bot)$ holds, then $C_{sub}$
In this case, at least one of the two parties falsely believes that the exploitation of the natural resources in the Kivu provinces cannot be jointly controlled by the RCD and the Kabila government. However, various forms of joint control over the natural resources in the Kivu provinces are possible and have been proposed by mediators.

Finally, we identify a factual dispute, a goal conflict, and a value conflict within the Second Congo War. Assume that $\text{Ba}_1 r \in C$ and $\text{Ba}_2 \neg r \in C$.

\begin{center}
$\text{r: Banyamulenge are threatened to be attacked by Hutu-aligned forces in Kivu}$
\end{center}

\begin{itemize}
  \item \{Ba$\text{r}$, $\text{Ba}_2 \neg r\} \subseteq C$
  \item \{r, $\neg r$\} $\subseteq C_h$
  \item $r \in C_h$ because $C \vdash \exists y \text{Byr}$
  \item $\neg r \in C_h$ because $C \vdash \exists y \text{By} \neg r$
  \item $C_h \vdash \bot$
\end{itemize}

$C_h$ expresses the factual dispute between the RCD and the Kabila government about the threat of Banyamulenge being attacked by Hutu-aligned forces in Kivu: the RCD believes that there is such a threat, whereas the Kabila government denies this.\(^{183}\)

A goal conflict within the Second Congo War has already been described earlier in the section. We just repeat it here, stressing the aspect of it being a goal conflict.

\begin{center}
$\text{Ga}_1 p, \text{Ga}_1 \neg q, \text{Ga}_2 p, \text{Ga}_2 (p \supset q) \subseteq C$
\end{center}

\begin{itemize}
  \item \{p, $\neg q$, $p \supset q$\} $\subseteq C_g$
  \item $p \in C_g$ because $C \vdash \exists y \text{Gyp}$
  \item $\neg q \in C_g$ because $C \vdash \exists y \text{Gy} \neg q$
  \item $p \supset q \in C_g$ because $C \vdash \exists y \text{Gy}(p \supset q)$
  \item $C_g \vdash \bot$
\end{itemize}

\(^{183}\) Indeed the main justification of the RCD for their military action was the protection of Banyamulenge against Hutu-aligned forces, such as the FDLR, in the Kivu provinces. See (ICG 2005).
Hence, $C_g$ is inconsistent and expresses the goal conflict between the RCD and the Kabila government with regard to the integration of the combatants of the RCD into the Congolese national army.

A value conflict between the two parties RCD and Hutu-aligned forces, represented by the agent constants $a_1$ and $a_3$, respectively, can be identified as follows.

<table>
<thead>
<tr>
<th>$a_1$: Rassemblement Congolais pour la Democratie (RCD)</th>
<th>$a_2$: Hutu-aligned forces</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u$: The perpetrators of the 1994 genocide in Rwanda are extradited from the territory of the DRC</td>
<td></td>
</tr>
<tr>
<td>${Na_1u, Na_3u} \subseteq C$</td>
<td></td>
</tr>
<tr>
<td>${u, \neg u} \subseteq C_n$</td>
<td></td>
</tr>
<tr>
<td>$u \in C_n$ because $C \vdash \exists y Nyu$</td>
<td></td>
</tr>
<tr>
<td>$\neg u \in C_n$ because $C \vdash \exists yNy \neg u$</td>
<td></td>
</tr>
<tr>
<td>$C_g \vdash \bot$</td>
<td></td>
</tr>
</tbody>
</table>

Hence, $C_n$ is inconsistent and expresses the value conflict between the RCD and the Hutu-aligned forces about the extradition of the perpetrators of the 1994 genocide in Rwanda from the territory of the DRC. The RCD considers it a moral obligation to extradite the perpetrators, whereas the Hutu-aligned forces consider it a moral obligation to keep them in the country.\(^{184}\)

### 5.11 Summary

The aim of this chapter was to provide a general conflict definition and a classification scheme for conflicts. This has been achieved to the extent that we have provided a general conflict definition based on the idea that the defining feature of a conflict is an

---

\(^{184}\) (Magsam 2005)
inconsistency in the set of propositions described by the attitudes of the agents involved in the conflict. Inconsistency, as such, is too weak to function as a general conflict definition because, as we showed in the chapter, there are consistent sets of formulae which, nevertheless, express conflicts. However, inconsistency is sufficient for a conflict definition as any inconsistent set immediately constitutes a conflict.

A number of definitions for specific sub-conflicts have been established by singling out certain subsets of a conflict and checking the consistency of the sets described by them. In particular, we have provided definitions for synchronic and diachronic conflicts, for subjective, objective, and recognised objective conflicts, for individual-, micro-, and macro level conflicts, for intra- and inter-agent conflicts, for factual disputes, conflicts over goals, and conflicts over values. Overall, we have established a classification scheme which allows us to identify the specific type of the sub-conflicts that constitute a conflict.

Having provided a general conflict definition and a classification scheme based on concepts of the syntax, semantics, and axiomatics of CML, we can now check whether a set of CML-formulae constitutes a conflict or not. Furthermore, we can identify sub-conflicts within a conflict and describe them along the dimensions of our classification scheme. The process of defining and classifying a conflict, as well as identifying its sub-conflicts, is the fourth step in the process of modelling and resolving conflicts.

In the following chapter, we will introduce two measures which can be used to assess the elements of conflicts: the claims made by the conflicting parties and the conflict as a whole. This will give us a quantitative measure for assessing conflicts.
CHAPTER 6

Measuring Conflicts

Potential Conflict Power and Degrees of Inconsistency

6.1 Introduction

The aim of chapter 6 is to introduce the concept of potential conflict power, a measure of how likely it is for a propositional formula to be inconsistent with arbitrary other propositional formulae, and the concept of degrees of inconsistency, a measure of the depth of inconsistency of a set of propositional formulae.

The concept of potential conflict power can be used to assess and compare conflict elements, such as goals, beliefs, norms, and emotions with respect to the role they play in a given conflict. We provide a semantic and a syntactic definition for the potential conflict power which we then prove to be equivalent to each other.

The degree of inconsistency of a set of propositional formulae is intended to express how easy it is to resolve the conflict described by the set. We discuss three proposals for measuring this concept and argue for the superiority of one of the three approaches.

As a result of the chapter, we will be able to give answers to the following questions.
Chapter 6  Measuring Conflicts

- How can the potential conflict power of propositional formulae be measured?
- How can the potential conflict power of propositional formulae be defined semantically?
- How can the potential conflict power of propositional formulae be defined syntactically, and why is this definition equivalent to the semantic definition?
- How can the degree of inconsistency of an inconsistent set of propositional formulae be measured?
- How can claims made by parties in the Second Congo War be assessed with respect to their potential conflict power and how can sub-conflicts of the Second Congo War be assessed with respect to their degree of inconsistency?

An outline of the chapter is as follows: We start with a background section on the AGM model, the prevalent theory of belief revision. In sections 6.2 and 6.3, we define the potential conflict power of propositional formulae, first semantically and then syntactically, and show that the two definitions are equivalent. In section 6.4, we explore some properties of the potential conflict power. In particular, we look at how the measure behaves with respect to the logical connectives and the consequence relation. We also compare it with Peter Gärdenfors’s binary relation of epistemic entrenchment. Section 6.5 deals with the definition of the degree of inconsistency of a set of propositional formulae.

Background: Belief Revision

The theory of belief revision has been developed as a formal framework to model changes taking place in the belief sets of agents. A typical situation in which such changes occur is a situation where an agent receives new information. If the new information is to
be added to the agent’s beliefs, it might be necessary to revise the agent’s already existing beliefs, i.e. we may have to delete those beliefs which are incompatible or inconsistent with the new information.

The most influential model of belief revision is the so-called AGM model. It was initially presented by Carlos Alchourrón, Peter Gärdenfors, and David Makinson in their 1985 paper *On the Logic of Theory Change: Partial Meet Contraction and Revision Functions*. Their model combines two research traditions in the area of belief change. In the philosophy of science tradition, philosophers were concerned with the development of scientific theories over time and criteria for rational belief change. In the computer science tradition, procedures to update databases were developed in order to maintain the consistency of a database.

Following the seminal paper by Alchourrón, Gärdenfors, and Makinson, a large number of papers on the subject has been published in which the AGM model has been further developed and applied to new areas. The model, which was originally based on classical propositional logic, has been applied to non-classical logics, such as paraconsistent logic, non-monotonic logic, and intuitionistic logic. A comprehensive monograph on the AGM model is Gärdenfors’s 1988 *Knowledge in Flux*. A textbook on the topic was published by Sven Hansson in 1999.

The basic concept of the AGM model is the notion of a belief set. In order to define belief sets, we first have to provide a representation for beliefs and specify the concept of

---

185 (Alchourrón et al. 1985)
186 (Mares 2002; Tanaka 2005)
187 (Bochman 2001; Rott 2001)
188 (Tennant 2005)
189 (Gärdenfors 1988)
190 (Hansson 1999b)
a consequence operator $Cn$. It is assumed that beliefs are represented by sentences $p$, $q$, $r$, etc. of a formal language, such as the language of classical propositional logic. Sets of beliefs are expressed by capital letters $A$, $B$, $C$, etc. A consequence operator $Cn$ is assumed to be a function assigning a set of beliefs $Cn(A)$ to every set of beliefs $A$.

In the original AGM model $Cn$ is assumed to be supraclassical, inclusive, monotone, and iterative, i.e. for all beliefs $p$ and all sets of beliefs $A$, it is assumed that

1. If $A \vdash_{rc} p$, then $p \in Cn(A)$ (Supraclassicality);
2. $A \subseteq Cn(A)$ (Inclusion);
3. If $A \subseteq B$, then $Cn(A) \subseteq Cn(B)$ (Monotony);
4. $Cn(A) = Cn(Cn(A))$ (Iterativity).\(^{192}\)

A belief set can now be defined as a set of beliefs $K$ which is closed under $Cn$, i.e. for which the identity $K = Cn(K)$ holds. Belief sets contain all the logical consequences of their members. In the following, $K$ will always be assumed to be a belief set.

As belief sets are infinite and require the epistemic agent holding the beliefs to be logically omniscient, they only provide an idealisation of the sets of beliefs held by real epistemic agents. They express what an agent is committed to believe rather than what he actually believes.\(^{193}\)

In addition to the concept of logically closed belief sets, the AGM theory also deals with simple sets of beliefs, i.e. collections of beliefs of the form $A = \{p, q, r, \text{etc.}\}$ which are

\(^{191}\) Note that sets of beliefs are arbitrary collections of beliefs, such as $A = \{p, q, r\}$, whereas the term “belief set” is reserved to only one specific type of sets of beliefs as defined below.

\(^{192}\) See (Hansson 2006).

\(^{193}\) (Levi 1991)
not closed under \( Cn \). Such sets are called belief bases. Although they more closely represent the beliefs held by real agents, they are technically harder to deal with.\(^{194}\)

For a given belief set \( K \), three basic change operators were proposed and defined in the original AGM model. These are expansion, contraction and revision. Expansion is the addition of a sentence \( p \) to a belief set \( K \) such that the newly obtained set \( K+p \) is the smallest, logically closed superset containing \( p \). Contraction is the removal of a sentence \( p \) from a belief set \( K \) such that the newly obtained set \( K\setminus p \) is a subset of \( K \) not containing \( p \). Revision is the addition of a sentence \( p \) to a belief set \( K \) such that the newly obtained set \( K*p \) is consistent and contains \( p \). In order to ensure the consistency of \( K*p \), it might be necessary to remove some elements of \( K \). All three change operators can be applied to either belief sets or belief bases.

Further operators of belief change have been proposed. Among them are operators for update, consolidation, semi-revision, selective revision, screened revision, shielded contraction, replacement, multiple contraction, multiple revision, indeterministic belief change, as well as special operations for extended languages.\(^{195}\) In the following, we will describe the three basic operations of expansion, contraction, and revision. Then, we will look at the operations of consolidation\(^ {196}\) and screened revision as these two operations are of particular interest for the process of conflict resolution.\(^ {197}\)

Expansion is the simplest operation of belief change. Indeed, there is not much to say about it as the expansion \( K+p \) of a belief set \( K \) by a sentence \( p \) is simply defined as the set \( Cn(K \cup \{p\}) \). As expansion, in contrast to revision, does not require making the newly

\(^{194}\) For a monograph on belief base change, see (Williams 1994).

\(^{195}\) For an overview of these and other operators of belief change, see (Williams and Rott 2001).

\(^{196}\) (Hansson 1991; Hansson 1999a; Fuhrmann 1997)

\(^{197}\) (Makinson 1997; Hansson 1999a)
obtained belief set consistent, we can just add the sentence $p$ to $K$, and then apply the consequence operator.

At first glance, the contraction operation looks equally simple. However, in actuality the situation is more complicated. In order to remove a sentence $p$ from a belief set $K$, it is not enough to just delete $p$ from $K$ and then apply the consequence operator $Cn$ because, being closed under logical consequence, the set $Cn(K\{p\})$ might contain $p$ again.\footnote{This is the case if $K\{p\} \vdash p$.} We have to make sure that the result of the contraction $K\lhd p$ does not logically imply $p$.

A trivial candidate for $K\lhd p$ is $\emptyset$. However, this would result in a complete loss of information; we would lose all beliefs previously held by the agent. In order to minimise, or at least to control the informational loss accompanying the contraction of a belief set $K$ by $p$, the authors of the AGM model proposed looking at inclusion-maximal subsets of $K$ not implying $p$. These sets can be defined as follows. If $K$ is a belief set, then $A \subseteq K$ is an inclusion-maximal subset of $K$ not implying $p$ if and only if $A \not\vdash p$ and there is no $B \subseteq K$ such that $B \not\vdash p$ and $A \subseteq B$. The set of all inclusion-maximal subsets of a belief set $K$ not implying $p$ is called the remainder set of $K$ and denoted by $K\perp p$.

With $K\perp p$, we have a set of potential candidates for $K\lhd p$ as every element of $K\perp p$ is a subset of $K$ not implying $p$. Hence, we can apply the consequence operator to elements of $K\perp p$ without thereby adding $p$. Furthermore, if minimising informational loss was the only criterion we cared about, $K\lhd p$ could be chosen as an element of $K\perp p$. However, there are reasons for choosing a selection of elements of $K\perp p$, and then defining $K\lhd p$ as the intersection of the selected elements. This approach to contraction is called partial meet contraction.
For partial meet contraction, a selection function \( \gamma \) is defined which assigns a selection of elements of the remainder set \( K \perp p \) to every pair consisting of a belief set \( K \) and a sentence \( p \), i.e. \( \gamma(K, p) \subseteq K \perp p \). It is not specified which elements of \( K \perp p \) are to be selected by \( \gamma \). Intuitively, \( \gamma \) is assumed to select the ‘best’ elements of \( K \perp p \), where ‘best’ is defined by some external criterion. For a given selection function \( \gamma \), partial meet contraction is defined by the intersection of the elements of \( K \perp p \) selected by \( \gamma \), i.e. \( K^+p = \bigcap \gamma(K, p) \).

Two extreme cases of partial meet contraction can be singled out. The case in which \( \gamma(K, p) \) has only one element, i.e. \( K^+p \in K \perp p \), is called maxichoice contraction. The case in which \( \gamma(K, p) = K \perp p \), i.e. in which \( \gamma \) selects all elements of \( K \perp p \), is called full meet contraction.

Partial meet contraction is characterised by the six basic AGM postulates, each of which describes a property of the partial meet contraction operation.\(^{199}\)

**Basic AGM Postulates**

1. \( K^+p = Cn(K^+p) \) (Closure);

2. If \( p \notin Cn(\emptyset) \), then \( p \notin Cn(K^+p) \) (Success);

3. \( K^+p \subseteq K \) (Inclusion);

4. If \( p \notin K \), then \( K^+p = K \) (Vacuity);

5. If \( p \equiv q \in Cn(\emptyset) \), then \( K^+p = K^+q \) (Extensionality);

6. \( K \subseteq (K^+p)^+p \) (Recovery).

\(^{199}\) (Alchourrón et al. 1985, p. 513)
The closure condition expresses that the result of partial meet contraction is itself a belief set. The success condition stipulates that p is actually removed by the contraction. This condition is proviso p not being a tautology (as expressed by the antecedent p \( \not\in \text{Cn}(\emptyset) \)). The inclusion condition ensures that nothing is added to the belief set by partial meet contraction. Vacuity claims that if p is not in K, and, therefore, not entailed by K, nothing needs to be removed. Extensionality expresses that logically equivalent sentences are treated equivalently with regard to partial meet contraction, i.e. they are either removed together or maintained together. The recovery condition, which has been disputed most among the six postulates, claims that the set obtained from contracting a belief set K by p, and then expanding the resulting set by p, contains at least all the sentences originally contained in K.

A representation theorem for partial meet contraction, expressing that an operator + is an operator of partial meet contraction if and only if it satisfies the basic AGM postulates, was proven by Alchourrón, Gärdenfors, and Makinson.²⁰⁰

Besides partial meet contraction, there are stronger contraction operations in the sense that they satisfy further postulates. Two such operations are transitively relational partial meet contraction and entrenchment-based contraction. It was proven that these two contraction operations are equivalent.²⁰¹ In the following, we briefly describe the entrenchment-based approach.

The idea behind entrenchment-based contraction is that not all beliefs of a belief set are of equal value. Some of the beliefs are more valuable than others to the extent that they have more ‘explanatory power’ or ‘informational value’ than others. In the context of

²⁰⁰ (Alchourrón et al. 1985, p. 513)
²⁰¹ (Gärdenfors and Makinson 1988)
contraction, this means that when deciding which beliefs to give up and which to maintain, one should give up beliefs of lower value and maintain beliefs of higher value.

In the AGM model the value relation between beliefs is expressed by a binary relation $\leq$ of epistemic entrenchment. Gärdenfors and Makinson describe the entrenchment relation as follows.

> “Even if all sentences in a knowledge set are accepted or considered as facts (so that they are assigned maximal probability), this does not mean that all sentences are of equal value for planning or problem-solving purposes. Certain pieces of our knowledge and beliefs about the world are more important than others when planning future actions, conducting scientific investigations, or reasoning in general. We will say that some sentences in a knowledge system have a higher degree of epistemic entrenchment than others. This degree of entrenchment will, intuitively, have a bearing on what is abandoned from a knowledge set, and what is retained, when a contraction or a revision is carried out.”

Expressions of the form $p \leq q$ are read as “$p$ is, at most, as entrenched as $q$”. Two further relations, $\equiv$ and $<$, which are read as “$p$ is equally entrenched as $q$” and “$p$ is less entrenched than”, respectively, are defined as follows.

1. $p \equiv q$ iff $p \leq q \land q \leq p$;
2. $p < q$ iff $p \leq q \land \neg(q \leq p)$.

The entrenchment relation is characterised by the five postulates for epistemic entrenchment, each of which expresses a property of the entrenchment relation.

**Postulates for Epistemic Entrenchment**

1. If $p \leq q$ and $q \leq r$, then $p \leq r$ (Transitivity)
2. If $p \vdash q$, then $p \leq q$ (Dominance)
3. For any $p$ and $q$, $p \leq p \land q$ or $q \leq p \land q$ (Conjunctiveness)

---

202 (Gärdenfors and Makinson 1988, p. 88)
203 (Gärdenfors and Makinson 1988, p. 89)
Chapter 6  Measuring Conflicts

(4) If K is consistent, then $p \not\in K$ iff $p \leq q$, for all $q$ (Minimality)

(5) If $q \leq p$ for all $q$, then $p \in Cn(\emptyset)$ (Maximality)

As a consequence of the postulates, it can be shown that the entrenchment relation is connective, i.e. for all $p$ and $q$ either $p \leq q$ or $q \leq p$.

Gärdenfors gives justifications for each of the five postulates. We just look at his justification for the dominance condition. If $p$ logically entails $q$, and we are forced to give up either $p$ or $q$, there would be less informational loss in giving up $p$ and maintaining $q$ than giving up $q$ since giving up $q$ would immediately force us to give up $p$ as well. Maintaining $p$ would force us to maintain $q$ because $q \in Cn(\{p\})$. Hence, if $p \vdash q$, we should rather give up $p$ than $q$, and, thus, $p$ is less entrenched than $q$.

The entrenchment relation can be combined with the contraction operation in two ways. Starting with a belief set $K$, we can define an epistemic entrenchment ordering using a given contraction operation, or we can use a given entrenchment relation to define an operation of entrenchment-based contraction. Gärdenfors provides conditions for both approaches and proves various theorems for them. We just state the conditions without explaining or justifying them. Motivations for the conditions can be found in Gärdenfors and Makinson’s paper Revisions of Knowledge Systems Using Epistemic Enrenchment.  

The condition for defining an entrenchment relation $\leq$ for a given belief set $K$ and contraction operation $\vdash$ is given as follows.

$$p \leq q \text{ iff } p \not\in K \vdash (p \land q) \text{ or } p \land q \in Cn(\emptyset).$$  

\[204\] (Gärdenfors and Makinson 1988)
\[205\] (Gärdenfors and Makinson 1988, p. 89)
The condition for the other direction is

\[ q \in K \Vdash p \text{ iff } q \in K \text{ and either } p \leq p \lor q \text{ or } p \in Cn(\emptyset). \]

Having presented partial meet contraction and entrenchment-based contraction, we now look at belief revision. In belief revision, a sentence \( p \) is added to belief set \( K \) such that the newly obtained set \( K*_{\cap}p \) is consistent. The revision operator \( * \) can be defined in terms of expansion and contraction. If we first contract \( K \) by \( \neg p \), i.e. we build \( K \Vdash (\neg p) \), we can add \( p \) to \( K \Vdash (\neg p) \) without loss of consistency. This works because \( K \Vdash (\neg p) \) does not imply \( \neg p \), i.e. \( \neg p \notin Cn(K \Vdash (\neg p)) \). Thus, the revision of a belief set \( K \) by a sentence \( p \) can be defined as \( K*_{\cap}p = K \Vdash (\neg p) + p \).

If the underlying contraction operator used to define a revision operator \( * \) is a partial meet contraction operator, then \( * \) is a partial meet revision operator. Similarly, if \( \Vdash \) is an entrenchment-based contraction operator, \( * \) is an operator for entrenchment-based revision. Similar to the basic AGM postulates for partial meet contraction, Gärdenfors has provided six postulates for partial meet revision and has proved a representation theorem for them.

In the last part of the section, we look at two further belief change operators: consolidation and screened revision. Consolidation is the operation of making an inconsistent belief set (or belief base) consistent. For a given belief set \( K \) (or belief base \( A \)), the result of applying the consolidation operator \( ! \) to \( K \) (or \( A \)) is a consistent belief set \( K! \) (or consistent belief base \( A! \)). Consistency is obtained by removing elements of the respective set until it is consistent.\(^{207}\) For a belief base \( A \), consolidation can be defined by

\(^{206}\) (Gärdenfors and Makinson 1988, p. 89)

\(^{207}\) The symbol \( ! \) for consolidation was introduced by Andre Fuhrman in (Fuhrmann 1997).
contracting $A$ by the falsum $\bot$, i.e. $A! = A \vdash \bot$. $A \vdash \bot$ must be consistent as otherwise we could deduce $\bot$ from it.\(^{208}\) This definition, however, works only for belief bases. In the case of a belief set $K$, we are faced with the problem that there exists only one inconsistent belief set, the set $Cn(\bot)$. As a consequence, consolidation would have the same result for all belief sets. Hansson describes this problem as follows.

\["Once an inconsistent belief set has been obtained, all distinctions have been lost, and consolidation cannot restore them."\(^{209}\)

A possible solution to this indiscriminatory aspect of classical logic with respect to maximal inconsistent sets would be to change the underlying logic. In a paraconsistent logic, for example, it is possible to distinguish between different types of maximally inconsistent sets of formulae. Hence, in these logics there is more than one inconsistent belief set.

Screened revision is a revision operation which has the property that certain pre-defined elements of the revised belief set $K$ are immune to revision, i.e. they are not removed in the process of revision. The idea of screened revision was introduced by Makinson in order to provide a revision operator which weighs the newly received information against old information contained in the belief set with no special priority assigned to the new information for its novelty.\(^{210}\) In screened revision, a set $X$ of potential core beliefs is specified. Core beliefs are immune to revision. To perform the screened revision $K\#p$ of a belief set $K$ by a sentence $p$, we have to go through two steps. In the first step it is decided whether $K$ is to be revised by $p$ or not. The condition for revising $K$ is that $p$ is consistent with the core beliefs contained in $K$, i.e. $K$ is only revised if the set $(K \cap X) \cup $

\(^{208}\) (Hansson 1991)
\(^{209}\) (Hansson 2006, p. 17)
\(^{210}\) (Makinson 1997)
{p} is consistent. If \((K \cap X) \cup \{p\}\) is inconsistent, then \(K\#p = K\). If \((K \cap X) \cup \{p\}\) is consistent, then \(K\) is revised by \(p\) with the proviso that no elements of \((K \cap X)\) are removed from \(K\).

### 6.2 The Semantic Definition of Potential Conflict Power

Propositional formulae can be assessed with respect to their potential conflict power. The idea behind this measure is that some claims made by an agent can be harder to satisfy than other claims. Claims that are hard to satisfy should be expected to have more potential for producing a conflict when confronted with other, arbitrary, claims than claims that are easy to satisfy. If claims are expressed in terms of propositional formulae, formulae expressing claims that are hard to satisfy should, therefore, have a higher degree of potential conflict power than formulae expressing claims that are easy to satisfy.

The potential conflict power, \(\text{conf}(\varphi)\), of a propositional formula \(\varphi\) is defined as a measure between 0 and 1 expressing how likely the formula produces a conflict when it represents a claim by an agent. Tautologies, being compatible with all other formulae, and contradictions, representing conflicts in themselves, make up the two boundaries of the measure, i.e. tautologies have the value 0 and contradictions the value 1.

In the following, we provide a first definition of \(\text{conf}\) which we call the semantic definition as it is based on the semantic concept of value assignment functions. According to the semantic definition, \(\text{conf}(\varphi)\) is defined as the ratio of the value assignments restricted to the propositional constants occurring in \(\varphi\) which make \(\varphi\) false to the total number of value assignments restricted to the propositional constants occurring in \(\varphi\). Restricting the domain of value assignments to constants occurring in \(\varphi\) is
required since value assignments are generally defined as functions from the set of all propositional constants into the set of truth values \{0, 1\}. Hence, there are infinitely many value assignments. This, however, makes it impossible to come up with a finite ratio.

First, we introduce some abbreviations. If \( \varphi \) is a propositional formula, let \( \text{pro}(\varphi) \) be the set of propositional constants occurring in \( \varphi \), \( E_{\varphi} \) be the set of value assignment whose domain is restricted to \( \text{pro}(\varphi) \), i.e. \( E_{\varphi} \) is the set of all function from \( \text{pro}(\varphi) \) into \{0, 1\}, and \( \text{mod}_{\varphi}(\varphi) \) be the set of value assignment restricted to \( \text{pro}(\varphi) \) which make \( \varphi \) true, i.e. \( \text{mod}_{\varphi}(\varphi) = \{ e \in E_{\varphi} \mid v(\varphi, e) = 1 \} \), where \( v \) is the usual evaluation function of classical propositional logic.

We can now define \( \text{conf}(\varphi) \) as follows.

**Definition 95 (Semantic Definition of Potential Conflict Power)**

The potential conflict power \( \text{conf}(\varphi) \) of a propositional formula \( \varphi \) is defined as

\[
\text{conf}(\varphi) = 1 - \frac{|\text{mod}_{\varphi}(\varphi)|}{2^{|	ext{pro}(\varphi)|}}.
\]

The definition makes use of the fact that \( |E_{\varphi}| = 2^{|	ext{pro}(\varphi)|} \), i.e. there are \( 2^{|	ext{pro}(\varphi)|} \) value assignments restricted to the propositional constants occurring in \( \varphi \), and that \( 2^{|	ext{pro}(\varphi)|} - |\text{mod}_{\varphi}(\varphi)| \) value assignments make \( \varphi \) false. If \( \varphi \) is a tautology, every value assignment makes \( \varphi \) true, i.e. \( |\text{mod}_{\varphi}(\varphi)| = |E_{\varphi}| = 2^{|	ext{pro}(\varphi)|} \), and, hence, \( \text{conf}(\varphi) = 0 \). If \( \varphi \) is a contradiction, every value assignment makes \( \varphi \) false, i.e. \( |\text{mod}_{\varphi}(\varphi)| = 0 \), and, hence, \( \text{conf}(\varphi) = 1 \). Note that \( \text{conf} \) is equally suitable for measuring the information content of propositions.

To illustrate the definition, we look at the two formulae \( p \land \neg q \) and \( p \supset q \). We have

\[
|\text{pro}(p \land \neg q)| = |\text{pro}(p \supset q)| = |\{p, q\}| = 2, \quad |\text{mod}_{p \land \neg q}(p \land \neg q)| = 1, \quad \text{and} \quad |\text{mod}_{p \supset q}(p \supset q)| = 3.
\]

Hence, we can calculate the potential conflict power of the two formulae as follows.
\[ \text{conf}(p \land \neg q) = 1 - \frac{1}{2^2} = 1 - .25 = .75; \]
\[ \text{conf}(p \supset q) = 1 - \frac{3}{2^2} = 1 - .75 = .25. \]

Although \text{conf} is defined relative to the set of propositional constants occurring in \( \varphi \), it does not depend fully on \( \text{pro}(\varphi) \). Indeed, we can look at value assignment restricted to any finite set of propositional constants as long as the constants occurring in \( \varphi \) are among them. For instance, if we take \{p, q, r\} as the set of propositional constants instead of \{p, q\}, in the example above, we also get

\[ \text{conf}(p \land \neg q) = 1 - \frac{2}{2^3} = .75; \]
\[ \text{conf}(p \supset q) = 1 - \frac{6}{2^3} = .25. \]

This partial independence is important as we want to be able to compare the potential conflict power of arbitrary formulae even if they do not have common propositional constants. In order to prove the partial independence, we generalise the concept of a set \( E_\varphi \) of value assignments whose domain is restricted to the set of propositional constants occurring in a formula \( \varphi \) to sets of propositional formulae \( \Phi \). We write \( E_\Phi \) to designate the set of value assignments restricted to the constants occurring in \( \Phi \) and \( \text{mod}_\Phi(\varphi) \) to refer to the set of all value assignments restricted to \( \text{pro}(\Phi) \) which make \( \varphi \) true. The following theorem expresses the partial independence of \( \text{conf}(\varphi) \) from the set of propositional constants occurring in \( \varphi \).

**Theorem 41 (Partial Independence of conf)**
Let \( \varphi \) be a propositional formula and \( P_1 \) and \( P_2 \) any two finite sets of propositional constants s. t. \( \text{pro}(\varphi) \subseteq P_1 \) and \( \text{pro}(\varphi) \subseteq P_2 \). Then, \( 1 - |\text{mod}_{P_1}(\varphi)| / 2^{|P_1|} = 1 - |\text{mod}_{P_2}(\varphi)| / 2^{|P_2|} \).

**Proof**
Assume, wlog, that \( |P_1| \leq |P_2| \). Then, \( |P_1| = |P_2| - d \), for some \( d \in \mathbb{N}_0 \). By combinatorics we get \( |\text{mod}_{P_1}(\varphi)| = |\text{mod}_{P_2}(\varphi)| / 2^d \). This gives us \( 1 -
\[
\frac{|\text{mod}_{P_1}(\phi)|}{2^{|P_1|}} = 1 - \left(\frac{|\text{mod}_{P_2}(\phi)|}{2^{|P_2|}}\right) = 1 - \frac{1}{2^{|P_2|}}\]

QED

6.3 The Syntactic Definition of Potential Conflict Power

In the semantic definition of \(\text{conf}\), the idea was to weight the number of situations in which a formula is unsatisfied against the total number of possible situations. The harder it is to satisfy a formula, the higher its degree of potential conflict power.

Now we present a syntactic approach to the potential conflict power of formulae which makes use of the syntactic notion of inconsistency. Here, the underlying idea is to compare a formula with relevant other formulae, and then count the number of formulae with which the formula is inconsistent. The more formulae a formula is inconsistent with, the higher its degree of potential conflict power should be. For a given formula \(\phi\), the main task will be to find an appropriate selection of formulae that we compare \(\phi\) with.

As there are infinitely many potential propositional formulae a particular formula can potentially be compared with, we have to limit the number of formulae that we look at. This is done in two steps. Starting with a formula \(\psi\) for which we want to compute \(\text{conf}(\psi)\), we first single out a set of formulae \(\text{FOR}_\psi\) containing all formulae that can be built using only propositional constants occurring in \(\psi\). As \(\text{FOR}_\psi\) is still an infinite set, we then ‘finitise’ \(\text{FOR}_\psi\) by factorising it modulo the syntactic equivalence relation \(\vdash\). This leaves us with a finite set of representants of the quotient set \(\text{FOR}_\psi \vdash\). The definition of \(\text{FOR}_\psi\) is given as follows.
Definition 96 (FOR\_φ)
Let \( \varphi \) be a propositional formula and \( \text{pro}(\varphi) \) be the set of propositional constants occurring in \( \varphi \), then \( \text{FOR}_\varphi \) is defined as the set \( \text{FOR}_\varphi = \{ \varphi | \text{pro}(\varphi) \subseteq \text{pro}(\varphi) \} \).

As the definition shows, \( \text{FOR}_\varphi \) is the set of propositional formulae that can be built using only propositional constants occurring in \( \varphi \). \( \text{FOR}_\varphi \) represents the set of all possible claims that can be made with respect to what is expressed by the basic propositions contained in \( \varphi \). If each element of \( \text{FOR}_\varphi \) represents, for instance, the content of a belief, \( \text{FOR}_\varphi \) represents the set of all possible beliefs concerning the basic propositions contained in \( \varphi \).

As the formation rules of propositional formulae allow infinite iterations, \( \text{FOR}_\varphi \) is an infinite set. To reduce \( \text{FOR}_\varphi \) to a finite set, thereby making it possible to compare \( \varphi \) to a finite number of formulae, we use the fact that in classical logic any two formulae \( \varphi \) and \( \psi \) that are provably equivalent, i.e. for which both \( \varphi \vdash \psi \) and \( \psi \vdash \varphi \) holds, express the same content. As usual, we write \( \varphi \vdash \psi \) as an abbreviation for \( \varphi \vdash \psi \) and \( \psi \vdash \varphi \). \( \vdash \) constitutes an equivalence relation on \( \text{FOR}_\varphi \), and, hence, we can factorise \( \text{FOR}_\varphi \) modulo \( \vdash \). The quotient set \( \text{FOR}_\varphi \backslash \vdash \) which is usually called the Lindenbaum algebra consists of \( 2^{|\text{pro}(\varphi)|} \) equivalence classes each being a set of provably equivalent formulae built from propositional constants occurring in \( \varphi \). By choosing one element from each equivalence class contained in \( \text{FOR}_\varphi \backslash \vdash \), we can define a finite set \( \text{REL}_\varphi \) of representants of \( \text{FOR}_\varphi \backslash \vdash \) as follows.

\[ \text{Definition 97 (REL} \_\varphi) \]
Let \( \text{FOR}_\varphi \backslash \vdash \) be the quotient set of \( \text{FOR}_\varphi \) modulo the syntactic equivalence relation \( \vdash \). Then, \( \text{REL}_\varphi \) is defined by \( \text{REL}_\varphi = \{ [\varphi] | [\varphi] \in \text{FOR}_\varphi \backslash \vdash \text{ and } [\varphi] \neq [\chi] \text{ for all } \chi \in \text{REL}_\varphi \} \).

With \( \text{REL}_\varphi \) we have a finite set of formulae containing only formulae that are relevant to \( \varphi \). Each cell of \( \text{FOR}_\varphi \backslash \vdash \) is represented in \( \text{REL}_\varphi \) by exactly one member. As \( \text{FOR}_\varphi \backslash \vdash \) is finite, \( \text{REL}_\varphi \) is finite too.
Before we provide the formal syntactic definition of $\text{conf}$ we still have to introduce the set $\text{INC}_\varphi$ of all formulae of $\text{REL}_\varphi$ with which $\varphi$ is inconsistent.

**Definition 98 (INC$_\varphi$)**

Let $\text{REL}_\varphi$ be the set of formulae relevant to $\varphi$ as defined in Definition 97. Then, $\text{INC}_\varphi$ is defined by $\text{INC}_\varphi(\varphi) = \{ \psi \in \text{REL}_\varphi \mid \varphi \land \psi \not\vdash \bot \}$.

$\text{INC}_\varphi$, which is a subset of $\text{REL}_\varphi$, can be obtained from $\varphi$ by consecutively comparing $\varphi$ with each element of $\text{REL}_\varphi$. In each case, we can decide whether or not $\varphi$ is consistent with the formula we are comparing $\varphi$ with. If the formula is inconsistent with $\varphi$, we add it to $\text{INC}_\varphi$.

Now we can syntactically define the potential conflict power $\text{conf}(\varphi)$ of a propositional formula $\varphi$ as follows.

**Definition 99 (Syntactic Definition of Potential Conflict Power)**

The potential conflict power $\text{conf}(\varphi)$ of a propositional formula $\varphi$ is syntactically defined as $\text{conf}(\varphi) = \log_2 |\text{INC}_\varphi(\varphi)|$.

The syntactical definition of $\text{conf}(\varphi)$ is based on the number $|\text{INC}_\varphi(\varphi)|$ of relevant formulae with which $\varphi$ is inconsistent. The logarithm function of base $2^{\frac{1}{|\text{pro}(\varphi)|}}$ is applied to $|\text{INC}_\varphi(\varphi)|$ in order to standardise the measure. This makes sure that $\text{conf}(\varphi)$ is always a number between 0 and 1. Furthermore, if $\varphi$ is a contradiction, then $|\text{INC}_\varphi(\varphi)| = 2^{\frac{1}{|\text{pro}(\varphi)|}}$ as $\varphi$ is inconsistent with every element of $\text{REL}_\varphi$. Hence, we get $\text{conf}(\varphi) = 1$. If $\varphi$ is a tautology, then $|\text{INC}_\varphi(\varphi)| = 1$ as $\varphi$ is only inconsistent with $\bot$. This results in $\text{conf}(\varphi) = 0$. As the logarithm function is monotonically increasing, the more formulae there are with which $\varphi$ is inconsistent, the higher $\text{conf}(\varphi)$. The definition reflects the intuitive assumption that the potential conflict power of a statement should be proportional to the number of statements with which it would potentially constitute a conflict.
Finally, we prove that the two definitions of potential conflict power, i.e. the semantic definition \( \text{conf}(\varphi) = 1 - \frac{|\text{mod}_\varphi(\varphi)|}{2^{|\text{pro}(\varphi)|}} \) and the syntactic definition \( \text{conf}(\varphi) = \log_2 2^{\frac{|\text{pro}(\varphi)|}{|\text{inc}_\varphi(\varphi)|}} \), indeed, pick the same number.

**Theorem 42**  (Semantic and Syntactic Conflict Power)

\[
1 - \frac{|\text{mod}_\varphi(\varphi)|}{2^{|\text{pro}(\varphi)|}} = \log_2 2^{\frac{|\text{pro}(\varphi)|}{|\text{inc}_\varphi(\varphi)|}} \quad \text{for every propositional formulae } \varphi.
\]

**Proof**

Let \( \varphi \) be a propositional formula, \( \text{pro}(\varphi) \) be the set of constants occurring in \( \varphi \), and \( \text{mod}_\varphi(\varphi) \) be the set of value assignment functions from \( \text{pro}(\varphi) \) that make \( \varphi \) true. Then, there are \( 2^{|\text{pro}(\varphi)|} - |\text{mod}_\varphi(\varphi)| \) value assignments from \( \text{pro}(\varphi) \) making \( \varphi \) false. By combinatorics, there are \( 2^{|\text{pro}(\varphi)| - |\text{mod}_\varphi(\varphi)|} \) formulae, up to semantic equivalence, built from \( \text{pro}(\varphi) \), which are unsatisfiable in conjunction with \( \varphi \). This is because every formula that is false under every value assignment under which \( \varphi \) is true, is unsatisfiable together with \( \varphi \). Holding \( |\text{mod}_\varphi(\varphi)| \), the number of value assignments making \( \varphi \) true, fixed, there are still \( 2^{|\text{pro}(\varphi)| - |\text{mod}_\varphi(\varphi)|} \) permutations of value assignments. Each of them represents one equivalence class of semantically equivalent formulae.

By completeness of propositional logic, there are \( 2^{2^{|\text{pro}(\varphi)| - |\text{mod}_\varphi(\varphi)|}} \) formulae, up to syntactic equivalence, built from \( \text{pro}(\varphi) \), which are inconsistent with \( \varphi \). Hence, \( |\text{INC}_\varphi(\varphi)| = 2^{|\text{pro}(\varphi)| - |\text{mod}_\varphi(\varphi)|} \). Substituting \( |\text{INC}_\varphi(\varphi)| \) in the syntactic definition of \( \text{conf}(\varphi) \), along with some stipulations on the logarithm function, gives us the following identities:

\[
\begin{align*}
\log_2 2^{\frac{|\text{pro}(\varphi)|}{|\text{inc}_\varphi(\varphi)|}} \cdot |\text{INC}_\varphi(\varphi)| \\
= \log_2 2^{\frac{|\text{pro}(\varphi)|}{2^{|\text{pro}(\varphi)| - |\text{mod}_\varphi(\varphi)|}}} \\
= \log_2 2^{\frac{|\text{pro}(\varphi)|}{2^{|\text{pro}(\varphi)| - |\text{mod}_\varphi(\varphi)|}}} \\
= \log_2 2^{\frac{|\text{pro}(\varphi)|}{2^{|\text{pro}(\varphi)|}}} - \log_2 2^{\frac{|\text{pro}(\varphi)|}{2^{|\text{mod}_\varphi(\varphi)|}}} \\
= 1 - \frac{|\text{mod}_\varphi(\varphi)|}{2^{|\text{pro}(\varphi)|}} \\
= 1 - \frac{|\text{mod}_\varphi(\varphi)|}{2^{|\text{pro}(\varphi)|}}.
\end{align*}
\]

QED

### 6.4 Some Properties of Potential Conflict Power

The potential conflict power, defined semantically or syntactically, is a function from the set of all propositional formulae \( \text{FOR} \) into the rational interval \( [0, 1] \subset \mathbb{Q} \). It is neither an injective function, as different formulae can have the same potential conflict power, nor
is it a surjective function since the image of conf contains only fractions of the form \(m/2^n\) where \(m, n \in \mathbb{N}\) and \(m < 2^n\), and not every element of the rational interval \([0, 1]\) can be expressed as such a fraction.

Assigning numbers to formulae, conf induces a linear ordering on FOR. Any two formulae \(\varphi\) and \(\psi\) can be compared with each other with respect to their potential conflict power, and one of three cases must hold: \(\text{conf}(\varphi) < \text{conf}(\psi)\), \(\text{conf}(\varphi) = \text{conf}(\psi)\), or \(\text{conf}(\psi) < \text{conf}(\varphi)\). Contradictions, having the highest degree of potential conflict power, are at the top end of the induced linear order. Tautologies, having no potential conflict power at all, are at the bottom end. Singular claims, expressed by singular propositional constants \(p, q, r, \text{etc.}\), are in the middle of the order as \(\text{conf}(p) = .5\) for all propositional constants \(p\).

In the following section, we look at some properties of conf. In particular, we are interested in its behaviour regarding the consequence relation \(\vdash\) and the propositional connectives \(\neg, \vee, \text{and } \wedge\).

Our first theorem expresses the relationship between conf and the consequence relation \(\vdash\).

**Theorem 43  (Conf and \(\vdash\))**

If \(\varphi \vdash \psi\), then \(\text{conf}(\psi) \leq \text{conf}(\varphi)\) for all propositional formulae \(\varphi, \psi\).

**Proof**

Let \(\varphi, \psi\) be two propositional formulae such that \(\varphi \vdash \psi\). Then, \(\text{mod}_{\varphi}(\varphi) \subseteq \text{mod}_{\varphi}(\psi)\) by completeness and the definition of \(\vdash\). Then, \(|\text{mod}_{\varphi}(\varphi)| \leq |\text{mod}_{\varphi}(\psi)|\) and, hence, \(\text{conf}(\psi) = 1 - |\text{mod}_{\varphi}(\psi)|/2^{|\text{var}(\varphi, \psi)|} \leq 1 - |\text{mod}_{\varphi}(\varphi)|/2^{|\text{var}(\varphi, \psi)|} = \text{conf}(\varphi)\).

**QED**
The above theorem can be interpreted as follows: If a propositional formula entails another formula, its potential conflict power is at least as high as the conflict power of the formula entailed by it. The theorem reflects the intuition that if a formula $\varphi$ entails another formula $\psi$, $\varphi$ contains more information than $\psi$ because $\varphi$ makes a more specific claim about the world than $\psi$. As more specific claims are harder to satisfy than less specific claims, their potential to produce a conflict is higher compared to less specific claim.

The theorem has some simple but interesting consequences. First, equivalent formulae have the same degree of potential conflict power because $\varphi \vdash \psi$ and $\psi \vdash \varphi$ implies $\text{conf}(\varphi) \leq \text{conf}(\psi)$ and $\text{conf}(\psi) \leq \text{conf}(\varphi)$, which amounts to $\text{conf}(\varphi) = \text{conf}(\psi)$. Second, the theorem reiterates the role of tautologies and contradictions in the linear ordering induced by conf. As contradictions entail any other formula $\varphi$, they can be characterised by the property of having an equal or higher conflict power than every other formula $\varphi$, i.e. $\text{conf}(\varphi) \leq \bot$ for all $\varphi$. In contrast, tautologies are entailed by any other formula $\varphi$. Hence, every other formula has an equal or higher potential conflict power than a tautology, i.e. $\top \leq \text{conf}(\varphi)$ for all $\varphi$. Third, the conflict power of a formula $\varphi$ is always at least as high as the conflict power of a disjunction of $\varphi$ with an arbitrary other formula $\psi$, i.e. $\text{conf}(\varphi \lor \psi) \leq \text{conf}(\varphi)$. Replacing a formula $\varphi$ by a disjunction $\varphi \lor \psi$ is, therefore, a way to reduce the potential conflict power of a formula. On the other hand, replacing a formula $\varphi$ by a conjunction $\varphi \land \psi$, where $\psi$ is an arbitrary formula, increases the potential conflict power, i.e. $\text{conf}(\varphi) \leq \text{conf}(\varphi \land \psi)$. This is because $\varphi \vdash \varphi \lor \psi$ and $\varphi \land \psi \vdash \varphi$.

As expressed by the following two theorems, the relationship between conf and the logical connectives can be characterised in a more specific way. For the negation symbol we can prove the following theorem.
Chapter 6  Measuring Conflicts

Theorem 44  \( \text{(Conf and } \neg \text{)} \)
conf\( (\neg \varphi) \) = 1 \( - \) conf\( (\varphi) \) for all propositional formulae \( \varphi \).

Proof
Let \( \varphi \) be a propositional formula. By the bivalence principle of propositional logic, we have \( |\text{mod}_{\text{pro}}(\neg \varphi)| = 2^{2^{\text{pro}(\varphi)}} \). Hence, we get:
\[
\text{conf}(\neg \varphi) = 1 - |\text{mod}_{\text{pro}}(\neg \varphi)|/2^{\text{pro}(\varphi)} = 1 - 2^{\text{pro}(\varphi)} = 1 - (1 - |\text{mod}_{\text{pro}}(\varphi)|/2^{\text{pro}(\varphi)}) = 1 - \text{conf}(\varphi).
\]
QED

The next theorem addresses the relationship between conf and the connectives \( \land \) and \( \lor \).

Theorem 45  \( \text{(Conf and } \land \text{ and } \lor \text{)} \)
conf\( (\varphi \land \psi) \) = conf\( (\varphi) \) + conf\( (\psi) \) \( - \) conf\( (\varphi \lor \psi) \) for all propositional formulae \( \varphi, \psi \).

Proof
Let \( \varphi, \psi \) be two propositional formulae. By definition of the truth conditions for \( \land \) and \( \lor \), we have the following identities:
\[
\text{mod}_{\text{pro}}(\varphi \land \psi) = \text{mod}_{\text{pro}}(\varphi) \cap \text{mod}_{\text{pro}}(\psi) \quad \text{and} \quad \text{mod}_{\text{pro}}(\varphi \lor \psi) = \text{mod}_{\text{pro}}(\varphi) \cup \text{mod}_{\text{pro}}(\psi).
\]
By set theory, \( |\text{mod}\{\varphi, \psi\}(\varphi \land \psi)| = |\text{mod}\{\varphi, \psi\}(\varphi)| + |\text{mod}\{\varphi, \psi\}(\psi)| - |\text{mod}\{\varphi, \psi\}(\varphi) \cup \text{mod}\{\varphi, \psi\}(\psi)|. \)
Consider now the following identity:
\[
\text{conf}(\varphi \land \psi) = 1 - |\text{mod}\{\varphi, \psi\}(\varphi \land \psi)|/2^{\text{var}\{\varphi, \psi\}} = 1 - |\text{mod}\{\varphi, \psi\}(\varphi)| + |\text{mod}\{\varphi, \psi\}(\psi)| - |\text{mod}\{\varphi, \psi\}(\varphi) \cup \text{mod}\{\varphi, \psi\}(\psi)|/2^{\text{var}\{\varphi, \psi\}}
\]
\[
= 1 - |\text{mod}\{\varphi, \psi\}(\varphi)|/2^{\text{var}\{\varphi, \psi\}} + 1 - |\text{mod}\{\varphi, \psi\}(\psi)|/2^{\text{var}\{\varphi, \psi\}} - (1 - |\text{mod}\{\varphi, \psi\}(\varphi) \cup \text{mod}\{\varphi, \psi\}(\psi)|/2^{\text{var}\{\varphi, \psi\}})
\]
\[
= \text{conf}(\varphi) + \text{conf}(\psi) \quad - \quad \text{conf}(\varphi \lor \psi).
\]
QED

The theorem allows us to calculate the conflict power of a conjunction if the individual conflict powers of its components and the conflict power of their disjunction are known.

Its proof relies on the fact that the set of models of a conjunction is the set theoretic intersection of the set of models of its constituents, whereas the set of models of a disjunction is the union of the sets of models of its constituents.

The theorem has two interesting consequences as it provides upper bounds for the conflict power of disjunctions and conjunctions. Both the conflict power of a disjunction
and the conflict power of a conjunction are at most as high as the sum of the conflict powers of their respective constituents, i.e. \( \text{conf}(\varphi \lor \psi) \leq \text{conf}(\varphi) + \text{conf}(\psi) \) and \( \text{conf}(\varphi \land \psi) \leq \text{conf}(\varphi) + \text{conf}(\psi) \) for all propositional formulae \( \varphi, \psi \).

Finally, we compare the concept of potential conflict power with Gärdenfors’s notion of epistemic entrenchment. As introduced in the background section on belief revision, Gärdenfors’s concept is intended to express how deeply sentences of a propositional language are epistemically entrenched in a given knowledge set. He explains this notion as follows.

“Even if all sentences in a knowledge set are accepted or considered as facts […] this does not mean that all sentences are of equal value for planning or problem-solving purposes. Certain pieces of our knowledge and beliefs about the world are more important than others when planning future actions, conducting scientific investigations, or reasoning in general. We will say that some sentences in a knowledge system have a higher degree of epistemic entrenchment than others.”

One might think that the ordering relation induced by \( \text{conf} \) also satisfies Gärdenfors’s postulates for epistemic entrenchment\(^{212}\) and, therefore, represents an entrenchment relation. The two relations have in common that they are both intended to reflect the informational value of propositional formulae. Gärdenfors considers sentences with a high amount of information more valuable for an epistemic agent than sentences containing less information. Similarly, propositional formulae that contain a high amount of information are harder to satisfy than formulae with a low degree of informational value and have, therefore, a higher potential conflict power than the latter ones.

Although there are similarities between the two concepts, there are also fundamental differences. Whereas entrenchment is a qualitative measure of how hard it is to give up a belief, \( \text{conf} \) is rather a measure of its information content. In fact, if we define a binary

\(^{211}\) (Gärdenfors and Makinson 1988)

\(^{212}\) See the background section on belief revision.
relation $<$ on FOR by $\phi \prec \psi$ iff $\text{conf}(\phi) < \text{conf}(\psi)$, then $<$ satisfies Gärdenfors’s postulates of transitivity and conjunctiveness, i.e. $\phi \prec \psi$ and $\psi \prec \chi$ implies $\phi \prec \chi$ and $\phi \preceq \psi$ or $\psi \preceq \phi \wedge \psi$\textsuperscript{213}. In the case of Gärdenfors’s dominance postulate, which states that if $p \vdash q$, then $q$ is epistemically more entrenched than $p$, the situation is inverted. To understand the difference between the two relations we have to look at Gärdenfors’s justification for his dominance postulate.

”[I]f $A$ entails $B$ and either $A$ or $B$ must be retracted from the belief set $K$, then it is a smaller change to give up $A$ and retain $B$ rather than to give up $B$, because then $A$ must also be retracted if we want the revised theory to be closed under logical consequences.”\textsuperscript{214}

For Gärdenfors, the question is how to minimise the overall informational loss when giving up certain elements of a belief set. This can lead to situations in which a sentence with more informational value than another sentence is given up in order to minimise the overall informational loss. In contrast, conf directly reflects the informational value of propositional formulae.

Also, Gärdenfors’s maximality postulate is satisfied by $\preceq$ in the opposite direction. Whereas for Gärdenfors, tautologies are maximally epistemically entrenched, i.e. all other formulae are less entrenched than them, conf assigns the minimal value of 0 to them. The difference is that Gärdenfors wants tautologies to be the last sentences to be given up if a belief set is revised; whereas in our framework they represent the least conflicting statements.

Gärdenfors’s minimality postulate explicitly refers to the notion of knowledge sets, a concept which is not directly applicable to the conflict power function conf.

\textsuperscript{213} $\preceq$ even satisfies $\phi \preceq \psi \wedge \phi$ and $\psi \preceq \phi \wedge \psi$, for all $\phi, \psi$.

\textsuperscript{214} (Gärdenfors 1988, p. 89)
6.5 Degrees of Inconsistency

With conf, we can assess the potential conflict power of propositional formulae, and, hence, the role they play within a conflict. However, the function does not provide a method for assessing conflicts, understood as inconsistent sets of formulae, as a whole. In fact, conf can only be applied to individual formulae and not to sets of formulae.

Generalising conf to sets of formulae, e.g. by defining conf(Φ) as the conflict power of the conjunction of all formulae contained in Φ, does not provide a solution as every inconsistent set would immediately have the potential conflict value 1.

The underlying idea of the concept of degrees of inconsistency is the assumption that not all conflicts are equally difficult to resolve. Some conflicts seem to have a more complex structure than others. In order to solve them, it is not sufficient to change just one of their parts. Instead, their overall structure is interlinked so deeply that resolving them requires changing a significant number of the components that constitute them.

To motivate the definition of the degree of inconsistency, we look at three inconsistent sets of propositional formulae.

\[ Φ_1 = \{p \lor \neg q, q, p \equiv q, \neg p \land q\}; \]
\[ Φ_2 = \{p \land q, \neg p \land q, p \equiv q, p \land \neg q\}; \]
\[ Φ_3 = \{p \land q, \neg p \land q, \neg p \land \neg q, p \land \neg q\}. \]

Each of them can be understood as a conflict among four agents. Each of the four formulae contained in a set represents one agent’s claim. To assess how deep the conflicts are, we can ask the question: what is the minimum number of formulae that need to be changed in order to make the set consistent? By changing a formula we mean
to replace it by a formula with a lower degree of potential conflict power that is as close as possible to the original formula.

In the case of $\Phi_1$, we can make the set consistent by replacing just one formula: $\neg p \land q$. We replace this formula by the weaker formula $q$. The new set $\Phi_1^* = \{p \lor \neg q, q, p \equiv q, q\}$ is consistent, $conf(q)$ is smaller than $conf(\neg p \land q)$ and $q$ is as close as possible at $\neg p \land q$ as it is logically entailed by $\neg p \land q$.

In the case of $\Phi_2$, it is not sufficient to replace just one formula to make the set consistent because the remaining three formulae will still be inconsistent with each other. However, there is a way to make $\Phi_2$ consistent by changing the two formulae $\neg p \land q$ and $p \land \neg q$. Replacing the former by $q$ and the latter by $p$, we obtain the consistent set $\Phi_2^* = \{p \land q, q, p \equiv q, p\}$. Again, the new formulae are entailed by the original ones and have a lower degree of potential conflict power.

$\Phi_3$ can only be made consistent by replacing three formulae simultaneously. Replacing $p \land q$ by $q$, $\neg p \land \neg q$ by $\neg p$ and $p \land \neg q$ by $p \equiv \neg q$ produces the required result. Weakening just one or two elements of $\Phi_3$ is not enough as the remaining formulae are still inconsistent with each other.

The process of replacing formulae by formulae with less potential conflict power corresponds to the process of weakening the agents’ claims until a solution to their conflict is found. Altogether, the example shows that not every inconsistent set of formulae can be made consistent equally easily. Taking the degree of change required for making an inconsistent set consistent as the degree of its inconsistency, we can conclude that $\Phi_1$ should have a lower degree of inconsistency than $\Phi_2$, and $\Phi_2$ should have a lower degree of inconsistency than $\Phi_3$.
In the following section we propose three different functions, inc-1, inc-2, and inc-3, to measure the degree of inconsistency of a set $\Phi$ of propositional formulae. For each of the three functions, we present an absolute and a relative version, depending on whether the size of $\Phi$ itself is taken into account. The relative version is indicated by an asterisk. After having introduced and discussed the three measures, we will argue for inc-1*, the relative version of the first proposal, as the most plausible function for expressing the degree of inconsistency of a set.

As in earlier sections, we denote the set of all value assignments from a fixed set of propositional constants into the set $\{0, 1\}$ by $E_\Phi$, i.e. if $\text{pro}(\Phi)$ is the set of propositional constants occurring in $\Phi$, then $E_\Phi = 2^{\text{pro}(\Phi)}$. $\text{mod}_\phi(\varphi)$ is the set of all truth value assignments restricted to $\text{pro}(\Phi)$ which make $\varphi$ true, i.e. $\text{mod}_\phi(\varphi) = \{e \in E_\Phi \mid v(\varphi, e) = 1\}$ where $\text{pro}(\varphi) \subseteq \text{pro}(\Phi)$.

Now we define the absolute density and relative density of a value assignment $e$ as follows.

**Definition 100 (Density of a Value Assignment)**
Let $\Phi$ be a set of propositional formulae and $e \in E_\Phi$ a value assignment restricted to $\text{pro}(\Phi)$, then:
- $d_\Phi(e) = |\{\varphi \in \Phi \mid e \notin \text{mod}_\phi(\varphi)\}|$ is the absolute density of $e$ with respect to $\Phi$;
- $d_\Phi^*(e) = |\{\varphi \in \Phi \mid e \notin \text{mod}_\phi(\varphi)\}| / |\Phi|$ is the relative density of $e$ with respect to $\Phi$.

$d_\Phi(e)$ is a function from the set of all value assignments $E_\Phi$ into $\mathbb{N}$. $d_\Phi^*(e)$ is a function from $E_\Phi$ into the rational interval $[0, 1] \subseteq \mathbb{Q}$. The density of a value assignment $e$ with respect to a set $\Phi$ expresses how many elements of $\Phi$ are false under $e$. In the absolute case $d_\Phi(e)$ reflects just the number of such formulae, whereas in the relative case $d_\Phi^*(e)$ reflects the proportion of such formulae to the total number of formulae contained in $\Phi$.

We have chosen the name ‘density’ for this function as it can graphically be illustrated as
the density of the layers of 0s generated by the elements of a set $\Phi$ of propositional formulae in the truth table for the formulae.

The following table illustrates the absolute and relative densities of the set $\Phi = \{p, p \land q, q, p \lor q\}$.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>p \land q</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>q</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>p \lor \neg q</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$d_\Phi(e_1)$</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$d_\Phi^*(e_1)$</td>
<td>0</td>
<td>.5</td>
<td>.75</td>
<td>.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E_\Phi$</th>
<th>$e_1$</th>
<th>$e_2$</th>
<th>$e_3$</th>
<th>$e_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>p=1 q=1</td>
<td>p=0 q=1</td>
<td>p=1 q=0</td>
<td>p=0 q=0</td>
<td></td>
</tr>
</tbody>
</table>

**Table 12: Absolute and Relative Density**

The table shows that $\Phi$ is consistent as there is a truth value assignment, $e_1$, such that $d_\Phi(e_1) = 0$.

Using the density function, we can now define three pairs of functions measuring the degree of inconsistency of a set $\Phi$ of propositional formulae: inc-1, inc-2, and inc-3.

**Definition 101 (Degree of Inconsistency)**
The degree of inconsistency of a set of propositional formulae $\Phi$ can be measured by the following three pairs of functions:

1. $\text{inc-1}(\Phi) = \min\{d_\Phi(e) \mid e \in E_\Phi\}$;
2. $\text{inc-1}^*(\Phi) = \min\{d_\Phi^*(e) \mid e \in E_\Phi\}$;
3. $\text{inc-2}(\Phi) = \sum_{e \in E_\Phi} d_\Phi(e) / 2^{\pro(\Phi)}$;
4. $\text{inc-2}^*(\Phi) = \sum_{e \in E_\Phi} d_\Phi^*(e) / 2^{\pro(\Phi)}$;
5. $\text{inc-3}(\Phi) = \prod_{e \in E_\Phi} d_\Phi(e)$;
6. $\text{inc-3}^*(\Phi) = \prod_{e \in E_\Phi} d_\Phi^*(e)$.
The first pair of functions, inc-1(Φ) and inc-1*(Φ), defines the degree of inconsistency of a set Φ of propositional formulae as the minimal absolute density of the set in the case of inc-1, and as the minimal relative density in the case of inc-1*. This definition directly reflects the consideration that the degree of inconsistency of a set is measured by the degree of difficulty to make the set consistent. inc-1(Φ) reflects the number of elements of Φ that need to be changed in order to make Φ consistent, whereas inc-1*(Φ) expresses the proportion of elements of Φ that need to be changed. In our previous example of the three sets Φ₁, Φ₂, and Φ₃, we get

\[
\begin{align*}
\text{inc-1}(\Phi_1) &= 1, \quad \text{inc-1}*(\Phi_1) = .25; \\
\text{inc-1}(\Phi_2) &= 2, \quad \text{inc-1}*(\Phi_2) = .5; \\
\text{inc-1}(\Phi_3) &= 3, \quad \text{inc-1}*(\Phi_3) = .75.
\end{align*}
\]

Using the minimal density as a measure of the degree of inconsistency, every consistent set has the value 0. Furthermore, inc-1*, which assigns only values between 0 and 1, assigns the value 1 to sets containing only contradictions. These extreme values reflect the intuition that consistent sets have a minimal degree of inconsistency, whereas the maximal value is only assigned to sets in which every element needs to be changed to make the set consistent.

The second proposal takes inc-2 to be the average absolute density of a set Φ and inc-2* as the average relative density. In our example we get

\[
\begin{align*}
\text{inc-2}(\Phi_1) &= 2, \quad \text{inc-2}*(\Phi_1) = .5; \\
\text{inc-2}(\Phi_2) &= 2.75, \quad \text{inc-2}*(\Phi_2) = .6875; \\
\text{inc-2}(\Phi_3) &= 3, \quad \text{inc-1}*(\Phi_3) = .75.
\end{align*}
\]
Comparing inc-1 with inc-2, we observe that the intuitively plausible order for the three sets is reflected by both functions. However, the distance between the respective values is much narrower in the case of inc-2 as opposed to inc-1.

The rationale behind inc-2 is to take into account not only the minimal density of a set, i.e. the minimal number of formulae that need to be changed to make the set consistent, but to look at every possible way to make the set consistent and then average the amounts of changes needed in each case. It has the advantage over inc-1 that it reflects in more detail the truth table for the set in question. However, it has the disadvantage that consistent sets can have a positive degree of inconsistency. Even more counterintuitively, a consistent set can have a higher degree of inconsistency than an inconsistent set as illustrated by the two sets $\Lambda = \{p \land q, q\}$ and $B = \{p, \neg p\}$. Here, we get $inc-2^*(\Lambda) = .625$ and $inc-2^*(B) = .5$, although $\Lambda$ is consistent, whereas $B$ is not.

The third proposal is intended to repair the defect of the second one. It still takes into account all value assignments relevant to the set $\Phi$, but does not have the property of assigning a degree of inconsistency greater than 0 to consistent sets. It defines inc-3 as the product of all absolute densities and inc-3* as the product of all relative densities. If a set is consistent, one of its densities is equal to 0, and, hence, the product of its densities automatically also equals 0. Looking at our example we get the following degrees of inconsistency.

\[
\begin{align*}
inc-3(\Phi_1) &= 12, \quad inc-3^*(\Phi_1) = .046875; \\
inc-3(\Phi_2) &= 54, \quad inc-3^*(\Phi_2) = .2109375; \\
inc-3(\Phi_3) &= 81, \quad inc-1^*(\Phi_3) = .31640625.
\end{align*}
\]

Again, inc-3 reflects the natural order of the three sets regarding their degree of inconsistency. However, the absolute version inc-3 assigns very large numbers to the
sets, whereas the relative version inc-3* does the opposite and assigns very small numbers to the three sets. In general, it remains unclear how the degrees of inconsistency assigned to sets by inc-3 can be interpreted in a natural way.

In conclusion, inc-1* seems to be the most plausible choice for expressing the degree of inconsistency of sets of propositional formulae. The reasons are as follows. First, inc-1* reflects the intuitive assumption that the degree of inconsistency of a set \( \Phi \) should be expressed by the complexity of the process of making \( \Phi \) consistent. Second, inc-1* assigns the minimal value 0 to consistent sets and the maximal value 1 to sets containing only contradictions. Third, inc-1* is standardised, i.e. its values are elements of the interval \([0, 1]\). Fourth, it seems justified to look at just the minimal density of \( \Phi \) as all other densities are unnecessary when it comes to making \( \Phi \) consistent.

**Example: Second Congo War**

In this section we apply the concept of potential conflict power and degree of inconsistency, as defined above, to our example of the Second Congo War. In the case of the definition of potential conflict power, we look at an example of a goal conflict within the Second Congo War which we have already examined. Suppose the agent RCD has the goal to become part of the national Congolese army without replacing its high ranking officers, and the Kabila government wants the RCD only to become part of the Congolese national army if it replaces its high ranking officers. Then, the former claim by the RCD is harder to satisfy than the latter claim by the Kabila government. This is evident because there are more scenarios in which the RCD’s claim is unsatisfied than there are scenarios in which the Kabila government’s claim is unsatisfied. We now justify this difference of the two goals by computing and comparing their potential conflict...
powers. If the RCD’s claim is, indeed, harder to satisfy than the claim made by the Kabila government, then the potential conflict power of the RCD’s claim should be higher than the potential conflict power of the claim by the Kabila government.

Using the language of CML, the two claims can be symbolised by the formulae $Ga_1(p \land \neg q)$ and $Ga_2(p \supset q)$, respectively. Hence, the goal conflict $C_g$ is given as follows.

\[
a_1: \text{RCD} \\
a_2: \text{Kabila government} \\
p: \text{The RCD is part of the national Congolese army} \\
q: \text{The RCD replaces its high ranking officers} \\
C: \{Ga_1(p \land \neg q), Ga_2(p \supset q)\} \\
C_g: \{p \land \neg q, p \supset q\}
\]

We calculate the potential conflict power of the two formulae via the semantic route, i.e. by looking at their truth table and taking the ratio of the situations in which they are satisfied and the number of all possible situations. The truth table for the formulae is shown below.

<table>
<thead>
<tr>
<th></th>
<th>$p \land \neg q$</th>
<th>$p \supset q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p=1$, $q=1$</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$p=1$, $q=0$</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$p=0$, $q=1$</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$p=0$, $q=0$</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 13: Truth Table for $C_g$**

The table shows that $p \land \neg q$ has one model, and $p \supset q$ has three models. Hence, we can compute $\text{conf}(p \land \neg q)$ and $\text{conf}(p \supset q)$ as follows.

\[
\text{conf}(p \land \neg q) = 1 - 1/2^2 = 1 - 1/4 = .75 \\
\text{conf}(p \supset q) = 1 - 3/2^3 = 1 - 3/4 = .25
\]
Altogether, we can conclude that the RCD’s claim of being part of the national Congolese army without replacing their high ranking officers has more potential conflict power than the conditional claim of the Kabila government according to which the RCD can only become part of the national Congolese army if they replace their high ranking officers.

Next, we illustrate the concept of degrees of inconsistency in the context of the Second Congo War. Again, we look at two sub-conflicts of the Second Congo War which have been used for the purpose of illustration before. We calculate the degree of inconsistency for the two conflicts and compare them with each other in order to judge which of the two conflicts is more difficult to solve.

The first conflict is, again, the goal conflict between the RCD and the Kabila government regarding the integration of combatants of the RCD into the national Congolese army.

\[
a_1: \text{Rassemblement Congolais pour la Democratie (RCD)} \\
a_2: \text{Kabila government} \\
p: \text{The combatants of the RCD are integrated into the Congolese national army.} \\
q: \text{High ranking officers of the RCD are replaced.} \\
C = \{Ga_1 p, Ga_1 \neg q, Ga_2 p, Ga_2 (p \supset q)\} \\
C_g = \{p, p, \neg q, p \supset q\}
\]

Note that \(p\) occurs twice in \(C_g\) as both parties have the goal of integrating the combatants of the RCD into the national Congolese army.

Using a horizontal arrangement for the truth table for the four formulae which displays the four possible value assignments at the bottom row, we can easily read off the densities of the value assignments by counting the 0s in the respective columns.
Chapter 6  Measuring Conflicts

The minimal densities are the densities of the two value assignments $e_1$ and $e_2$, each having a relative density with respect to $C_g$ of .25. Using inc-1*, the most plausible measure for the degree of inconsistency of a set, we can compute inc-1*(C_g) as follows.

\[
\text{inc-1}^*(C_g) = \min \{ d_{C_g}^*(e) \mid e \in E_{C_g} \} = .25
\]

The goal conflict between the RCD and the Kabila government regarding the integration of combatants of the RCD into the national Congolese army has a degree of inconsistency of .25. This number can be interpreted as follows. In order to come to a solution, 25% of the claims made by the two agents need to be weakened.

The second conflict for which we calculate the degree of inconsistency is the simple factual dispute between the RCD and the Kabila government regarding the potential threat of Banyamulenge being attacked in the Kivu provinces. The RCD believes that there is a threat, whereas the Kabila government believes that there is none.

<table>
<thead>
<tr>
<th>p</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>p \rightarrow q</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>\neg q</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$d_{C_g}^*(e_i)$</td>
<td>.25</td>
<td>.75</td>
<td>.25</td>
<td>.5</td>
</tr>
<tr>
<td>$E_{C_g}$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
</tr>
<tr>
<td>p=1 q=1</td>
<td>p=0 q=1</td>
<td>p=1 q=0</td>
<td>p=0 q=0</td>
<td></td>
</tr>
</tbody>
</table>

Table 14: Relative Densities of $C_g$

The minimal densities are the densities of the two value assignments $e_1$ and $e_2$, each having a relative density with respect to $C_g$ of .25. Using inc-1*, the most plausible measure for the degree of inconsistency of a set, we can compute inc-1*(C_g) as follows.

\[
\text{inc-1}^*(C_g) = \min \{ d_{C_g}^*(e) \mid e \in E_{C_g} \} = .25
\]

The goal conflict between the RCD and the Kabila government regarding the integration of combatants of the RCD into the national Congolese army has a degree of inconsistency of .25. This number can be interpreted as follows. In order to come to a solution, 25% of the claims made by the two agents need to be weakened.

The second conflict for which we calculate the degree of inconsistency is the simple factual dispute between the RCD and the Kabila government regarding the potential threat of Banyamulenge being attacked in the Kivu provinces. The RCD believes that there is a threat, whereas the Kabila government believes that there is none.

\[a_1: \text{Rassemblement Congolais pour la Democratic (RCD)}\]
\[a_2: \text{Kabila government}\]
\[r: \text{Banyamulenge are threatened to be attacked in the Kivu provinces}\]
\[C = \{Ba_1r, Ba_2\neg r\}\]
\[C_b = \{r, \neg r\}\]
The truth table for $C_{i0}$ displaying the relative densities of the value assignments, is given as follows.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$\neg r$</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$d_{C_{i0}}(e_1)$</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>$E_{C_{i0}}$</td>
<td>$e_1$ r=1</td>
<td>$e_2$ r=0</td>
</tr>
</tbody>
</table>

Table 15: Relative Densities of $C_{i0}$

The minimal densities are the densities of the two value assignments $e_1$ and $e_2$ each having a relative density with respect to $C_{i0}$ of .5. Hence, we get $\text{inc-1}^*(C_{i0})$ as follows.

$$\text{inc-1}^*(C_{i0}) = \min \{ d_{C_{i0}}(e) \mid e \in E_{C_{i0}} \} = .5$$

Thus, the factual dispute between the RCD and the Kabila government regarding the security situation of the Banyamalunge in the Kivu provinces has a degree of inconsistency of .5 which means that 50% of the beliefs need to be weakened in order to come to a solution.

Comparing the degree of inconsistency of the two conflicts, we can conclude that the goal conflict is easier to solve than the factual dispute in the sense that a smaller percentage of elements needs to be weakened.

6.6 Summary

The aim of this chapter was to introduce the concept of potential conflict power as a tool to assess the role propositional formulae play within a conflict as well as the concept of degrees of inconsistency as a tool to measure the depth of a conflict. This has been
achieved to the extent that we have given two independent, but equivalent, definitions of the potential conflict power of propositional formulae and three definitions for the degree of inconsistency of a set of propositional formulae, from which we have chosen inc-1* as the most plausible definition.

The semantic definition of potential conflict power is based on the idea that formulae that are hard to satisfy should have a higher degree of potential conflict power than formulae that are easy to satisfy. The idea behind the syntactic definition is to look at the number of formulae that are inconsistent with the formula in question and to assign a degree of potential conflict power to the formula that is equal to the logarithm of that number.

Having provided definitions for the potential conflict power and the degree of inconsistency, we can assess how much a claim, expressed by a propositional formula, contributes to a conflict, and, thus, which role it plays in a conflict. Furthermore, we can assess how difficult it is to solve a conflict expressed by a set of propositional formulae. The degree of inconsistency gives us a hint about how many elements of the conflict must be changed in order to resolve it.

Assessing formulae and sets of formulae with respect to the two introduced measures constitutes the fifth step in the process of modelling and resolving a conflict. It is the first step that goes beyond a purely descriptive and classificatory level.

In the following chapter, we will use the notions of potential conflict power and degrees of inconsistency as part of our definitions of five resolution algorithms which generate potential solutions to conflicts modelled by CML.
CHAPTER 7

Generating Solutions

The Resolution Algorithms of Conflict Resolution Logic

7.1 Introduction

The aim of chapter 7 is to specify five algorithms jointly referred to as Conflict Resolution Logic (CRL). The algorithms generate potential solution sets to conflicts having been modelled by CML. Each algorithm consists of two stages. In the first stage, a set of value assignments is computed for a given input of a set of propositional formulae. In the second stage, the algorithms generate a consistent solution set for each of the identified value assignments. Each of the five algorithms realises a different principle of conflict resolution by choosing different value assignments in the first stage.

As a result of the chapter we will be able to give answers to the following questions.

- What are the general conditions for a resolution algorithm, i.e. what kind of input does it require and what is its output?
- How can a consistent set of propositional formulae be generated on the basis of a given value assignment function?
Chapter 7  Generating Solutions

- How can a set of value assignment functions be selected on the basis of a set of propositional formulae?
- How do the five algorithms of CRL generate solution sets to conflicts previously modelled by CML?
- What complexity class are the algorithms in?
- How can the algorithms of CRL generate potential solutions to sub-conflicts of the Second Congo War?

An outline of the chapter is as follows: We start with a background section on dialogue logic. Then, we provide a general characterisation for the resolution algorithms of CRL in section 7.2. In section 7.3, we describe the part of the resolution algorithms of CRL which generates solution sets on the basis of given value assignments. This part is identical for all five algorithms. In section 7.4 we introduce the specific algorithms. Each of them selects a set of value assignments on the basis of an input set of propositional formulae. We describe each of the five algorithms and explain the principle realised by them in a separate section. CRL-1 is introduced in 7.4.1, CRL-2 in 7.4.2, CRL-3 in 7.4.3, CRL-4 in 7.4.4, and CRL-5 in 7.4.5. In section 7.5, we briefly address the aspect of computability and complexity of the algorithms. Finally, we illustrate the five algorithms by generating solutions to some sub-conflicts of the Second Congo War.

**Background: Dialogue Logic**

Dialogue logic, or dialogical logic, was developed by Paul Lorenzen and Kuno Lorenz in the 1960s as a logic intended to model dialogues, or arguments, among interlocutors. Their aim was to find a logical representation for the correct defence or refutation of assertions made by a disputant against an opponent disputant, which would represent
more closely the dynamics of the argument than a proof theoretical deduction from a number of premises or an axiom system. With their system of dialogue logic, Lorenzen and Lorenz also provided a new semantic characterisation of classical and intuitionist logic which incorporated pragmatic aspects into the semantics, such as the speech acts of asserting, attacking, and defending statements.

Their dialogue logic was designed as a game to be played by two players. The game can be won or lost by either of the two players, and, thereby, decides whether the initially asserted statement is refuted or defended. The game character of dialogue logic made it possible to link it with other game-theoretic approaches in logic, in particular with Jaakko Hintikkas ‘Semantic Games’.  

215

The seminal text on dialogue logic is Lorenz’s 1961 doctoral thesis *Arithmetic and Logic as Games*.  

216

His ideas were further developed by Lorenzen, Wilhelm Kamlah, Walther Kindt, Wolfgang Stegmueller, Walter Felscher, Shahid Rahman, Erik Krabbe, as well as himself, in a number of papers.  

217

The most comprehensive monograph on the topic is Lorenz and Lorenzen’s *Dialogische Logik* from 1978.  

218

In the following, we introduce the classical and the intuitionist versions of dialogue logic and illustrate them with some examples.

Dialogue logic is based on the idea of a dialogue between a proponent P, who asserts an initial thesis, and an opponent O, who tries to attack the thesis in order to prove the proponent wrong. The proponent P and the opponent O start a sequence of moves,

---

215 (Hintikka 1973; Hintikka and Sandu 1997)
216 (Lorenz 1961)
217 For an overview over recent developments in dialogue logic see, for instance, (Krabbe 2006; Felscher 2003; Hodges 2004).
218 (Lorenzen and Lorenz 1978)
Chapter 7  Generating Solutions

each of which is the utterance of a certain assertion. Each assertion is either an attack on one of the other player’s previous assertions or a defence of an assertion previously attacked by the other player. Altogether, a dialogue is constituted by a series of attacks and defences made by the two players P and O. A player wins the dialogue if the other player has no more moves left, i.e. if he is not able to attack or defend any more.

Graphically, dialogues are displayed in two columns. The left column contains assertions made by the opponent O, the right column contains the assertions made by the proponent P. Assertions are expressed by closed formulae of first-order logic. P starts the game by asserting a closed first-order formula \( \varphi \). The dialogue proceeds by alternating moves which are governed by six particle rules. The particle rules describe how the players can attack a formula asserted by the other player or defend themselves against an attack made by the other player. The type of attack or defence depends on the main operator of the formula that is attacked or defended. There are particle rules for each of the logical connectives \( \neg, \land, \lor, \supset \) and for the quantifiers \( \exists \) and \( \forall \). The particle rules are displayed in the following table.\(^{219}\)

<table>
<thead>
<tr>
<th>Attack</th>
<th>Defence</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \neg \varphi )</td>
<td>( \varphi )</td>
</tr>
<tr>
<td>( \varphi \land \psi )</td>
<td>( ?l )</td>
</tr>
<tr>
<td></td>
<td>( ?r )</td>
</tr>
<tr>
<td>( \varphi \lor \psi )</td>
<td>( ? )</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>( \varphi \supset \psi )</td>
<td>( ? )</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>( \exists x \varphi )</td>
<td>( ? )</td>
</tr>
<tr>
<td>( \forall x \varphi )</td>
<td>( ?n )</td>
</tr>
</tbody>
</table>

Table 16: Particle Rules

\(^{219}\) The table is a slightly different version of the table presented in (Lorenzen 1969).
Attacks are symbolised by the question mark. Atomic formulae can neither be attacked nor defended. Negated formulae can only be attacked but not defended. Conjunctions can be attacked and defended in two ways. Either the left or the right constituent can be attacked and must be defended accordingly. Disjunctions can be attacked only as a whole but there are two possible defences for them. In the case of an existentially quantified formula, the defender can chose an individual constant $n$ with which he defends the formula. If the formula is quantified by a universal quantifier the attacker can choose an arbitrary $n$. Note that not every attack involves the assertion of a new statement. Only attacking a negation or an implication requires the attacker to assert a statement himself which can then be attacked by the other player.

In addition to the particle rules, which govern the individual moves of the players, there are further rules describing how the whole process of a dialogue proceeds. These global rules are called structural rules for formal dialogues. We quote them from a paper by Shahid Rahman and Walter Carnielli as they provide a particularly simple presentation of the structural rules.\(^{220}\)

**Starting Rule**

The Proponent begins by asserting a thesis.

**Moves**

The players make their moves alternately. Each move, with the exception of the starting move, is an attack or a defence.

\(^{220}\) (Rahman and Carnielli 2000)
**Formal Rule**

Atomic statements cannot be attacked. The Proponent may use an atomic statement in a move if and only if the Opponent has already stated the same statement before.

**Winning Rule**

If any player cannot make any further move (without producing a repetition of identical rounds) the other has won.

**Intuitionistic Rule**

In any move, each player may attack a (complex) statement asserted by his partner or he may defend himself against the last not already defended attack, according to the particle and the other structural rules.

**Classical Rule**

In any move, each player may attack a (complex) statement asserted by his partner or he may defend himself against any attack (including already defended ones) of the Opponent, according to the particle and the other structural rules.

Using the particle rules and the structural rules, we can construct a dialogue for any closed first-order formula \( \varphi \) asserted as a thesis by the proponent P.

Before we illustrate the rules with an example, we look at the link between dialogues, as described by dialogue logic, and classical and intuitionistic logic. A formula \( \varphi \) is classically valid if there is a winning strategy for a proponent P starting with \( \varphi \) as a thesis in a dialogue that satisfies the Classical Rule, i.e. it is possible for P to make moves that necessarily bring O into a situation in which he cannot make any more moves. A formula
\( \varphi \) is intuitionistically valid if there is a winning strategy for a proponent \( P \) starting with \( \varphi \) as a thesis in a dialogue that satisfies the Intuitionistic Rule.

Attempts to prove equivalence theorems between dialogue logic and classical/intuitionistic logic were made by Lorenz.\(^{221}\) However, his proofs contained mistakes. A first correct proof was given by Kindt.\(^{222}\) Further simpler proofs were presented by Krabbe, Barth, and Felscher.\(^{223}\)

To illustrate dialogue logic, we construct dialogues for the three formulae \( \neg \neg p \supset p \), \( \neg p \land q \), and \( \neg (p \land \neg p) \). The players’ moves are numbered in brackets. Numbers in square brackets behind the attacks and defences refer to the formulae that are attacked and defended, respectively. A question mark symbolises an attack, whereas an exclamation mark represents a defence.

A dialogue for \( \neg \neg p \supset p \) is given as follows.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1) ( \neg \neg p \supset p )</td>
<td>(2) ( \neg \neg p ) [?]</td>
<td>(3) ( \neg p ) [?]</td>
</tr>
<tr>
<td>(4) p [?]</td>
<td>(5) p [!]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O has no move left</td>
<td>P wins</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 17: Dialogue for \( \neg \neg p \supset p \)

Move (5) is not allowed according to the intuitionistic rule as the formula \( \neg p \) was attacked by \( O \) in move (3) but not defended by \( P \). As the intuitionistic rule allows \( P \) to defend himself only against the last not already defended attack, \( P \) is required to defend

---

\(^{221}\) (Lorenz 1961; Lorenz 1968)  
\(^{222}\) (Kindt 1972)  
\(^{223}\) (Krabbe 1982; Barth and Krabbe 1982; Felscher 1985)
\(\neg p\) first before defending \(\neg \neg p \supset p\). The dialogue is compatible with the classical rule, though, and so \(\neg \neg p \supset p\) is classically, but not intuitionistically, valid. Note that P would not be allowed to defend himself immediately against O’s attack of \(\neg \neg p \supset p\) in move (3) by asserting \(p\) because O has not stated \(p\) yet at that stage.

A dialogue for \(\neg p \land q\) is given as follows.

<table>
<thead>
<tr>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) (\neg p \land q)</td>
<td></td>
</tr>
<tr>
<td>(2) (?l [?1])</td>
<td>(3) (\neg p [?1])</td>
</tr>
<tr>
<td>(4) (p [?2])</td>
<td>(5) P has no move left</td>
</tr>
</tbody>
</table>

**Table 18: Dialogue for \(\neg p \land q\)**

In this dialogue, every move is compatible with both the classical and the intuitionist rule. However, the dialogue does not prove that the formula is either classically or intuitionistically valid since O wins. This does not necessarily mean that the formula is invalid as there might be other winning strategies for P.

Finally we look at a formula which is both classically and intuitionistically valid. A dialogue in which P has a winning strategy for the formula \(\neg(p \land \neg p)\) is shown in the following table.

<table>
<thead>
<tr>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) (\neg(p \land \neg p))</td>
<td></td>
</tr>
<tr>
<td>(2) (p \land \neg p [?1])</td>
<td>(3) (?r [?2])</td>
</tr>
<tr>
<td>(4) (\neg p [?2])</td>
<td>(5) (?l [?2])</td>
</tr>
<tr>
<td>(6) (p [?2])</td>
<td>(7) (p [?4])</td>
</tr>
<tr>
<td>(8) O has no move left</td>
<td>P wins</td>
</tr>
</tbody>
</table>

**Table 19: Dialogue for \(\neg(p \land \neg p)\)**
As every move in the dialogue is compatible with both the classical and the intuitionistic rule, P has a winning strategy for $\neg(p \land \neg p)$ in classical and intuitionistic dialogue logic. Hence, the formula is classically and intuitionistically valid.

Dialogue logic has been extended in various directions. Extensions have proven particularly fruitful for the modelling of dialogues involving further linguistic elements, such as questions, answers, or commands.\textsuperscript{224} Dialogue logic has also been applied as a constructive proof theory for other logical systems, such as modal logic and paraconsistent logic.\textsuperscript{225}

### 7.2 General Characterisation of the Resolution Algorithms

For a given conflict C, we can generate solutions by making consistent the inconsistent sets representing its sub-conflicts. To do this, we first have to transform the sub-conflicts into propositional form. The consistent sets we generate represent potential solutions to the sub-conflicts which the parties involved can reach. The consistency of the solution set guarantees that it is possible to realise all the propositions that are either directly contained in the set or follow logically from it.

Once a solution is generated, we can look at the agents’ original propositional attitudes and check which of them are fully satisfied according to the solution, which are partly satisfied, and which are not satisfied at all. For instance, if a party a has the goal Ga(p $\land$ q), this goal would be fully satisfied if p and q were in the solution set, it would be partially satisfied if only p were contained in the solution set, and it would not be satisfied

\textsuperscript{224} (Atkinson et al. 2008; Girle 1996)

\textsuperscript{225} (Rueckert and Rahman 1999; Rahman and Carnielli 2000)
at all if neither p nor q were in the solution set. Obviously, not all of the original propositional attitudes making up the conflict can be fully satisfied in a solution as then the parties would not have had a conflict in the first place. However, for each party some of its propositional attitudes might be fully or partly satisfied.

A generated solution can be proposed to the agents involved in the conflict. As the solution indicates which propositional attitudes need to be abandoned, as well as which ones will be partly and fully satisfied, the solution set provides the means of convincing the agents to accept the solution by pointing out their potential gains.

At the core of the process of generating solutions to a given conflict is the task of transforming an inconsistent set of propositional formulae into a consistent one. This task is accomplished by a number of algorithms which we jointly call Conflict Resolution Logic (CRL). Each algorithm of CRL allows us to generate possible solutions to a conflict given in propositional form.

As there are many different ways of making an inconsistent set of formulae consistent, we first characterise the resolution algorithms of CRL by a general condition. This general characterisation outlines the overall procedure of how consistency is obtained and introduces the main concepts of an input set and a solution set.

The general characterisation of CRL is given as follows.

**Definition 102 (General Characterisation of CRL)**
The resolution algorithms of CRL take as an input a finite set $\Phi$ of propositional formulae and generate as an output a number of finite, consistent, sets $\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*$ such that each $\Phi_i^*$ is obtained from $\Phi$ by replacing the formulae $\varphi \in \Phi$ by formulae $\varphi^*$ satisfying the three conditions:

1. $\varphi \vdash \varphi^*$;
2. $\text{conf}(\varphi^*) \leq \text{conf}(\varphi)$;
3. $\text{pro}(\varphi^*) \subseteq \text{pro}(\Phi)$;

until a consistent set of formulae is obtained.
The general characterisation of CRL specifies that the algorithms of CRL take an arbitrary finite set $\Phi$ of propositional formulae as an input and output a number of finite consistent sets $\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*$ of propositional formulae. It is not excluded that the input set $\Phi$ is already consistent. If this is the case, the algorithms just output $\Phi$. The process of obtaining the output sets $\Phi_i^*$ involves running through a sequence of steps by which a number of formulae contained in the input set $\Phi$ are replaced by other formulae.

The process of replacing a formula $\varphi$ by $\varphi^*$ is governed by three conditions. First, $\varphi^*$ has to be a logical consequence of $\varphi$. This makes sure that $\varphi$ is replaced by a formula which is as close as possible at the original formula. Being a logical consequence of $\varphi$, $\varphi^*$ represents a statement whose truth has already been claimed by $\varphi$. Second, the potential conflict power of $\varphi^*$ has to be smaller than that of $\varphi$. This condition makes sure that the set containing the new formula $\varphi^*$ is more likely to be consistent than the set containing $\varphi$. Indeed, if we did not reduce the potential conflict power of the formulae of $\Phi$, we would not make $\Phi$ consistent. The third condition, according to which the new formula $\varphi^*$ is not allowed to contain propositional constants other than the constants already contained in $\Phi$, is required in order to make sure that formulae are not replaced by new formulae containing arbitrary new propositions that are irrelevant to the original conflict.

Before we turn to the details of the five resolution algorithms of CRL in the following section, we compare CRL with the AGM model of belief revision and Lorenz’s dialogue logic.

As described in the background section on belief revision, the operation of making an inconsistent set consistent is captured by the consolidation operator $!$ in the AGM model. We also showed that $!$ can only be defined reasonably for belief bases. For a given contraction operator $\cdot$, $!$ is defined as contraction by the falsum $\bot$, i.e. if $A$ is a belief
Chapter 7  Generating Solutions

base, we get $A\! = A \lor \bot$. Both $\!$ and the algorithms of CRL require sets of propositional formulae as input and outputs consistent sets. With regard to the input, CRL requires a finite set of propositional formulae, whereas $\!$ can be applied to any set of propositional formulae. With regard to the output, CRL produces a number of consistent solution sets, whereas $\!$ produces only one consistent set.

The main difference between the two approaches lies in the way consistency is achieved. In contrast to the AGM model, where sets are made consistent by removing formulae, the algorithms of CRL only replace the elements of an inconsistent set by formulae that have a lower degree of potential conflict power. Thus, the algorithms of CRL achieve consistency with a smaller loss of information compared to the update operator of AGM. As a result, the agents holding propositional attitudes towards propositions contained in a set are not forced to completely abandon their propositional attitudes once the set is made consistent. They only have to weaken some of their attitudes by replacing them with attitudes having a lower degree of potential conflict power.

Although both models are intended to model arguments, CRL has a fundamentally different focus than dialogue logic. In order to compare dialogue logic with CRL, we interpret dialogues as decision algorithms that take a proponent’s initial thesis as an input and come up with an answer to the question of whether or not the initial thesis is a theorem of classical/intuitionistic logic. Understood this way, dialogues are decision procedures that can be applied to single formulae. In contrast, the algorithms of CRL are applied to sets of formulae.

The main difference between CRL and dialogue logic is that dialogues always have a winner and a loser, i.e. the formula is either defended or refuted by the dialogue, whereas the solution sets generated by the algorithms of CRL offer gains and losses to all parties
involved in a conflict. Hence, in CRL there are no winners and losers as in the arguments modelled by dialogue logic.

### 7.3 The Algorithm Generating Solution Sets

The general characterisation of CRL given thus far does not define an algorithm as there are infinitely many ways to replace the elements of the input set $\Phi$. Thus, we have to specify the conditions of the general characterisation of CRL. In particular, we have to determine the unique output sets $\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*$ generated by CRL for a given input set $\Phi$.

The specification of CRL is based on the semantic concept of value assignment functions. For a given input set $\Phi$ each algorithm of CRL first identifies a set of value assignments $e_0, e_1, \ldots, e_n \in E_\Phi$. Each of these value assignments $e_i$ determines a unique consistent solution set $\Phi_i^*$, which is generated by the algorithms in the second step. Finally, the algorithms output all the generated solution sets $\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*$.

The first part of the algorithms, the selection of the value assignments, is carried out differently by each of the five algorithms. The second part of the algorithms, the process of generating a solution set $\Phi_i^*$ on the basis of the input set $\Phi$ and a value assignment $e_i$, is identical for all five algorithms of CRL and is carried out by an independent algorithm called SOL\textsuperscript{226} which is defined as follows.

**Definition 103 (SOL)**

The algorithm SOL takes as an input a finite set $\Phi$ of propositional formulae and a value assignment $e \in E_\Phi$. It runs through the following steps:

Step 0: Input $\Phi, e$;

\textsuperscript{226} The acronym SOL indicates that the algorithm generates solution sets to a given conflict.
Step 1: Generate the set $LIT_e = \{p \mid e(p) = 1\} \cup \{-p \mid e(p) = 0\}$;
Step 2: Generate the formula $\alpha_e = \bigwedge_{p \in LIT_e} \beta$;
Step 3: Generate the set $\Phi^* = \{\phi \lor \alpha_e \mid \phi \in \Phi\}$;
Step 4: Output $\Phi^*$.

For a given input $\Phi$, $e$, SOL collects all the propositional constants true under $e$ and the negation of all the constants false under $e$ (step 1). This is a finite process as the value assignment $e$ is restricted to the finite set of constants occurring in the finite set $\Phi$. Next, the algorithm generates the finite conjunction over all these literals and names this conjunction $\alpha_e$ (step 2). Then, the algorithm generates the set of all disjunctions $\phi \lor \alpha_e$ constituted by a formula $\phi \in \Phi$ and $\alpha_e$ and names this set $\Phi^*$ (step 3). Finally, the algorithm outputs $\Phi^*$ (step 4).

Note that due to step 3, every formula contained in $\Phi^*$ is satisfied by $e$. Furthermore, if a formula $\phi$ contained in $\Phi$ is already satisfied by $e$, it remains unchanged, up to logical equivalence, in the solution set because $\phi \lor \alpha_e$ is logically equivalent to $\phi$ if $\alpha_e \vDash \phi$.

In the following we prove that the output set $\Phi^*$ satisfies the three conditions specified in the general characterisation of CRL, i.e. it is finite, consistent, and, for every $\phi^* \in \Phi^*$, there is an element $\phi \in \Phi$ such that $\phi \vdash \phi^*$, $\text{conf}(\phi^*) < \text{conf}(\phi)$, and $\text{pro}(\phi^*) \subseteq \text{pro}(\Phi)$.

**Theorem 46** (SOL and the General Characterisation of CRL)

Let $\Phi$ be a finite set of propositional formulae and $e \in E_\Phi$, a value assignment restricted to $\text{pro}(\Phi)$. Then, the output set $\Phi^*$, i.e. the result of applying the algorithm SOL to $\Phi$ and $e$, is finite and consistent and for every $\phi^* \in \Phi^*$ there is a $\phi \in \Phi$ such that $\phi \vdash \phi^*$, $\text{conf}(\phi^*) < \text{conf}(\phi)$, and $\text{pro}(\phi^*) \subseteq \text{pro}(\Phi)$.

**Proof**

(Finiteness) $\Phi^*$ is finite as $\Phi$ is finite and, by definition of $\Phi^*$, for every element $\phi^* \in \Phi^*$ there is a unique element $\phi \in \Phi$ such that $\phi^*$ is of the form $\phi \lor \alpha_e$.

(Consistency) By definition of $\alpha_e$, $e$ is a model of $\alpha_e$. Hence, by the truth condition for $\lor$, $e$ is a model of any formula of the form $\phi \lor \alpha_e$. By definition of
\( \Phi^* \) every element \( \varphi^* \in \Phi^* \) is of the form \( \varphi \lor \alpha_e \) for some \( \varphi \in \Phi \). Hence, \( e \) is a model of \( \Phi^* \). Therefore, \( \Phi^* \) is consistent by completeness.

\( (\varphi \vdash \varphi^*) \) By definition of \( \Phi^* \) for every element \( \varphi^* \in \Phi^* \) there is an element \( \varphi \in \Phi \) such that \( \varphi^* \) is of the form \( \varphi \lor \alpha_e \). Then, \( \varphi \vdash \varphi \lor \alpha_e \), and, hence, \( \varphi \vdash \varphi^* \).

\( (\text{conf}(\varphi^*) < \text{conf}(\varphi)) \) By definition of \( \Phi^* \) for every element \( \varphi^* \in \Phi^* \) there is an element \( \varphi \in \Phi \) such that \( \varphi^* \) is of the form \( \varphi \lor \alpha_e \). As \( \text{conf}(\varphi \lor \varphi_e) < \text{conf}(\varphi) \) we get \( \text{conf}(\varphi^*) < \text{conf}(\varphi) \).

\( (\text{pro}(\varphi^*) \subseteq \text{pro}(\Phi)) \) By definition of \( \Phi^* \) every element \( \varphi^* \in \Phi^* \) is of the form \( \varphi \lor \alpha_e \) for some \( \varphi \in \Phi \). By definition of \( \alpha_e \) we get \( \text{pro}(\varphi_e) \subseteq \text{pro}(\Phi) \) and \( \text{pro}(\varphi) \subseteq \text{pro}(\Phi) \) as \( \varphi \in \Phi \). Hence, \( \text{pro}(\varphi^*) = \text{pro}(\varphi \lor \alpha_e) \subseteq \text{pro}(\alpha_e) \cup \text{pro}(\varphi) \subseteq \text{pro}(\Phi) \).

QED

### 7.4 Algorithms Selecting Value Assignments

With SOL we have an algorithm which produces a consistent set of formulae \( \Phi^* \) if applied to a set of formulae \( \Phi \) and a value assignment function \( e \). As mentioned above, SOL provides the second part of each of the five algorithms CRL-1, CRL-2, CRL-3, CRL-4, and CRL-5.

In the first part of the algorithms, a set \( E \) of value assignments is determined on the basis of a set of propositional formulae \( \Phi \). This process differs for each of the five algorithms.

Before we describe the algorithms, we reiterate the concept of absolute density of a value assignment \( e \) with respect to a set of formulae \( \Phi \), as introduced in Definition 100. By the absolute density \( d_\Phi(e) \) of a value assignment \( e \in E_\Phi \) with respect to the set of formulae \( \Phi \), we mean the number of formulae of \( \Phi \) which are false under \( e \), i.e.

\[
|e \notin \mod_\Phi(e)|.
\]
7.4.1 The Density Minimising Algorithm

We now describe the first algorithm CRL-1, called the density minimising algorithm. Its formal definition is given as follows.

**Definition 104 (CRL-1)**
The algorithm CRL-1 takes a finite set $\Phi$ of propositional formulae as input and outputs a finite set $\{\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*\}$ of finite and consistent sets $\Phi_i^*$. It runs through the following steps:

1. **Step 0**: Input $\Phi$.
2. **Step 1**: Generate the set $E_{\Phi}$.
3. **Step 2**: Generate the set $\text{DEN} = \{d_\Phi(e) \mid e \in E_{\Phi}\}$.
4. **Step 3**: Generate the set $E_{\text{minDen}} = \{e \mid d_\Phi(e) = \min(\text{DEN})\}$.
5. **Step 4**: For each $e_i \in E_{\text{minDen}}$ apply SOL to $\Phi, e_i$.
6. **Step 5**: Collect the outputs of step 4 in the set $\{\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*\}$.
7. **Step 6**: Output $\{\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*\}$.

For a given input $\Phi$, CRL-1 first generates the set $E_{\Phi}$ of all possible value assignments restricted to the constants occurring in $\Phi$ (step 1). It then computes, for every value assignment $e \in E_{\Phi}$, its absolute density $d_\Phi(e)$ and collects them in the set $\text{DEN}$ (step 2). Next, the algorithm identifies those value assignments whose density is minimal, i.e. those $e \in E_{\Phi}$ for which $d_\Phi(e) = \min(\text{DEN})$ holds, and collects them in the set $E_{\text{minDen}}$ (step 3). Note that $E_{\text{minDen}}$ can contain more than one element as there can be more than one value assignment with the same minimal absolute density. Next, the algorithm SOL is applied to each $e_i \in E_{\text{minDen}}$ (step 4). As a result, for each value assignment $e_i$ contained in $E_{\text{minDen}}$, a corresponding consistent solution set $\Phi_i^*$ is generated. All generated solution sets are collected in a set (step 5) which constitutes the output of CRL-1 (step 6).

Without providing the details, it is assumed that each step of CRL-1 is computable. In each step the algorithm only applies simple computable operations to finite sets generating further finite sets. Hence, in each step only a finite number of stipulations is executed.
With CRL-1 we have an algorithm which generates a set of consistent sets \{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\} for a given finite set \Phi. Each of these sets can be suggested to the parties involved in the conflict as a potential solution. We will provide an interpretation of the type of solutions generated by CRL-1 in the last chapter. Note that if the input set \Phi is already consistent, CRL-1 just produces \Phi itself as an output.

**7.4.2 The Density Minimising, \(\Delta\) Preserving Algorithm**

The second algorithm of CRL-2 called density minimising, \(\Delta\) preserving algorithm is defined as follows.

**Definition 105 (CRL-2)**
The algorithm CRL-2 takes a finite set \(\Phi\) of propositional formulae and a consistent subset \(\Delta \subseteq \Phi\) as input and outputs a finite set \{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\} of finite and consistent sets \(\Phi_i^*\). It runs through the following steps:

Step 0: Input \(\Phi, \Delta\);

Step 1: Generate the set \(E_{\Phi}\);

Step 2: Generate the set \(\text{mod}_{\Phi}(\Delta)\);

Step 3: Generate the set \(\text{DEN}_{\Delta} = \{d_{\Phi}(e) \mid e \in \text{mod}_{\Phi}(\Delta)\}\);

Step 4: Generate the set \(E_{\text{minDen}_{\Delta}} = \{e \mid d_{\Phi}(e) = \min(\text{DEN}_{\Delta})\}\);

Step 5: For each \(e_i \in E_{\text{minDen}_{\Delta}}\) apply SOL to \(\Phi, e_i\);

Step 6: Collect the outputs of step 5 in the set \{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\};

Step 7: Output \{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}.

The algorithm CRL-2 differs from the other algorithms with respect to the input it requires. In addition to the input set \(\Phi\), it requires a set \(\Delta\) of propositional formulae which is a consistent subset of \(\Phi\). For a given input, CRL-2 generates the set \(\text{mod}_{\Phi}(\Delta)\) of all values assignments that satisfy the formulae contained in \(\Delta\) (step 2). It then generates the set \(\text{DEN}_{\Delta}\) of all absolute densities of these value assignments (step 3). From here it continues in a similar way to CRL-1, i.e. it generates the set \(E_{\text{minDen}_{\Delta}}\) of value assignments contained in \(\text{mod}_{\Phi}(\Delta)\) whose absolute density is minimal in \(\text{DEN}_{\Delta}\) (step 4) and produces a consistent solution set \(\Phi_i^*\) for each \(e_i \in E_{\text{minDen}_{\Delta}}\) by virtue of the algorithm SOL.
We will postpone the interpretation of the additional input $\Delta$, as well as the type of solutions generated by CRL-2, to the final chapter. Again, if the input set $\Phi$ is already consistent, CRL-2 just produces $\Phi$ itself as an output.

### 7.4.3 The Minimax Algorithm

The third algorithm, CRL-3, called the minimax algorithm, requires a finite, exhaustive collection $\{\Phi_i\}_{i \in I}$ of subsets $\Phi_i \subseteq \Phi$ as an input in addition to the set $\Phi$. By exhaustive we mean that the union $\bigcup_{i \in I} \Phi_i$ is equal to $\Phi$, where $I$ can be any finite index set. We call elements of $\{\Phi_i\}_{i \in I}$ cells. Note that $\{\Phi_i\}_{i \in I}$ is not required to be exclusive, i.e. elements of $\Phi$ can appear in different cells $\Phi_i, \Phi_j \in \{\Phi_i\}_{i \in I}$. From now on, we assume that $\{\Phi_i\}_{i \in I}$ always refers to a finite, exhaustive collection of subsets $\Phi_i \subseteq \Phi$.

CRL-3 is defined as follows.

**Definition 106 (CRL-3)**

The algorithm CRL-3 takes a finite set $\Phi$ of propositional formulae and a finite, exhaustive collection $\{\Phi_i\}_{i \in I}$ of subsets $\Phi_i \subseteq \Phi$ as input and outputs a finite set $\{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$ of finite and consistent sets $\Phi_i^*$. It runs through the following steps:

- **Step 0:** Input $\Phi, \{\Phi_i\}_{i \in I}$;
- **Step 1:** Generate the set $E_\Phi$;
- **Step 2:** For every $e \in E_\Phi$, generate the set $\{d_\Phi(e) | i \in I\}$;
- **Step 3:** For every $e \in E_\Phi$, identify $\max_e = \max\{d_\Phi(e) | i \in I\}$;
- **Step 4:** Generate the set $\text{MAX} = \{\max_e | e \in E_\Phi\}$;
- **Step 5:** Generate the set $E_{\text{minMax}} = \{e | d_\Phi(e) = \min(\text{MAX})\}$;
- **Step 6:** For each $e_i \in E_{\text{minMax}}$, apply SOL to $\Phi, e_i$;
- **Step 7:** Collect the outputs of step 6 in the set $\{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$;
- **Step 8:** Output $\{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$.

In contrast to CRL-1, CRL-3 does not generate the absolute densities of value assignments with respect to the whole set $\Phi$. Instead, for every value assignment $e \in E_\Phi$, \ldots
it calculates the set of all densities $d_{\Phi_i}(e)$ with respect to the cells $\Phi_i$ of $\{\Phi_i\}_{i \in I}$ (step 2). It then identifies the maxima of these sets and collects them in the set $\text{MAX}$. Finally, it proceeds, like the previous algorithms, by generating the set $E_{\text{minMax}}$ of value assignments whose absolute density is minimal in $\text{MAX}$ (step 5) and producing a consistent solution set $\Phi_i^*$, for each $e_i \in E_{\text{minMax}}$, by virtue of the algorithm SOL.

The interpretation of the additional input $\{\Phi_i\}_{i \in I}$ and the type of solutions generated by CRL-3, will be given in the final chapter. If the input set $\Phi$ is already consistent, CRL-3 produces $\Phi$ itself as an output.

7.4.4 The Difference Minimising Algorithm

Similar to CRL-3, the fourth algorithm, CRL-4, called difference minimising algorithm, also requires $\{\Phi_i\}_{i \in I}$ as an input in addition to $\Phi$. It is defined as follows.

**Definition 107 (CRL-4)**

The algorithm CRL-4 takes a finite set $\Phi$ of propositional formulae and a finite, exhaustive collection $\{\Phi_i\}_{i \in I}$ of subsets $\Phi_i \subseteq \Phi$ as input and outputs a finite set $\{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$ of finite and consistent sets $\Phi_i^*$. It runs through the following steps:

- **Step 0**: Input $\Phi$, $\{\Phi_i\}_{i \in I}$;
- **Step 1**: Generate the set $E_{\Phi}$;
- **Step 2**: For every $e \in E_{\Phi}$, generate the set $\{d_{\Phi_i}(e) \mid i \in I\}$;
- **Step 3**: For every $e \in E_{\Phi}$, identify $\max_e = \max(\{d_{\Phi_i}(e) \mid i \in I\})$;
- **Step 4**: For every $e \in E_{\Phi}$, identify $\min_e = \min(\{d_{\Phi_i}(e) \mid i \in I\})$;
- **Step 5**: Generate the set $\text{DIF} = \{\max_e - \min_e \mid e \in E_{\Phi}\}$;
- **Step 6**: Generate the set $E_{\text{minDIF}} = \{e \mid \max_e - \min_e = \min(\text{DIF})\}$;
- **Step 7**: For each $e_i \in E_{\text{minDIF}}$, apply SOL to $\Phi_i$, $e_i$;
- **Step 8**: Collect the outputs of step 7 in the set $\{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$;
- **Step 9**: Output $\{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$.

CRL-4 starts, like CRL-3, by computing, for every value assignment $e \in E_{\Phi}$, the sets of absolute densities $d_{\Phi_i}(e)$ with respect to the cells $\Phi_i$ of the partition $\{\Phi_i\}_{i \in I}$ (step 2). It
then identifies the maximum \( \max_e \) (step 3) and the minimum \( \min_e \) (step 4) of these sets and generates the set DIF of all differences of the form \( \max_e - \min_e \) (step 5). Next, the algorithm generates the set \( E_{\min\text{DIF}} \) of all value assignments \( e \) for which \( \max_e - \min_e \) is a minimum of DIF (step 6). From then, it proceeds like the other algorithms.

### 7.4.5 The Conflict Power Balancing Algorithm

Finally, we provide a definition for the last of the five algorithms, CRL-5, called conflict power balancing algorithm. CRL-5 also requires \( \{\Phi_i\}_{i \in I} \) as an input in addition to \( \Phi \). Its formal definition is given as follows.

**Definition 108 (CRL-5)**

The algorithm CRL-5 takes a finite set \( \Phi \) of propositional formulae and a finite, exhaustive collection \( \{\Phi_i\}_{i \in I} \) of subsets \( \Phi_i \subseteq \Phi \) as input and outputs a finite set \( \{\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*\} \) of finite and consistent sets \( \Phi_i^* \). It runs through the following steps:

1. **Step 0:** Input \( \Phi, \{\Phi_i\}_{i \in I} \);
2. **Step 1:** Generate the set \( E_{\Phi} \);
3. **Step 2:** Generate the set \( \{\text{conf}(\varphi) \mid \varphi \in \Phi\} \);
4. **Step 3:** Generate the set \( \{\text{mod}_\Phi(\varphi) \mid \varphi \in \Phi\} \);
5. **Step 4:** For every \( e \in E_{\Phi} \), generate the set \( \Phi_e = \{\varphi \in \text{mod}_\Phi(\varphi) \mid \varphi \in \Phi\} \);
6. **Step 5:** For every \( e \in E_{\Phi} \), generate the set \( \{\sum_{\varphi \in \Phi_i \cap \Phi} \text{conf}(\varphi) \mid i \in I\} \);
7. **Step 6:** For every \( e \in E_{\Phi} \), identify \( \max_e = \max\{\sum_{\varphi \in \Phi_i \cap \Phi} \text{conf}(\varphi) \mid i \in I\} \);
8. **Step 7:** For every \( e \in E_{\Phi} \), identify \( \min_e = \min\{\sum_{\varphi \in \Phi_i \cap \Phi} \text{conf}(\varphi) \mid i \in I\} \);
9. **Step 8:** Generate the set \( \text{DIFCON} = \{\max_e - \min_e \mid e \in E_{\Phi}\} \);
10. **Step 9:** Generate the set \( E_{\min\text{DIFCON}} = \{e \mid \max_e - \min_e = \min(\text{DIFCON})\} \);
11. **Step 10:** For each \( e_i \in E_{\min\text{DIFCON}} \), apply SOL to \( \Phi, e_i \);
12. **Step 11:** Collect the outputs of step 10 in the set \( \{\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*\} \);
13. **Step 12:** Output \( \{\Phi_1^*, \Phi_2^*, \ldots, \Phi_n^*\} \).

CRL-5 makes use of the concept of potential conflict power \( \text{conf} \). For every cell \( \Phi_i \) of \( \{\Phi_i\}_{i \in I} \), and every value assignment \( e \in E_{\Phi} \), it calculates the sum \( \sum_{\varphi \in \Phi_i \cap \Phi} \text{conf}(\varphi) \) of the potential conflict power of formulae which are elements of this cell true under \( e \) and
collects these sums in a set (step 5). The algorithm then continues by identifying the maximum \( \max_e \) (step 6) and minimum \( \min_e \) (step 7) of these sets for every \( e \in E_\Phi \). Next, it generates the set \( \text{DIFCON} \) of differences of the form \( \max_e - \min_e \) for every \( e \in E_\Phi \) (step 8) and identifies the set \( E_{\min \text{DIFCON}} \) of value assignments for which \( \max_e - \min_e \) is minimal in \( \text{DIFCON} \) (step 9). From there, it proceeds in a similar way to the other algorithms.

### 7.5 Computability and Complexity

In this section we have a brief look at some computational aspects of the five algorithms of CRL. We have described them all in the same way. For a given input, they all run through a number of steps and produce a finite set \( \{ \Phi_1^*, \Phi_2^*, \ldots, \Phi_n^* \} \) of finite and consistent sets \( \Phi_i^* \). They all require a finite set \( \Phi \) of propositional formulae as an input. CRL-2 requires a consistent subset \( \Delta \subseteq \Phi \) as an additional input, whereas CRL-3, CRL-4, and CRL-5 require a finite, exhaustive collection, \( \{ \Phi_i \}_{i \in I} \) of subsets \( \Phi_i \subseteq \Phi \) as an additional input. In all cases, the input consists of a finite string of symbols of the language of propositional logic.

The steps executed by the algorithms represent lists of sub-tasks. Each sub-task is achieved by a number of sub-steps. In our description, we have not provided descriptions of the sub-steps for the sake of clarity. The sub-tasks represented by the steps are simple, in most cases set theoretic, operations on sets. They include tasks like generating the set of all value assignments \( E_\Phi \), generating the set of all absolute densities \( \{d_\Phi(e) \mid e \in E_\Phi\} \), or identifying the minimum of a set. In general, they start with a given finite set and generate another finite set.
As an example, we describe one sub-task in more detail. Step 1 of each algorithm represents the sub-task of generating the set of all value assignments $E_\Phi$ restricted to the set of constants occurring in $\Phi$. This task can be performed by the following sub-steps.

**Definition 109 (Sub-steps of Step 1)**

Step 1.1: Input $\Phi$;
Step 1.2: Generate the set $\text{pro}(\Phi)$;
Step 1.3: Calculate $|\text{pro}(\Phi)|$;
Step 1.4: Enumerate $\text{pro}(\Phi) = \{p_i\}_{i \in \{1, \ldots, |\text{pro}(\Phi)|\}}$;
Step 1.5: Generate the Cartesian product $\{0, 1\}^{|\text{pro}(\Phi)|}$;
Step 1.6: Enumerate $\{0, 1\}^{|\text{pro}(\Phi)|} = \{e_i\}_{i \in \{1, \ldots, 2^{|\text{pro}(\Phi)|}\}}$;
Step 1.7: Output $E_\Phi = \{e_i\}_{i \in \{1, \ldots, 2^{|\text{pro}(\Phi)|}\}}$.

The definition illustrates which sub-steps are needed in order to generate $E_\Phi$. However, it is still not at the most basic level as the sub-steps represent tasks which might not be primitive in the implementation language of the algorithm. For instance, the sub-step 1.2, generating the set of constants occurring in $\Phi$, consists of further sub-tasks, such as scanning the symbols of $\Phi$, deciding whether they are constants or not and collecting the constants in a set. The question of which tasks count as primitive and which do not depends on the actual implementation language used to program the algorithm.

For us the crucial facts are the following.

1. Each of the five algorithms of CRL is comprised of a finite number of steps.
2. Each step can be executed by a finite number of sub-steps.
3. Each sub-step is deterministic, i.e. for a given input, each sub-step will always produce the same output and will always pass through the same sequence of states.
4. Each sub-step terminates.
From this it follows that every step is deterministic and terminates, and, hence, each of the five algorithms of CRL is deterministic and terminates. Note that we have assumed the four facts without proof.

Having informally shown that all five versions of CRL are deterministic and terminate, we now address the question of the complexity of the five algorithms. All of them follow exhaustive strategies, i.e. in each of their steps, they run through the entire problem space. For instance, in step 1 each of them generates the whole set of all possible value assignments $E_{\Phi}$. Although their exhaustiveness guarantees that they always compute the correct solution for any input, it is also the reason for their computational intractability. In fact, all five algorithms are intractable in the sense that the amount of time and space required to produce an output increases at a faster-than-polynomial rate as the input size grows.

The size of the input can be measured in terms of the number of propositional constants occurring in the input set $\Phi$, i.e. $|\text{pro}(\Phi)|$. In the first step, each of the algorithms requires us to compute the set $E_{\Phi}$, which has the size $2^{|\text{pro}(\Phi)|}$. Hence, the number of substeps it takes to carry out step 1 exceeds $2^{|\text{pro}(\Phi)|}$ and cannot be expressed in terms of a polynomial function of $|\text{pro}(\Phi)|$. The same is true for other steps.

For every algorithm of CRL we can express the amount of time required for running it in terms of an exponential function $2^{f(|\text{pro}(\Phi)|)}$, where $f(|\text{pro}(\Phi)|)$ is a polynomial function of the number of variables occurring in $\Phi$. Hence, applying any of the five algorithms to a given input constitutes a problem which is at least in the complexity class EXPTIME.

The fact that CRL is intractable makes the algorithm unsuitable for practical purposes. The algorithm is only applicable to inputs of very small sizes as its running time would otherwise be too long. However, the purpose of the chapter was to show that it is
possible in principle to compute potential solutions to a given conflict. This has been achieved by specifying the exhaustive algorithms CRL-1 to CRL-5.

Answering the question of how to design algorithms that can actually perform the tasks in polynomial-time, or how to produce close enough approximations to the solutions, exceeds the scope of the thesis.

**Example: Second Congo War**

In this section, we examine the goal conflict between the RCD and the Kabila government with regard to the integration of combatants of the RCD into the national Congolese army and apply each of the five algorithms to it. As a result, we get five sets of potential solutions to the conflict. We will evaluate the solutions in the final chapter of the thesis.

The goal conflict is given as follows.

\[ a_1: \text{Rassemblement Congolais pour la Democratie (RCD)} \]
\[ a_2: \text{Kabila government} \]
\[ p: \text{The combatants of the RCD are integrated into the Congolese national army.} \]
\[ q: \text{High ranking officers of the RCD are replaced.} \]
\[ C = \{Ga_1p, Ga_1\neg q, Ga_2p, Ga_2(p \supset q)\} \]
\[ C_g = \{p, p, \neg q, p \supset q\} \]

The Kabila government wants the combatants of the RCD to be integrated into the national Congolese army \( (Ga_1(p)) \), but only if they replace their high ranking officers \( (Ga_1(p \supset q)) \). The RCD also wants its combatants to be integrated into the national Congolese army \( (Ga_2(p)) \), but it does not want to replace its high ranking officers \( (Ga_2(\neg q)) \).
For the algorithm CRL-1, we take the set $C_g$ as an input. The following table displays all the relevant information required for running the algorithm.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p \lor q$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$p$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\neg q$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DEN</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>$E_\Phi e_1$</td>
<td>$p=1, q=1$</td>
<td>$e_2$</td>
<td>$p=1, q=0$</td>
<td>$e_3$</td>
</tr>
</tbody>
</table>

Table 20: Information for CRL-1

CRL-1 selects the two value assignments, $e_1$ and $e_2$, with minimal density and constructs, for each of them, a solution set $\Phi_1^*$ and $\Phi_2^*$. The two solution sets can be simplified by replacing their elements with logically equivalent, but simpler, formulae, and are then given by $\Phi_1^* = \{p, p \lor q, p, q \lor p\}$ and $\Phi_2^* = \{p, p \lor p, p, \neg q\}$, respectively.

Algorithm: CRL-1

| Input: $C_g = \{p, p, \neg q, p \lor q\}$ |
|---|---|---|---|---|
| Output: $\Phi_1^* = \{p, p \lor q, p, q \lor p\}$ and $\Phi_2^* = \{p, p \lor p, p, \neg q\}$ |

According to the solutions generated by CRL-1, either the RCD has to give up its claim of not replacing its high ranking officers as a consequence of solution $\Phi_1^*$, or the Kabila government has to give up its insistence that combatants of the RCD can only be integrated into the national Congolese army if the RCD replaces its high ranking officers as a consequence of accepting the solution $\Phi_2^*$.

For the algorithm CRL-2 we need to specify a subset of $\Delta \subseteq C_h$ such that all elements of $\Delta$ are contained in the solution set without being changed. For our example we choose $\Delta = \{\neg q\}$, i.e. we only look at solutions in which the high ranking officers of the RCD are not replaced.
Chapter 7  Generating Solutions

The relevant information for the algorithm is displayed in the following table.

<table>
<thead>
<tr>
<th>p</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>p ⊃ q</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>p</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Δ</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DEN</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>E₀</td>
<td>e₁</td>
<td>e₂</td>
<td>e₃</td>
<td>e₄</td>
</tr>
<tr>
<td>p=1 q=1</td>
<td>p=1 q=0</td>
<td>p=0 q=1</td>
<td>p=0 q=0</td>
<td></td>
</tr>
</tbody>
</table>

**Table 21: Information for CRL-2**

CRL-2 selects the value assignment $e_2$ as this is the only value assignment with minimal density among the value assignments $e_2$ and $e_4$ that make $Δ$ true. Then, the algorithm constructs a solution set $Φ_1^*$ based on $e_2$. The solution set is logically equivalent to the set $Φ_2^* = \{p, p ⊃ p, p, ¬q\}$.

```
Algorithm: CRL-2
Input: $C_g = \{p, p ⊃ q\}$, $Δ = \{¬q\}$
Output: $Φ_1^* = \{p, p ⊃ q, p \lor p\}$ and $Φ_2^* = \{p, p ⊃ p, p, ¬q\}$
```

According the solution generated by CRL-2, the Kabila government has to give up its claim that the combatants of the RCD can only be integrated into the national Congolese army if the RCD replaces its high ranking officers.

For the remaining three algorithms, we additionally need a finite, exhaustive collection of subsets of $C_g$ which reflects the origin of the claims, i.e. a distinction between claims made by the RCD and claims made by the Kabila government. The collection is given as follows: $C_{g1} = \{p, p ⊃ q\}$ and $C_{g2} = \{p, ¬q\}$.

We deal with the two algorithms CRL-3 and CRL-4 simultaneously. The following table displays all the information needed for running CRL-3 and CRL-4.
The algorithm CRL-3 minimises MAX by selecting the value assignments $e_1$, $e_2$ and $e_4$ and constructing solution sets based on these value assignments that are logically equivalent with the sets $\Phi_1^* = \{p, p \supset q, p, q \supset p\}$, $\Phi_2^* = \{p, p \supset p, p, \neg q\}$, and $\Phi_4^* = \{q \supset p, p \supset q, q \supset p, \neg q\}$.

Algorithm: CRL-3

Input: $C_9 = \{p, p, \neg q, p \supset q\}$, $C_{g1} = \{p, p \supset q\}$ $C_{g2} = \{p, \neg q\}$

Output: $\Phi_1^* = \{p, p \supset q, p, q \supset p\}$, $\Phi_2^* = \{p, p \supset p, p, \neg q\}$ and $\Phi_4^* = \{q \supset p, p \supset q, q \supset p, \neg q\}$.

According to the first solution generated by CRL-3, the RCD has to give up its goal not to replace its high ranking officers. According to the second solution, the Kabila government has to give up its claim that the combatants of the RCD can only be integrated into the national Congolese army if the RCD replaces its high ranking officers. According to the third solution, both parties have to give up their goal of integrating the combatants of the RCD into the national Congolese army.

CRL-4 minimises DIF by selecting the value assignment $e_4$ and constructing a solution set based on $e_4$ that is logically equivalent with the sets $\Phi_4^* = \{q \supset p, p \supset q, q \supset p, \neg q\}$. 

---

<table>
<thead>
<tr>
<th>$C_{g1}$</th>
<th>p</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>p $\supset q$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$C_{g2}$</td>
<td>p</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\neg q$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\text{den}_{g1}(e)$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\text{den}_{g2}(e)$</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>DIF</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$E_\Phi$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
<td></td>
</tr>
<tr>
<td>$p=1$</td>
<td>$q=1$</td>
<td>$p=1$</td>
<td>$q=0$</td>
<td>$p=0$</td>
<td>$q=1$</td>
</tr>
</tbody>
</table>

Table 22: Information for CRL-3 and CRL-4
Chapter 7 Generating Solutions

Algorithm: CRL-4
Input: \( C_g = \{ p, p, \neg q, p \supset q \}, \{ C_{g1} = \{ p, p \supset q \} \} \) 
Output: \( \Phi_4^* = \{ q \supset p, p \supset q, q \supset p, \neg q \} \)

The solution suggests that both parties should give up their goal of integrating the combatants of the RCD into the national Congolese army.

Finally, we provide a table containing the relevant information for running CRL-5.

<table>
<thead>
<tr>
<th>( C_{g1} )</th>
<th>( p )</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>( p \supset q )</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_{g2} )</td>
<td>( p )</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>( \neg q )</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>( \sum \text{conf}(C_{g1} \cap C_e) )</td>
<td>.75</td>
<td>.5</td>
<td>.25</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sum \text{conf}(C_{g2} \cap C_e) )</td>
<td>.5</td>
<td>1</td>
<td>0</td>
<td>.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIFCON</td>
<td>.25</td>
<td>.5</td>
<td>.25</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E_{\Phi} )</td>
<td>( e_1 )</td>
<td>( p=1 \ q=1 )</td>
<td>( e_2 )</td>
<td>( p=1 \ q=0 )</td>
<td>( e_3 )</td>
<td>( p=0 \ q=1 )</td>
<td>( e_4 )</td>
<td>( p=0 \ q=0 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 23: Information for CRL-5

The algorithm CRL-5 minimises DIFCON by selecting the value assignments \( e_1 \), \( e_3 \) and \( e_4 \) and constructing corresponding solution sets. The solution sets are logically equivalent to the sets \( \Phi_1^* = \{ p, p \supset q, p \supset q \} \), \( \Phi_3^* = \{ p \lor q, p \supset q, q \supset p, \neg q \} \) and \( \Phi_4^* = \{ q \supset p, p \supset q, q \supset p, \neg q \} \).

Algorithm: CRL-4-SOL
Input: \( C_g = \{ p, p, \neg q, p \supset q \}, \{ C_{g1} = \{ p, p \supset q \} \} \) 
Output: \( \Phi_4^* = \{ p, p \supset q, p \supset q \}, \Phi_3^* = \{ p \lor q, p \supset q, q \supset p, \neg q \} \) and \( \Phi_4^* = \{ q \supset p, p \supset q, q \supset p, \neg q \} \).

According to the first solution generated by CRL-5, the RCD has to give up its goal not to replace its high ranking officers. The second solution suggests that both parties should
give up their goal of integrating the combatants of the RCD into the national Congolese army and the RCD should give up its goal of not replacing its high ranking officers. If the third solution were accepted, both parties would have to give up the goal of integrating the combatants of the RCD into the national Congolese army.

In conclusion, we list again the input and output for the five algorithms if applied to the goal conflict between the RCD and the Kabila government.

<table>
<thead>
<tr>
<th>Input</th>
<th>Algorithm</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_p = {p, p \supset q, p, \neg q}$</td>
<td>CRL-1</td>
<td>$\Phi_1^* = {p, p \supset q, p, q \supset p}$ $\Phi_2^* = {p, p \supset p, p, \neg q}$</td>
</tr>
<tr>
<td>$C_p = {p, p \supset q, p, \neg q}$ $\Delta = {\neg q}$</td>
<td>CRL-2</td>
<td>$\Phi_2^* = {p, p \supset p, p, \neg q}$</td>
</tr>
<tr>
<td>$C_p = {p, p \supset q, p, \neg q}$ $C_{q_1} = {p, p \supset q}$ $C_{q_2} = {p, \neg q}$</td>
<td>CRL-3</td>
<td>$\Phi_1^* = {p, p \supset q, p, q \supset p}$ $\Phi_2^* = {p, p \supset p, p, \neg q}$ $\Phi_3^* = {q \supset p, p \supset q, p \supset p, \neg q}$</td>
</tr>
<tr>
<td>$C_p = {p, p \supset q, p, \neg q}$ $C_{q_1} = {p, p \supset q}$ $C_{q_2} = {p, \neg q}$</td>
<td>CRL-4</td>
<td>$\Phi_4^* = {q \supset p, p \supset q, q \supset p, \neg q}$</td>
</tr>
<tr>
<td>$C_p = {p, p \supset q, p, \neg q}$ $C_{q_1} = {p, p \supset q}$ $C_{q_2} = {p, \neg q}$</td>
<td>CRL-5</td>
<td>$\Phi_4^* = {q \supset p, p \supset q, q \supset p, \neg q}$</td>
</tr>
</tbody>
</table>

Table 24: Input and Output of the Five Algorithms

7.6 Summary

The aim of this chapter was to specify five algorithms of CRL which generate potential solution sets to conflicts previously modelled by CML. This has been achieved to the extent that we have provided definitions for the five algorithms CRL-1, CRL-2, CRL-3, CRL-4, and CRL-5, and have shown how each of them computes a set of consistent sets to a given input set of propositional formulae. The consistent sets computed by the algorithms can be suggested to the conflicting parties as possible ways to resolve their
conflict. Each of the five algorithms has been designed according to a particular principle which is realised in the type of solutions generated by the algorithm.

It has been informally shown that the algorithms are all computable, i.e. that they halt after a finite number of steps for any given input. Furthermore, we have shown that the algorithms are intractable and are at least in the complexity class EXPTIME.

Having defined resolution algorithms, we can now generate possible solutions to conflicts expressed in terms of consistent sets of propositional formulae. This gives us a constructive tool to address the question of how to solve conflicts. Computing solutions to a conflict constitutes the sixth step in the process of modelling and resolving a conflict.

In the following chapter, we will interpret the types of solutions generated by the five algorithms and show how they provide minimally invasive solutions, compromises, or legal solutions to a conflict.
CHAPTER 8

Evaluating Solutions

The Resolution Principles of Conflict Resolution Logic

8.1 Introduction

The aim of chapter 8 is to interpret the solutions generated by the five algorithms of CRL. As each of the five algorithms realises a different principle when generating a solution set to a given conflict, the solution sets satisfy different conditions of conflict resolution. Some of the solutions have the property of requiring only a minimal amount of change from the conflicting parties. Other solutions are guaranteed to be compatible with certain legal or moral norms, or provide compromise solutions.

As a result of the chapter we will be able to give answers to the following questions.

- How can minimally invasive solutions to a conflict be generated and how do they minimise collective and individual change, respectively?
- How can solutions satisfying certain, specified, legal norms be generated?
- How can compromise solutions be generated and how do they balance individual change and individual non-change?
• How can potential solutions to sub-conflicts of the Second Congo War be evaluated?

An outline of the chapter is as follows: In section 8.2 we show why solutions generated by the two algorithms CRL-1 and CRL-3 are minimally invasive solutions. In particular, we explain how solutions generated by the first algorithm minimise collective change (8.2.1), and solutions generated by the second algorithm minimise individual change (8.2.2). In section 8.3, we look at the type of solutions generated by CRL-2 and show how they are guaranteed to be compatible with norms contained in the set Δ. In section 8.4, we analyse the solutions generated by the two algorithms CRL-4 and CRL-5, both of which produce compromise solutions. First, we look at the type of compromises generated by CRL-4 (8.4.1), and then at those generated by CRL-5 (8.4.2). Finally, we evaluate the solutions generated for the sub-conflicts of the Second Congo War.

8.2 Minimally Invasive Solutions

Solutions produced by the algorithms CRL-1 and CRL-3 can be characterised as minimally invasive solutions. We first look at the type of solutions generated by CRL-1 and then at those generated by CRL-3.

8.2.1 Solutions Minimising Collective Change

CRL-1 takes as an input a finite set Φ of propositional formulae and outputs a set \{Φ₁\*, Φ₂\*, … Φₙ\*\} of consistent sets Φᵢ\*. In the case of CRL-1, each Φᵢ\* has the property of being as close as possible at the original set Φ, i.e. any other consistent set obtained from Φ by replacing its elements by formulae with a lower degree of potential conflict power
has less elements in common with \( \Phi \) than \( \Phi_1^* \). As formulae express the content of propositional attitudes held by the agents, the overall number of propositional attitudes that need to be changed if a set generated by CRL-1 is taken as a solution to a conflict is minimal. In other words: proposing a solution generated by CRL-1 to the parties involved in a conflict requires them to change their propositional attitudes only as much as absolutely necessary. Thus, we call solutions generated by CRL-1 minimally invasive solutions.

The algorithm produces this kind of solution because it selects value assignments with minimal density and then alters just those formulae of \( \Phi \) which are not satisfied by the selected value assignments. If a value assignment has minimal density with respect to a set \( \Phi \), it satisfies a maximal number of elements of \( \Phi \) because the density of \( e \) is defined as the number of formulae not satisfied by \( e \).

To illustrate the property of minimal invasiveness of the solutions generated by the algorithm CRL-1, we look at the set \( \Phi = \{p \lor q, p, q, p \land \neg q\} \). The following truth table for \( \Phi \) also shows the absolute densities of the four possible value assignments \( e_1, e_2, e_3, \) and \( e_4 \) with respect to \( \Phi \).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p \lor q )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>( p )</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( q )</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>( p \land \neg q )</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DEN</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>( e_1 )</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>( e_2 )</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>( e_3 )</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>( e_4 )</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 25: Solutions of CRL-1

CRL-1 picks the two value assignments with minimal density, \( e_1 \) and \( e_2 \), and generates the two corresponding solutions \( \Phi_1^* = \{p \lor q \lor (p \land q), p \lor (p \land q), q \lor (p \land q), (p \land \neg q) \lor \}

\
\( (p \land q) \) and \( \Phi_2^* = \{ p \lor q \lor (p \land \neg q), p \lor (p \land \neg q), q \lor (p \land \neg q), (p \land \neg q) \lor (p \land \neg q) \} \). Simplifying the formulae by replacing them with simpler formulae that are logically equivalent gives us the sets \( \Phi_1^* = \{ p \lor q, p, q, p \} \) and \( \Phi_2^* = \{ p \lor q, p \lor q, q, p \land \neg q \} \).

Both \( \Phi_1^* \) and \( \Phi_2^* \) are consistent. In \( \Phi_1^* \), only the formula \( p \land \neg q \) has been replaced by the weaker formula \( p \), whereas in \( \Phi_2^* \), the formula \( q \) has been replaced by the weaker formula \( p \lor q \). Note, that the density \( \text{den}_{\Phi_1}(e_1) = \text{den}_{\Phi_2}(e_2) = 1 \) is equal to the degree of inconsistency of \( \Phi \), Inc-1(\( \Phi \)).

Any solution based on other value assignments than \( e_1 \) and \( e_2 \) results in a higher number of formulae required to be changed. For instance, the solution based on \( e_3 \), \{ \( p \lor q, p \lor q, q, p \equiv \neg q \} \), contains two formulae different from the formulae contained in \( \Phi \). Solutions based on \( e_4 \) require three formulae to be changed.

The example shows that solutions generated by the algorithm CRL-1 require as little overall change as possible. However, these kinds of solutions do not take into account which of the parties is required to change its propositional attitudes. Although it is guaranteed that only a minimal number of the original claims must be changed, the required changes might be distributed unequally among the parties. For instance, one party might be required to change all its propositional attitudes, whereas another party might not be required to change any of its attitudes. This deficiency is accommodated for in solutions generated by CRL-3.

With respect to the AGM model, solutions generated by the algorithm CRL-1 are closely related, though not equivalent, to the result of applying an update consolidation operator \( ! \) based on a maxichoice contraction operator \( \div \) to a belief base \( \Lambda \).
8.2.2 Solutions Minimising Individual Change

CRL-3 takes as input not only the set $\Phi$, but also a finite, exhaustive collection $\{\Phi_i\}_{i \in I}$ of subsets $\Phi_i \subseteq \Phi$. $\{\Phi_i\}_{i \in I}$ contains information about the origins of the propositional attitudes whose contents are collected in $\Phi$. Each cell $\Phi_i$ represents the content of the propositional attitudes held by exactly one agent $a_i$. For instance, the collection $\{\Phi_1 = \{p\}, \Phi_2 = \{q\}, \Phi_3 = \{p \land \neg q\}\}$ represents a situation in which an agent $a_1$ has a propositional attitude towards $p$, an agent $a_2$ has a propositional attitude towards $q$, and an agent $a_3$ has an attitude towards $p \land \neg q$.

Solutions generated by CRL-3 can be characterised as minimally invasive for individuals. They have the property that the number of propositional attitudes required to be changed by any individual agent involved in the conflict is kept to a minimum. In any set which is not a solution generated by CRL-3, there is an agent who has to change more propositional attitudes than the agent who has to change the most attitudes according to any solution generated by CRL-3.

This can be illustrated by looking at the set $\Phi = \{-p \lor q, -p, -p \equiv q, -p \lor q, q, p \land q, p \lor q, p \land -q\}$ and the collection $\{\Phi_1 = \{-p \lor q, -p, -p \equiv q\}, \Phi_2 = \{-p \lor q, q, p \land q\}, \Phi_3 = \{p \lor -q, p, p \land -q\}\}$. The truth table of this set is expressed by the table below. The table also shows the individual densities for each of the three agents, $d_{\Phi_1}(e), d_{\Phi_2}(e)$ and $d_{\Phi_3}(e)$, the densities with respect to the whole set $\Phi$, $d_{\Phi}(e)$, and the set MAX of maxima of the set $\{d_{\Phi_1}(e), d_{\Phi_2}(e), d_{\Phi_3}(e)\}$. 

<table>
<thead>
<tr>
<th>$\Phi_i$</th>
<th>$d_{\Phi_1}(e)$</th>
<th>$d_{\Phi_2}(e)$</th>
<th>$d_{\Phi_3}(e)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$\Phi_2$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$\Phi_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

| $d_{\Phi}(e)$ | 0.5 |

| MAX of maxima | 0.5 |
The algorithm first calculates the densities $d_{\Phi_1}(e)$, $d_{\Phi_2}(e)$, and $d_{\Phi_3}(e)$, and then identifies the respective maxima of the three densities. The maxima are displayed in the row MAX. The algorithm then chooses the minimum of MAX, $e_4$, and generates a solution set based on $e_4$ that is logically equivalent to the simplified set $\Phi^* = \{\neg p \lor \neg q, \neg p, \neg p \lor \neg q, \neg p \lor q, \neg p \lor q, \neg p \lor q, p \equiv q, p \lor \neg q, p \lor \neg q, \neg q\}$.

The solution generated by CRL-3 requires the agent $a_1$ to replace his attitude towards $\neg p \equiv q$ by weakening its content to $\neg p \lor \neg q$, and it requires the agents $a_2$ and $a_3$ to change two of their attitudes each. The required changes for $a_2$ consist of replacing the formula $q$ by $\neg p \lor q$ and the formula $p \land q$ by $p \equiv q$. $a_3$ has to replace $p$ with $p \lor \neg q$ and $p \land \neg q$ by $\neg q$. Altogether, the maximum number of attitudes that any of the three agents need to change, according to the solution $\Phi^*$, is two.
Solutions constructed on the basis of $e_1$, $e_2$, or $e_3$ require at least one of the agents to change all three of his propositional attitudes. For instance, in the solution based on $e_1$, the agent $a_1$ has to change all of his three attitudes, whereas the agent $a_2$ is not required to change any of his attitudes.

If the simpler algorithm CRL-1 were applied to $\Phi$, it would pick the three value assignments $e_1$, $e_2$ and $e_3$, and generate solutions based on these value assignments as the overall density of $\Phi$ is minimal in each of them.

8.3 Legal Solutions

Solutions generated by the algorithm CRL-2 can be interpreted as solutions that are compatible with certain, pre-defined, moral or legal norms. The norms, which are represented by a set $\Delta$ of propositional formulae, can be moral principles, such as principles of justice, non-violence, no-harm etc., or legal principles, such as certain legal conventions, human rights declarations, civil rights, elements of the international humanitarian law, etc. The solutions generated by CRL-2 are guaranteed to be compatible with all the previously designated norms contained in $\Delta$.

In principle, we can include in $\Delta$ any proposition $p$ for which we want to be guaranteed that it is represented in the solution set of a conflict. Apart from norms, we might also want to include certain facts or trivially true beliefs in $\Delta$ in order to avoid solutions which contradict obvious facts or trivially true beliefs shared by all parties. However, in the following, we focus on legal or moral norms.

If a formula $\varphi$ is contained in $\Delta$, it is assumed that it ought to be the case that $\varphi$, i.e. that we are obliged to restrict our solutions to only those solutions in which the truth of $\varphi$ is
not contested. As a minimal condition, $\Delta$ itself must be consistent. Otherwise, we would be forced to automatically violate at least some of the norms contained in $\Delta$. Furthermore, $\Delta$ is assumed to be a subset of $\Phi$, so it is finite as $\Phi$ is finite.

The algorithm CRL-2 takes a finite consistent set $\Phi$ of propositional formulae together with a consistent subset $\Delta \subseteq \Phi$ of norms as an input. It then computes a set $\{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$ of consistent sets $\Phi_i^*$ such that $\Delta$ is contained in all of the solution sets $\Phi_i^*$, i.e. $\Delta \subseteq \Phi_i^*$ for all $\Phi_i^* \in \{\Phi_1^*, \Phi_2^*, \ldots \Phi_n^*\}$.

CRL-2 systematically looks at only those value assignments which satisfy $\Delta$, and then proceeds according to the density minimising principle of the algorithm CRL-1. The following example illustrates how the algorithm works and shows that the solutions generated by it all satisfy the norms contained in $\Delta$. We provide the truth table for the set $\Phi = \{p \equiv q, q, \neg p \land q, \neg q, p \lor \neg q\}$, including the rows representing the formulae contained in the set of norms $\Delta = \{\neg q, p \lor \neg q\}$.

<table>
<thead>
<tr>
<th>$p \equiv q$</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$\neg p \land q$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$\Delta$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$p \lor \neg q$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$\text{DEN}$</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>$E_\Phi$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
</tr>
<tr>
<td>$p=1 q=1$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
</tr>
<tr>
<td>$p=1 q=0$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
</tr>
<tr>
<td>$p=0 q=1$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
</tr>
<tr>
<td>$p=0 q=0$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
</tr>
</tbody>
</table>

Table 27: Solutions of CRL-2

As the table shows, only the value assignments $e_2$ and $e_4$ satisfy the formulae contained in $\Delta$. These are picked by the algorithm first. Their respective densities are then compared and the one with the lower density, $e_4$, is chosen. The corresponding solution set,
generated by CRL-2 is the set \( \Phi^* = \{ p \equiv q, q \lor \neg p, \neg p, \neg q, p \lor \neg q \} \) which contains as a subset the norms expressed in \( \Delta \).

Note that the algorithm CRL-1 would have produced two solution sets based on the value assignments \( e_1 \) and \( e_4 \), but one of the solutions would not have been compatible with the norms contained in \( \Delta \).

With respect to the AGM model, the algorithm CRL-2 is closely related to the operation of screened revision. The set \( \Delta \) of pre-defined norms directly corresponds to the set \( X \) of potential core beliefs, which is also guaranteed to be part of the revised set \( K#p \) of screened revision. Moreover, the consistency condition applies to both \( \Delta \) and \( X \).

### 8.4 Compromises

The two algorithms CRL-4 and CRL-5 both generate solutions to a conflict which can be interpreted as compromise solutions. CRL-4 produces compromises in the sense that the difference between the number of satisfied claims for the agent with the highest number of satisfied claims and the number of satisfied claims for the agent with the lowest number of satisfied claims is kept minimal. In other words: the number of satisfied claims for each agent is as equal as possible. CRL-5 produces more sophisticated compromises than CRL-4. Here, not only is the number of satisfied claims for each agent taken into account, but the sum of the potential conflict powers of these claims is also factored in. CRL-5 can, therefore, be regarded as a weighted version of CRL-4.

Both algorithms take a set \( \Phi \) of propositional formulae and a finite, exhaustive collection \( \{ \Phi_i \}_{i \in I} \) of subsets \( \Phi_i \subseteq \Phi \). \( \{ \Phi_i \}_{i \in I} \) as input. As in the case of CRL-3, the collection
reflects the different propositional attitudes held by the various agents involved in the conflict. A cell $\Phi_i$ contains the content of the propositional attitudes held by the agent $a_i$.

We first describe and illustrate the type of solutions generated by CRL-4, and then we characterise the second compromise algorithm CRL-5.

### 8.4.1 Solutions Equalling Individual Change

Solutions generated by CRL-4 have the property that every agent involved in the conflict has approximately the same number of propositional attitudes represented in the solution. The number of satisfied propositional attitudes for the agent with the maximal amount of satisfied attitudes is as close as possible to the number of satisfied propositional attitudes for the agent with the minimal number of satisfied attitudes. This makes sure that the solution is fair to the extent that no agent is forced to give up significantly more propositional attitudes than the other agents.

The algorithm, and the type of solutions it generates, can be illustrated by looking at the set $\Phi = \{ \neg p \lor q, \neg p \equiv q, \neg p \lor \neg q, p \equiv q, p \lor \neg q, \neg p, \neg q \}$ and the collection $\{ \Phi_1 = \{ \neg p \lor q, \neg p \equiv q, \neg p \lor \neg q \}, \Phi_2 = \{ p \equiv q, p \lor \neg q, \neg p, \neg q \} \}$. The truth table is shown below. The table also shows the densities $d_{\Phi_1}(e)$ and $d_{\Phi_2}(e)$ as well as the differences between the maximal density and the minimal density for every $e \in E_{pq}$. 
CRL-4 first computes the densities with respect to the two sets $\Phi_1$ and $\Phi_2$. Then, it identifies the maximum and minimum of these densities. As there are only the two cells, $\Phi_1$ and $\Phi_2$, in the example, the row DIF just displays the differences between $d_{\Phi_1}(e)$ and $d_{\Phi_2}(e)$ for every $e \in E_{\Phi}$. The algorithm then picks the value assignment with the smallest difference, which is $e_1$ in the example, and generates a solution set based on $e_1$. The solution set is logically equivalent to the simpler set $\Phi^* = \{p \lor q, p \lor \neg p, \neg p \lor q, p \lor q\}$, which is satisfied by $e_1$. According to the solution set, each of the two agents has to change two of his original propositional attitudes. The agent $a_1$ has to replace his attitudes towards $\neg p \equiv q$ and $\neg p \lor \neg q$ by attitudes towards $p \lor q$ and $p \lor \neg p$. In the second case he has to completely abandon his original attitude as he is required to replace it by a tautological attitude. The agent $a_2$ is required to replace his attitudes towards $\neg p$ and $\neg q$ by attitudes towards $\neg p \lor q$ and $p \lor \neg q$.

The example shows that $e_1$ is the only value assignment which forces the two agents to change an equal amount of attitudes. In a solution based on $e_2$, for example, $a_1$ has to change only one attitude, whereas $a_2$ has to change two attitudes. Hence, the solution

<table>
<thead>
<tr>
<th>$\Phi_1$</th>
<th>$\neg p \lor q$</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\neg p \equiv q$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$\neg p \lor \neg q$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Phi_2$</th>
<th>$p \equiv q$</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p \lor \neg q$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\neg p$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\neg q$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

| $d_{\Phi_1}(e)$ | 2 | 1 | 0 | 1 |
| $d_{\Phi_2}(e)$ | 2 | 2 | 3 | 0 |
| DIF | 0 | 1 | 3 | 1 |

<table>
<thead>
<tr>
<th>$E_{\Phi}$</th>
<th>$e_1$</th>
<th>$e_2$</th>
<th>$e_3$</th>
<th>$e_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p=1$ $q=1$</td>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
</tr>
<tr>
<td>$p=1$ $q=0$</td>
<td>$p=0$ $q=1$</td>
<td>$p=0$ $q=0$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 28: Solutions of CRL-4
generated by CRL-4 constitutes a situation in which every party has to compromise to the same extent, and, therefore, represents a compromise solution. Note that the solution generated by CRL-4 is actually maximally invasive as the overall amount of changes is lower in each of the other solutions. However, it is the solution which distributes the amount of change most equally among the agents.

### 8.4.2 Solutions Equalling Individual Non-Change

Now we turn to the algorithm CRL-5. Before we illustrate it with an example, we describe a deficiency of the algorithm CRL-4 which will be dealt with by CRL-5.

In the solution proposed by CRL-4, \( a_1 \) does not have to change his attitude towards \( \neg p \lor q \) and \( a_2 \) is not required to change his attitudes towards the propositions \( p \equiv q \) and \( p \lor \neg q \). This can be regarded as unfair. Especially, when it is taken into account that the claim expressed by the formula \( \neg p \lor q \) is much weaker than the claim expressed by the two formulae \( p \equiv q \) and \( p \lor \neg q \), and, hence, \( a_2 \) is able to satisfy much stronger claims than \( a_1 \) according to the solution generated by CRL-4. The strength of a claim can be measured by its potential conflict power. It is immediately obvious that \( a_1 \) can satisfy weaker claims than \( a_2 \) when we look at the potential conflict power of the formulae involved: \( \text{conf}(\neg p \lor q) = .25 < \text{conf}(p \equiv q) + \text{conf}(p \lor \neg q) = .5 + .25 = .75 \).

Solutions generated by CRL-5 represent compromises in which the strength of the claims satisfied by each individual party is as equal as possible. This can be realised by keeping the difference between the strength of the satisfied claims for the party with the strongest satisfied claims and the strength of the satisfied claims for the party with the weakest satisfied claims minimal.
As the strength of a party’s claims can be measured by the sum of the potential conflict powers of the claims, the algorithm first calculates the conflict powers of the formulae that are satisfied by a certain value assignment \( e \). It calculates this sum for each party and each \( e \in E_{\phi} \) individually. Then, it proceeds in a similar fashion to the algorithm CRL-4, i.e. it builds the differences between the maximum and the minimum of the sums and identifies the value assignment with the minimal difference.

Applying CRL-5 to the input set \( \Phi = \{ \neg p \lor q, \neg p \equiv q, \neg p \lor \neg q, p \equiv q, p \lor \neg q, \neg p, \neg q \} \) and the collection \( \{ \Phi_1 = \{ \neg p \lor q, \neg p \equiv q, \neg p \lor \neg q \}, \Phi_2 = \{ p \equiv q, p \lor \neg q, \neg p, \neg q \} \} \) gives us the following truth table. The table also displays the sums \( \sum_{\forall \in \Phi_1 \cap \Phi} \text{conf} (\forall) \) and \( \sum_{\forall \in \Phi_1 \cap \Phi_2} \text{conf} (\forall) \), where \( \Phi_e \) is the set of formulae contained in \( \Phi \) which are satisfied by \( e \).

| \( \Phi_1 \) | \( \neg p \lor q \) | 1 | 0 | 1 | 1 |
|\( \neg p \equiv q \) | 0 | 1 | 1 | 0 |
| \( \neg p \lor \neg q \) | 0 | 1 | 1 | 1 |
| \( p \equiv q \) | 1 | 0 | 0 | 1 |
| \( p \lor \neg q \) | 1 | 1 | 0 | 1 |
| \( \neg p \equiv q \) | 0 | 0 | 1 | 1 |
| \( \neg q \equiv q \) | 0 | 1 | 0 | 1 |

| \( \sum_{\forall \in \Phi_1 \cap \Phi} \text{conf} (\forall) \) | .5 | .75 | 1 | .5 |
| \( \sum_{\forall \in \Phi_1 \cap \Phi_2} \text{conf} (\forall) \) | .75 | .75 | .5 | 1.75 |
| Max - Min | .25 | 0 | .5 | 1.25 |
| \( E_{\Phi} \) | \( e_1 \) | p=1 q=1 | \( e_2 \) | p=1 q=0 | \( e_3 \) | p=0 q=1 | \( e_4 \) | p=0 q=0 |

Table 29: Solutions of CRL-5

The solution generated by CRL-5 on the basis of the value assignment function \( e_2 \) is a set which is logically equivalent to the simpler set \( \Phi^* = \{ \neg p \lor p, \neg p \equiv q, \neg p \lor \neg q, p \lor \neg q, p \lor \neg q, \neg p \lor \neg q, \neg q \} \).
If the solution generated by CRL-5 were accepted by the two agents \( a_1 \) and \( a_2 \), \( a_1 \) would have to replace his attitude towards \( \neg p \lor q \) by an attitude towards \( \neg p \lor p \) whereas \( a_2 \) would have to replace his attitudes towards \( p \equiv q \) and \( \neg p \lor \neg q \) by attitudes towards \( p \lor \neg q \) and \( \neg p \lor \neg q \). At the same time, \( a_1 \) would be able to maintain without change claims with a combined potential conflict power of .75, i.e. the two propositional attitudes towards \( \neg p \equiv q \) and \( \neg p \lor \neg q \), whereas \( a_2 \) would be able to maintain the two attitudes towards \( p \lor \neg q \) and \( \neg p \lor \neg q \) with a combined potential conflict value also of .75. Hence, the strength of the claims that can be satisfied without change is equal for the two parties.

In conclusion, we can observe that both CRL-4 and CRL-5 generate compromise solutions. CRL-4 focuses on equality regarding the number of changes necessary to be made by the agents involved in the conflict, whereas CRL-5 makes sure that equality is guaranteed with respect to the strength of the claims that do not require change by the agents.

**Example: Second Congo War**

In this section we evaluate the solutions generated by the five algorithms for the goal conflict between the RCD and the Kabila government. For each of the solutions, we give a justification of why it is a good solution. In conclusion, we make some comments about the solutions in general and how they can be seen as a step towards finding a sustainable solution to a conflict like the Second Congo War as a whole.

The justification for choosing a solution generated by CRL-1 is its closeness to the original conflict. As a consequence, the agents have to change their attitudes only to a minimal extent. The solution \( \Phi_1^* = \{p, p \supset q, q \supset p\} \), for instance, only requires the
RCD to abandon its goal not to replace its high ranking officers. This is the only change required according to the solution. In any other solution, except the two generated by CRL-1, more changes are required, i.e. both the RCD and the Kabila government have to change their goals.

The motivation for choosing the solution \( \Phi_2^* = \{p, p \supset p, p, \neg q\} \), generated by CRL-2, is that it is the closest solution to the original conflict which also guarantees that the high ranking officers of the RCD are not required to be replaced. In any other solution in which the high ranking officers of the RCD are not required to be replaced, more than one change is required, whereas \( \Phi_2^* \) only requires the Kabila government to abandon its claim regarding the condition for the combatants of the RCD to be integrated into the national Congolese army. In the example, we chose \( \Delta = \{\neg q\} \) arbitrarily. A possible justification for that choice could be that dismissing high ranking officers of the RCD is considered morally wrong as they have achievements equal to those of the officers in the national Congolese army and should, therefore, be treated equally. There might also be legal reasons why the officers of the RCD cannot be dismissed. Obviously, other moral claims, even the opposite claim that they should be dismissed, might be justifiable. This would alter the set \( \Delta \) and, hence, the solutions generated by CRL-2.

Solutions generated by CRL-3 have the property that the individual change required from each agent is kept minimal. For instance, according to the solution \( \Phi_1^* = \{p, p \supset q, p, q \supset p\} \), the maximal amount of claim changes required from either of the two agents is one. The solution only requires the agent RCD to give up its goal of not having to replace its high ranking officers. In contrast, according to the solution \( \Phi_3^* = \{p \lor q, p \supset q, q \supset p, \neg q\} \) the agent RCD has to abandon both of its claims.
The justification for the solution \( \Phi^*_4 = \{ q \supset p, p \supset q, q \supset p, \neg q \} \), generated by CRL-4, can be described as follows: If \( \Phi^*_4 \) were accepted by the RCD and the Kabila government as a solution to their conflict, each of the two agents would have to give up one of their two goals and could realise the other one. Both the RCD and the Kabila government would have to abandon the goal of integrating the combatants of the RCD into the national Congolese army. However, the Kabila government could realise its goal of having combatants of the RCD integrated into the national Congolese army only if the RCD replaces its high ranking officers, and the RCD would not have to replace its high ranking officers. Thus, the solution \( \Phi^*_4 \) represents a fair solution as each party would have to give up, and could realise, the same amount of goals.

Solutions generated by CRL-5 are good because they realise another principle of justice. Here, the RCD and the Kabila government can realise goals of similar strength. For instance, in the solution \( \Phi^*_3 = \{ p \lor q, p \supset q, q \supset p, \neg q \} \), the Kabila government can realise only the weak goal according to which combatants of the RCD are only integrated into the national Congolese army if the RCD replaces its high ranking officers. This goal has a potential conflict power of .25. The RCD cannot realise any of its goals. In other words, both parties are strongly encouraged to abandon their goals. Similar justifications can be given for the other solutions generated by CRL-5.

Altogether, the solutions generated by the five algorithms seem trivial. Either they suggest obvious solutions, like the RCD having to give up its goal of not having to replace its high ranking officers (\( \Phi^*_1 \)) or the Kabila government having to give up its condition for integrating the combatants of the RCD into the national Congolese army (\( \Phi^*_2 \)), or they seem counterintuitive as in the case of \( \Phi^*_3 \), where the parties are required to abandon the goal of integrating the RCD combatants into the national Congolese army, even though this is a goal shared by both parties.
The reason for the triviality of the solutions generated by the algorithms is the small size of the input set. In small conflicts that do not include many elements (goals, beliefs, norms, etc.) there are not many potential solutions available, and those that are available seem trivial. However, if the number of conflict elements increases, the number of possible solutions increases too. This opens the door for a large number of bargaining situations in which the different goals, beliefs, norms, etc. can be traded with each other. For these kinds of conflicts, the algorithms of CRL can produce various package solutions which satisfy the underlying principle of the respective algorithm.

If we compare the solutions generated for the small goal conflict, we observe that some solutions are picked by more algorithms than others. For instance, the solution $\Phi_3^*$ is picked only once. This can be interpreted as a hint that $\Phi_3^*$ is not an appropriate solution to the conflict. Indeed, $\Phi_3^*$ seems also to be the intuitively least convincing solution. We also observe that the solutions $\Phi_1^*$ and $\Phi_2^*$ are picked more frequently by the algorithms CRL-1, CRL-2, and CRL-3, whereas the solutions $\Phi_3^*$ and $\Phi_4^*$ are picked more frequently by CRL-4 and CRL-5. This can be explained by the fact that the former three algorithms focus on minimising change, whereas the latter two focus on principles of justice.

### 8.5 Summary

The aim of chapter 8 was to provide an interpretation of the solutions generated by the five algorithms of CRL. This has been achieved to the extent that we have explained why solutions generated by the algorithms CRL-1 and CRL-3 can be regarded as minimally invasive solutions, solutions generated by CRL-2 as legal solutions, and solutions generated by CRL-4 and CRL-5 as compromise solutions. We have illustrated each type
Chapter 8  Evaluating Solutions

of solution by looking at the solution sets generated by each of the five algorithms when applied to an inconsistent set of propositional formulae.

Having provided an interpretation for the types of solutions, we are now in the position of being able to assess the solutions that we have generated to a conflict and choose the solution which best fits the situation. Furthermore, we can look at solutions that have worked for previous conflicts and identify their type. Interpreting the solutions for a conflict can be seen as the last step in the process of modelling and resolving a conflict.
In conclusion, we critically assess the work presented in this thesis by providing answers to the following questions.

- What has been achieved in the thesis?
- What are the main results of the thesis?
- How can the theoretical framework, developed in this thesis, be used in practice?
- What questions have been raised in the thesis that would be interesting for further investigations?

The aim of the thesis was to show how formal logic can be applied to the issues of conflict modelling and conflict resolution. After reviewing current theories of conflict (chapter 1), we developed the syntax (chapter 2), semantics (chapter 3), and axiomatics (chapter 4) of the logical system CML which is intended to model conflicts by combining ideas from classical propositional and first-order logic, alethic modal logic, branching-time temporal logic, and a number of propositional attitude operators. We have demonstrated how the different components and concepts of CML can be applied to real
conflicts by reconstructing aspects of the Second Congo War within the framework of CML.

We have established a classification scheme based on CML in order to identify the specific types of sub-conflicts constituting a conflict (chapter 5) and we have introduced two numeric measures for assessing the role conflict elements play within a conflict, as well as the depth of conflicts in terms of how easy it is to resolve them (chapter 6). In the last two chapters we have developed CRL, a set of five algorithms that generates potential solutions to conflicts previously modelled by CML. We have stated the specific steps through which the algorithms run (chapter 7) and explored the type of solutions generated by each of the algorithms (chapter 8).

As the main theoretical result of the thesis, we have proven a completeness theorem for the logical system CML. Other theoretical results of the thesis include the definition of the potential conflict power of propositional formulae and the associated equivalence theorem for the semantic and syntactic versions of it, the definition of the degree of inconsistency of sets of propositional formulae, and the definition of the algorithms transforming inconsistent sets of propositional formulae into consistent sets by reducing the potential conflict power of their elements.

A peculiarity of the semantics of the temporal fragment of CML is the property that states and time points are represented by two different semantic entities. In contrast to most branching-time temporal logics, in which states and time points are the same entity, in t-CML we can differentiate between the different modal states a conflict can possibly be in at a single time point.

We now turn to the question of how the theoretical framework of CML and CRL can be used in practice. In general, conflict scholars who are inclined to systematic research can
use CML to represent particular conflict cases in a unified and systematic way that straightforwardly displays the logical structure of the case. CRL can be used to generate solutions to a conflict which can then serve as a source of inspiration for resolving the conflict. Solutions generated by CRL can, obviously, not completely solve a conflict. However, they can be used as a starting point for finding a sustainable solution accepted by all parties. The advantage of solutions generated by CRL is that the parties are forced to make their beliefs, goals, norms, and emotions explicit. The solutions generated by CRL only reflect what has been fed as an input to the algorithms. The explicit character of the approach, combined with the feature that the type of solution, i.e. minimal invasive solutions, legal solutions, or compromise solutions, can be chosen by choosing the respective algorithm, provides a high degree of controllability. Conflict scholars working with the framework can control the whole process of modelling a conflict case, generating solutions to it, and evaluating the solutions. As a consequence, they can ‘play around’ with the model, i.e. they can generate different types of solutions, they can observe how different representations of a conflict produce different solutions as an output, and they can weigh the different aspects of a conflict by altering the input.

The model can either be used descriptively or prescriptively. In its descriptive use, conflict scholars can describe conflicts with CML and classify them according to the classification scheme. Furthermore, they can generate solutions and compare these generated solutions with the actual solutions that were realised in the conflict. In its prescriptive use, conflict scholars can use the generated solutions as an inspiration for solving the conflict. They can suggest the solutions to the parties and explain why they are good solutions, i.e. they can explain what each party gains from the solution and what each party has to give up. Choosing a certain type of solution makes sure that the underlying principle of the respective algorithm is realised.
Conclusion

Practitioners can use the model as a fact-finding tool. They can collect data about a conflict, either in co-operation with the parties or through independent observation, using the framework of CML. CML provides a guideline to the question of which data are of interest and how the data can be represented in an efficient way.

Although our logic-based model as a whole, in addition to particular components of it, provides a number of different possible uses for conflict scholars and practitioners, we will summarise the way it can be used to model and resolve conflicts by going through a list of steps that establish a systematic method of modelling and resolving a conflict on the basis of CML and CRL. In the steps we assume that a particular conflict is to be modelled and resolved.

Step 1: Collect data about the conflict. In particular data concerning the issues the conflict is about, the agents involved in the conflict, as well as their beliefs, goals, norms, and emotions.

Step 2: Track the temporal occurrence of the collected data, in particular the dates at which the agents’ beliefs, goals, norms, and emotions appear.

Step 3: Describe the conflict in the language of CML by translating the collected data into CML-formulae using the syntax of CML.

Step 4: Identify subsets of the set of all formulae that represents the conflict and check their consistency/satisfiability using the axiomatics/semantics of CML.

Step 5: Identify inconsistent/unsatisfiable subsets of the set of formulae that represents the conflict and classify the sub-conflicts that they represent using the classification scheme of CML.
Conclusion

Step 6: Transform the identified sub-conflicts into propositional form.

Step 7: Assess the claims made in the sub-conflicts with respect to their potential conflict power and the depth of the sub-conflicts with respect to their degree of inconsistency.\(^{227}\)

Step 8: Generate potential solutions to the identified sub-conflicts by using their propositional form as an input to the five algorithms of CRL.

Step 9: Assess the type of the solutions to the sub-conflicts generated by the five algorithms of CRL.

Step 10: Identify, for each solution and each party, the set of claims the party is required to change, and the set of claims the party can directly realise, according to the solution.

Step 11: Suggest the solution to the parties and point out what each party gains and loses in terms of claims that require change and claims that do not.

Finally, we look at some areas that were examined in the thesis but have the potential for further investigation. We identify six such areas of further investigation. Four of the areas are of theoretical nature, i.e. they are of interest from a logical perspective rather than the perspective of a user of the model.

First, the branching-time framework of the temporal fragment of CML provides an interesting basis for further research as it combines, in a unique way, the Rt-approach of Rescher’s temporal systems with the semantics of typical branching-time temporal logics,

\(^{227}\) Step 7 is optional in the sense that it is not necessarily required for the generation of solutions to a conflict.
such as CTL. As a consequence, the system has a syntactical representation for specific time points and its axiomatisation contains a complete description of the forwards-branching, backwards-linear, tree structures that constitute its semantics. Further questions include the relationship of the system to other temporal logics and also to logics such as hybrid logic and first-order modal logic, or the potential of the system to incorporate STIT theory and thereby constitute a system for a temporal modal logic of agency.

A further theoretical question to elaborate concerns the relationship between the notion of degrees of inconsistency and paraconsistency. As a consequence of having a measure for the inconsistency of sets of formulae, the class of inconsistent sets can be ordered into classes of equally inconsistent sets. This ordering of inconsistency parallels the view in paraconsistent logic according to which not every inconsistent set of formulae is trivial, i.e. there are degrees of inconsistency such that some inconsistent sets are trivial, whereas others are not.

The third theoretical question that might be interesting for further exploration concerns the relationship between the algorithm for making an inconsistent set consistent, as presented in the thesis, and other algorithms that perform this task, in particular the algorithms used in the AGM model for the consolidation operator. The method applied in our algorithm differs from established algorithms as it just replaces formulae from the inconsistent set by weaker formulae, instead of completely removing them. It can be argued that the result of our algorithm is, therefore, closer to the original set than the results produced by the updating operator of the AGM model.

The fourth theoretical concern about our model deals with the relationship between the informational input, the description of a conflict in terms of CML-formulae, and the
informational output, the solutions generated by the algorithms of CRL. Obviously, the output depends on the input, i.e. the types of solutions we can generate for a particular conflict with our model depend on how we describe the conflict in the first place. Although this dependency is not surprising, it might lead to controversial results. In particular, it might be interesting to address the question of monotonicity with regard to the informational input and output of our model. As a whole, our model is non-monotone in that input given in addition to an already provided conflict description influences the solutions generated by the model. Thus, it might be worthwhile to standardise the way conflicts are to be described by CML.

From a practical point of view, we could use our logic-based model to design a database for conflicts that contains not only structural information of conflicts, such as the time they break out/end, the number of casualties, or the agents involved in them, but also information about their content, i.e. information about the beliefs, goals, norms, and emotions that constitute them and the logical relationships between those elements. In a further project we could link this CML-based database to already existing conflict databases\textsuperscript{228}, such as COW\textsuperscript{229}, COSIMO\textsuperscript{230}, UCDP\textsuperscript{231}, etc. Once a database has been created, we can collect data about conflicts, store it in the database, and use it to simulate conflict resolutions by applying the algorithms of CRL to selected data of the database.

\textsuperscript{228} For a guide to conflict data and conflict datasets, see (Eck 2005).
\textsuperscript{229} The Correlates of War project (COW), which originated in the University of Michigan is the largest databank on wars available covering wars between 1816 and 1997. See (Bremer and Cusack 1995).
\textsuperscript{230} COSIMO is the databank used by the Heidelberg Institute for International Conflict Research in which the institute records information on political conflicts from 1945 onward. See (Pfetsch and Rohloff 2000).
\textsuperscript{231} The Uppsala Conflict Data Program run by the University of Uppsala has collected data on minor armed conflicts, intermediate armed conflicts, and wars covering the years from 1946 onwards. See (Harbom and Wallensteen 2007).
This brings us to the last area of further investigation of implementing the algorithms. As presented here, the five algorithms of CRL are intractable. We could make them workable by finding tractable algorithms for their sub-tasks and then programming them.
BIBLIOGRAPHY


HEIDER, F. (1958) The psychology of interpersonal relations, New York, Wiley.


KAMP, J. A. W. (1968) Tense Logic and the Theory of Order. Los Angeles, UCLA.


Bibliography


334


THURSTON, L. L. & CHAVE, E. J. (1929) The measurement of attitude; a psychophysical method and some experiments with a scale for measuring attitude toward the church, Chicago, Ill., The University of Chicago Press.


