Ramification analysis using causal mapping

Anthony Hunter
Department of Computer Science
University College London
Gower Street
London WC1E 6BT
(a.hunter@cs.ucl.ac.uk)

June 8, 1999

Abstract

To operate in the real-world, intelligent agents constantly need to absorb new information, and to consider the ramifications of it. This raises interesting questions for knowledge representation and reasoning. Here we consider ramification analysis in which we wish to determine both the likely outcomes from events occuring and the less likely, but very significant outcomes, from events occuring. To formalize ramification analysis, we introduce the notion of causal maps for modelling "causal relationships" between events. In particular, we consider existential event classes, for example `presidential-election`, with instances being *true, false*, or *unknown*, and directional events classes, for example `inflation`, with instances being *increasing, decreasing* or *unchanging*. Using causal maps, we can propagate new information to determine possible ramifications. These ramifications are also described in terms of events. Whilst causal maps offer a lucid view on ramifications, we also want to support automated reasoning, to address problems of incompleteness, and to represent further conditions on ramifications. To do this, we translate causal maps into default logic, and use the proof theory and automated reasoning technology of default logic. In this paper, we provide a syntax and semantics for causal mapping, and a translation into default logic, and discuss an integration of the approach with langauge engineering.

Keywords: default logic, non-monotonic logic, handling incompleteness, modelling causality, graphical knowledge representation and reasoning, knowledge engineering, language engineering.

1 Introduction

In order to operate in an environment, an intelligent agent needs to consider the effects of events in that environment. These events may result from the agent's actions or result from exogenous causes. We call this assessment of events in an environment **ramification analysis**. Consider for example the news that there is a train drivers' strike tomorrow. Ramifications include many people will have difficulty getting into work on time which in turn means meetings may be running late and be poorly attended, and this is likely to have a knock-on effect on schedules for meetings in the subsequent weeks.

Formalizing ramification analysis for an intelligent agent is a difficult problem in general. It calls for comprehensive common-sense reasoning and general knowledge. However, there is the possibility to develop ramification analysis for constrained problems, where the set of events under consideration is limited.

In ramification analysis, we wish to determine both the likely outcomes from news and also the less likely, but very significant, outcomes from news. The aim is not to take news and determine just the most likely outcomes. Rather it is to explore possibilities. We are therefore not proposing some qualitative form of probabilistic reasoning. Nor are we adopting some form of possibility theory. Rather we are treating possibilities equally and focussing our attention on their interactions. In addition, we are interested in addressing issues of incompleteness in reasoning about news in a logic-based framework.

To do ramification analysis, we exploit symbolic representations. This obviates some of the complexity of integer or continuous real-valued variables and differential calculus-modelling by only considering the qualitative nature of the systems being modelled. Useful inferences can be drawn from a model with a limited number of values for a qualitative variable.

Whilst we describe new information as news, we don't intend this to restrict the new information to be from just news sources in the narrow sense of the word, i.e news such as in newspapers, TV news, and newsfeeds. Rather we regard news in a broader sense where an agent operating in the real world is receiving new information from communication with other agents and from observing the world directly such as by seeing and hearing. However, we do see handling news in the narrow sense of the word as being a particularly important application of ramification analysis in the short term. We address this application area in more detail in Section 4. In particular, we will discuss how information extraction technology can be harnessed to indentify logical representations of news from news reports in the form of free text. Information extraction is based on natural language processing, and by focussing on restricted domains such as on news reports from online newsfeeds, it can give impressive performance [8, 14, 2].

However, handling textual news reports is only one application of ramification analysis. Potentially, the approach could be used by intelligent systems such as robotic systems, surveillance systems, and monitoring systems, where the news is obtained from various kinds of sensors such as acoustic, vision and sonar sensors.

In this paper, we have used causal mapping as an abstraction, or outline specification of a default theory. We believe that the structure and the graphical representation offered by causal mapping facilitates the development of default theories. In this way, we have used causal mapping as part of an approach for knowledge engineering for default logic. In addition, we regard causal maps as a means for explaining events and their ramifications.

In the following sections, we consider the need for ramification analysis for identifying possible ramifications of news, we show how we can use causal mapping to represent and reason with the ramifications of events, and then show how we can capture causal maps in default logic. We also discuss integration of causal mapping with language engineering.

2 Causal mapping

In **causal mapping**, we capture causal relationships between events. In Section 2.1, we define the syntax for events, news reports, and causal maps. In section 2.2, we present the semantics in terms of truth values. In section 2.3, we consider conflicts arising in the defeasible reasoning with causal maps.

2.1 Syntax for causal maps

Definition 2.1 *An* **event** *is a ground monadic predicate where the predicate symbol is the* **event class** *and the argument is the* **event value***. An event is an instance of an event class. The set of event classes is unlimited but the set of event values is* $\{+, -, 0\}$*. There are two types of event class* **existential** *and* **directional***. An event class* α *is directional iff* α *is not existential.*

Example 2.1 *Consider the event classes* `inflation` *and* `presidential-election`*. Let the first be a directional and the second be existentional. Events that can be formed from the event classes are:*

```
inflation(+)
inflation(0)
inflation(-)
presidential-election(+)
presidential-election(0)
presidential-election(-)
```

The intuition for the event values for the existential event class is described as follows:

- Positive existence, denoted by $\alpha(+)$, represents it is known that α holds.
- Zero existence, denoted by $\alpha(0)$, represents it is unknown whether α holds or not.
- Negative existence, denoted by $\alpha(-)$, represents it is known that α does not hold.

The intuition for the event values for the directional event class is described as follows:

- Positive change, denoted by $\alpha(+)$, represents it is known that α is rising.
- Zero change, denoted by $\alpha(0)$, represents it is known that α is unchanged.
- Negative change, denoted by $\alpha(-)$, represents it is known that α is falling.

Existential events refer to observables that only hold for discrete periods of time, and so do not always exist, whereas directional events refer to observables that are measured continuously or, at regular points in time, and so always exist, and only change in direction. Since, we regard ramification analysis as looking at a snapshot of the observerables in the world, we believe that it is intuitive to assign positive, zero, and negative, event values for both the existential and directional event classes.

Definition 2.2 *A* **news report** *is a set of events.*

A news report is information that is regarded as certain. It is taken at a point in time. The events in the news report reflect observables in the real-world. Ramification analysis is intended to identify possible ramifications of the events in the news report.

Example 2.2 *For the event classes* `presidential-election, inflation`*, and if there is new information that* `presidential-election` *is occuring, and the* `inflation` *is decreasing, then the corresponding news report is captured as:*

{`presidential-election(+), inflation(-)`}

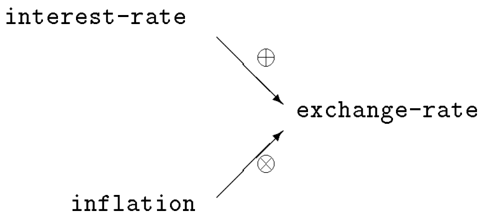

Figure 1: A causal map where if `interest-rate` increases then `exchange-rate` increases, if `interest-rate` decreases then `exchange-rate` decreases, if `inflation` increases then `exchange-rate` decreases, and if `inflation` decreases then `exchange-rate` increases.

For causal mapping, we consider two types of causal relationship, namely increasing causality and decreasing causality, which are defined as follows.

Definition 2.3 Increasing causality: *Let α and β be event classes. A causal relation from α to β is increasing iff whenever the event value for α is $+$ (or respectively $-$), then the event value for β is $+$ (or respectively $-$).*

Definition 2.4 Decreasing causality: *Let α and β be event classes. A causal relation from α to β is decreasing iff whenever the event value for α is $+$ (or respectively $-$) then the event value for β is $-$ (or respectively $+$).*

So increasing causality can equivalently be described as a strictly increasing monotonic function. Similarly, decreasing causality can equivalently be described as a strictly decreasing monotonic function. (Recall, a function $f : A \rightarrow B$ is a strictly increasing monotonic function iff for all $x, y \in A$ such that $x > y$, then $f(x) > f(y)$; and a function $f : A \rightarrow B$ is a strictly monotonic decreasing function if for all $x, y \in A$ such that $x > y$, then $f(x) < f(y)$.) For this, we assume that $+ < 0 < -$ holds as the ordering over event values.

A causal map is a set of event classes connected by arrows, with the direction of the arrow indicating the direction of influence or causality. We define a causal map as follows.

Definition 2.5 *A* **causal map** *is a directed graph (N, A) where N is a set of event classes and A is a set of labelled directed arcs. Each node represents an event class. Each arc is annotated with either $\oplus$ or $\otimes$. The direction of each arc indicates the direction of influence or causality. The annotation on an arc indicates the type of causality, where $\oplus$ denotes increasing causality and $\otimes$ denotes decreasing causality. For each arc $\oplus(\alpha, \beta) \in A$, and similarly for each arc $\otimes(\alpha, \beta) \in A$, the event class α is the antecedent, and the event class β is the consequent.*

Example 2.3 *Consider Figure 1. Here* `inflation` *has a causal effect on the* `exchange-rate`*. This effect is decreasing causality. Also* `interest-rate` *has a causal effect on* `exchange-rate`*, and this effect is increasingly causality. So increasing* `inflation` *decreases the* `exchange-rate` *whereas increasing* `interest-rate` *increases the* `exchange-rate`*.*

Definition 2.6 *A* **confluence** *in a causal map (N, A) is a subgraph of (N, A) of one of the following three forms:*

$$(\{\alpha, \beta, \gamma\}, \{\oplus(\alpha, \gamma), \oplus(\beta, \gamma)\})$$
$$(\{\alpha, \beta, \gamma\}, \{\oplus(\alpha, \gamma), \otimes(\beta, \gamma)\})$$
$$(\{\alpha, \beta, \gamma\}, \{\otimes(\alpha, \gamma), \otimes(\beta, \gamma)\})$$

The **head** *of the confluence is the event class γ, and the* **tails** *of the confluence are the event classes α and β. An event class can be the head of more than one confluence. Similarly, an event class can be the tail of more than one confluence. For a head γ, the set of tails of γ is $Tails(\gamma)$.*

A confluence such as that in Figure 1 is problematical if both `interest-rate` and `inflation` have event value $+$. Do we increase or decrease the event value for `exchange-rate`? In fact, any confluence is problematical. Consider the causal map $(\{\alpha, \beta, \gamma\}, \{\oplus(\alpha, \gamma), \oplus(\beta, \gamma)\})$. Here if α has event value $+$, and β has event value $-$, then what is the event value for γ. We will address the problem in Section 2.3, but first we need to consider the semantics for causal maps.

2.2 Semantics for causal maps

We adopt a three-valued semantics for causal maps.

Definition 2.7 *There are three truth-values:* **positive***, denoted $+$,* **negative***, denoted $-$, and* **zero***, denoted 0.*

Definition 2.8 *An interpretation is a function h from the set of event classes into the set of truth-values $\{+, -, 0\}$.*

In the following defintion, we give the conditions required to ensure that an interpretation is a model of a news report together with a causal map. In the definition, there is a priority for events in the news reports over events derived from a causal map. This reflects the intuition that causal relationships capture possibilities rather than certainties, whereas news reports are assumed to be certainties. To support this, the definition incorporates a notion of rebut. Essentially, an event with event class γ is rebutted if, and only if, either there is an event in the news report that has the same event but different event value, or γ is part of a confluence such that the assignment of event values to the tails leads to a conflict or ambiguity about the event value for γ.

Definition 2.9 *For a causal map (N, A), a news report P, and an interpretation h, h is a model for (P, N, A) if and only if for all $\alpha \in N$ the following nine conditions hold:*

1. *If $\alpha(+) \in P$, then $h(\alpha) = +$.*
2. *If $\alpha(0) \in P$, then $h(\alpha) = 0$.*
3. *If $\alpha(-) \in P$, then $h(\alpha) = -$.*
4. *If $h(\alpha) = +$, and $\oplus(\alpha, \gamma) \in A$, and $\gamma(+)$ is not rebutted in (P,N,A) for h, then $h(\gamma) = +$.*
5. *If $h(\alpha) = -$, and $\oplus(\alpha, \gamma) \in A$, and $\gamma(-)$ is not rebutted in (P,N,A) for h, then $h(\gamma) = -$.*
6. *If $h(\alpha) = +$, and $\otimes(\alpha, \gamma) \in A$, and $\gamma(-)$ is not rebutted in (P,N,A) for h, then $h(\gamma) = -$.*
7. *If $h(\alpha) = -$, and $\otimes(\alpha, \gamma) \in A$, and $\gamma(+)$ is not rebutted in (P,N,A) for h, then $h(\gamma) = +$.*
8. *If $h(\alpha) = 0$, and $(\oplus(\alpha, \gamma) \in A$ or $\otimes(\alpha, \gamma) \in A)$, and $\gamma(0)$ is not rebutted in (P,N,A) for h, then $h(\gamma) = 0$.*
9. *If $\alpha(+) \notin P$, and $\alpha(-) \notin P$, and there is no β such that $(\oplus(\beta, \alpha) \in A$ or $\otimes(\beta, \alpha) \in A)$, then $h(\alpha) = 0$.*

where rebutted is defined by the following three conditions:

1. *$\gamma(+)$ is rebutted in (P,N,A) for h iff $\gamma(-) \in P$, or $\gamma(0) \in P$, or there is an $\otimes(\beta,\gamma) \in A$ such that $h(\beta) = +$, or there is an $\oplus(\beta,\gamma) \in A$ such that $h(\beta) = -$.*

2. *$\gamma(0)$ is rebutted in (P,N,A) for h iff $\gamma(+) \in P$, or $\gamma(-) \in P$, or there is an $\otimes(\beta,\gamma) \in A$ such that $h(\beta) = +$, or there is an $\oplus(\beta,\gamma) \in A$ such that $h(\beta) = -$, or there is an $\otimes(\beta,\gamma) \in A$ such that $h(\beta) = -$, or there is an $\oplus(\beta,\gamma) \in A$ such that $h(\beta) = +$.*

3. *$\gamma(-)$ is rebutted in (P,N,A) for h iff $\gamma(+) \in P$, or $\gamma(0) \in P$, or there is an $\otimes(\beta,\gamma) \in A$ such that $h(\beta) = -$, or there is an $\oplus(\beta,\gamma) \in A$ such that $h(\beta) = +$.*

Note, in Definition 2.9, there is a priority for propagating the $+$ and $-$ truth-values over the 0 truth-value. So for example, in Figure 3, if $P = \{\texttt{A(0)}, \texttt{B}(+)\}$, then any model for the causal map and P would have $h(\texttt{C}) = +$.

The intuition for this aspect of the definition for the semantics is that we are interested in propagating perturbations (changes) through the causal map and do not want unchanged event classes to block the range of possibilities. In this way, apart from items in the news report, the zero truth-value is a default truth-value: If no perturbation can be identified for an event class, it has the zero truth-value by default.

Example 2.4 *Let $P = \{\texttt{A}(+)\}$ and (N, A) be the causal map given in Figure 2. For (P, N, A), there is one model h_1 below:*

$$h_1(\texttt{A}) = +, h_1(\texttt{B}) = +, h_1(\texttt{C}) = -, h_1(\texttt{D}) = -$$

Example 2.5 *Let $P = \{\texttt{B}(-)\}$ and (N, A) be the causal map given in Figure 2. For (P, N, A), there is one model h_1 below:*

$$h_1(\texttt{A}) = 0, h_1(\texttt{B}) = -, h_1(\texttt{C}) = +, h_1(\texttt{D}) = +$$

Example 2.6 *Let $P = \{\texttt{A}(+)\}$ and (N, A) be the causal map given in Figure 3. For (P, N, A), there is one model h_1 below:*

$$h_1(\texttt{A}) = +, h_2(\texttt{B}) = 0, h_1(\texttt{C}) = +$$

Example 2.7 *Let $P = \{\texttt{A}(+), \texttt{B}(+)\}$ and (N, A) be the causal map given in Figure 3. For (P, N, A), models include h_1, h_2, and h_3 below:*

$$\begin{array}{lll} h_1(\texttt{A}) = + & h_1(\texttt{B}) = + & h_1(\texttt{C}) = + \\ h_2(\texttt{A}) = + & h_2(\texttt{B}) = + & h_2(\texttt{C}) = 0 \\ h_3(\texttt{A}) = + & h_3(\texttt{B}) = + & h_3(\texttt{C}) = - \end{array}$$

Definition 2.10 *For a causal map (N, A), and a news report P, if there is no h such that h is a model for (P, N, A), then (P, N, A) is inconsistent. (P, N, A) is consistent iff it is not consistent.*

Definition 2.11 *For a news report P, if there is an event category α, such that one or more of the following conditions hold, (1) $\alpha(+) \in P$ and $\alpha(0) \in P$, or (2) $\alpha(+) \in P$ and $\alpha(-) \in P$, or (3) $\alpha(0) \in P$ and $\alpha(-) \in P$, then P is contradictory.*

Proposition 2.1 *For any news report P, and any casual map (N, A), P is contradictory iff (P, N, A) is inconsistent.*

Proof: ($\Rightarrow$) If P is contradictory, there is a pair of events $\phi, \psi \in P$, where the event class of ϕ and ψ is the same, say α, but the event values are different. However, there is no interpretation h such that $h(\alpha)$ has more than one image, and so Definition 2.9 is not satisfiable. ($\Leftarrow$) If (P, N, A) is inconsistent, then one of the Conditions 1...9 in Definition 2.9 has been the means for the violation by (P, N, A). Because of the rebut condition in each of Conditions 4...8, none of these can give a violation by any (P, N, A). Similarly, the antecedent of Condition 9 cannot give a violation by any (P, N, A). The only conditions that can give a violation are Conditions 1..3. Hence, P is contradictory.

Definition 2.12 *The transitive closure of (N, A), denoted $\Sigma(N, A)$, is the smallest set satisfying the following:*

1. *If $\oplus(\alpha, \beta) \in A$, then $\oplus(\alpha, \beta) \in \Sigma(N, A)$.*
2. *If $\oplus(\alpha, \beta) \in A$, then $\otimes(\alpha, \beta) \in \Sigma(N, A)$.*
3. *If $\oplus(\alpha, \beta) \in \Sigma(N, A)$ and $\oplus(\beta, \gamma) \in \Sigma(N, A)$, then $\oplus(\alpha, \gamma) \in \Sigma(N, A)$.*
4. *If $\otimes(\alpha, \beta) \in \Sigma(N, A)$ and $\oplus(\beta, \gamma) \in \Sigma(N, A)$, then $\otimes(\alpha, \gamma) \in \Sigma(N, A)$.*
5. *If $\oplus(\alpha, \beta) \in \Sigma(N, A)$ and $\otimes(\beta, \gamma) \in \Sigma(N, A)$, then $\otimes(\alpha, \gamma) \in \Sigma(N, A)$.*
6. *If $\otimes(\alpha, \beta) \in \Sigma(N, A)$ and $\otimes(\beta, \gamma) \in \Sigma(N, A)$, then $\oplus(\alpha, \gamma) \in \Sigma(N, A)$.*

Definition 2.13 *For any causal map (N, A), (N, A) is coherent iff the following two conditions hold:*

1. *there are no event classes α, β such that $\oplus(\alpha, \beta) \in \Sigma(N, A)$ and $\otimes(\alpha, \beta) \in \Sigma(N, A)$.*
2. *there is no event class α such that $\otimes(\alpha, \alpha) \in \Sigma(N, A)$ or $\oplus(\alpha, \alpha) \in \Sigma(N, A)$.*

Example 2.8 *The following two causal maps are incoherent:*

$$(\{\texttt{A}, \texttt{B}\}, \{\otimes(\texttt{A}, \texttt{A}), \otimes(\texttt{A}, \texttt{A})\})$$
$$(\{\texttt{A}, \texttt{B}, \texttt{C}, \texttt{D}\}, \{\otimes(\texttt{A}, \texttt{B}), \oplus(\texttt{A}, \texttt{C}), \oplus(\texttt{B}, \texttt{D}), \oplus(\texttt{C}, \texttt{D})\})$$

Proposition 2.2 *For any causal map (N, A), if (N, A) is coherent then (N, A) is an acyclic graph.*

Proof: Follows directly from condition 2 of Definition 2.13.

Whilst acyclic graphs are desirable for causal maps for ramification analysis, we do not necessarily require other constraints on causal maps such as connectivity.

Definition 2.14 *For a causal map (N, A), and a news report P, the entailment relation $\models$ is defined as follows,*

1. *$(P, N, A) \models \alpha(+)$ iff for all h, if h is a model for (P,N,A), then $h(\alpha) = +$*

2. $(P,N,A) \models \alpha(0)$ *iff for all h, if h is a model for (P,N,A), then* $h(\alpha) = 0$

3. $(P,N,A) \models \alpha(-)$ *iff for all h, if h is a model for (P,N,A), then* $h(\alpha) = -$

Example 2.9 *Consider the causal map given in Figure 3 and let* $P_1 = \{\texttt{A}(+), \texttt{B}(-)\}$, $P_2 = \{\texttt{A}(+), \texttt{B}(+)\}$, *and* $P_3 = \{\texttt{A}(-)\}$:

$$\begin{array}{lll} (P_1,N,A) \not\models \texttt{C}(+) & (P_1,N,A) \not\models \texttt{C}(0) & (P_1,N,A) \models \texttt{C}(-) \\ (P_2,N,A) \not\models \texttt{C}(+) & (P_2,N,A) \not\models \texttt{C}(0) & (P_2,N,A) \not\models \texttt{C}(-) \\ (P_3,N,A) \models \texttt{C}(+) & (P_3,N,A) \not\models \texttt{C}(0) & (P_3,N,A) \not\models \texttt{C}(-) \end{array}$$

Proposition 2.3 *For a casual map* (N,A), *and news report* P, *and an event class* α,

$$\begin{array}{l} \text{If } (P,N,A) \models \alpha(+), \text{ then } (P,N,A) \not\models \alpha(0) \text{ and } (P,N,A) \not\models \alpha(-) \\ \text{If } (P,N,A) \models \alpha(-), \text{ then } (P,N,A) \not\models \alpha(+) \text{ and } (P,N,A) \not\models \alpha(0) \\ \text{If } (P,N,A) \models \alpha(0), \text{ then } (P,N,A) \not\models \alpha(+) \text{ and } (P,N,A) \not\models \alpha(-) \end{array}$$

Proof: Follows from Definition 2.14 and the constraint that an interpretation is a function.

Definition 2.15 *Let h be a model for* (P,N,A). *h is a* **complete model** *for* (P,N,A) *iff for all event classes* α, $\alpha \in N$ *iff* α *is in the domain of h. h is a* **unique complete model** *iff h is a complete model* (P,N,A) *and for all g, if g is a complete model for* (P,N,A), *then* $h = g$.

Proposition 2.4 *For a causal map (N,A) and a news report P, if* (N,A) *is coherent and P is not contradictory, then there is not necessarily a unique complete model for* (P,N,A).

Proof: Let (N,A) be $(\{\texttt{A},\texttt{B},\texttt{C}\}, \{\oplus(\texttt{A},\texttt{C}), \otimes(\texttt{B},\texttt{C})\})$, and $P = \{\texttt{A}(+), \texttt{B}(+)\}$. (N,A) is coherent and P is not contradictory. Yet there is no unique complete model for (P,N,A).

Unsuprisingly, if a causal map is not coherent, then there is not necessarily a unique complete model, as illustrated by the following example.

Example 2.10 *Consider the following example. Here,* (N,A) *is not coherent, and* (P,N,A) *does not have a unique complete model.*

$$\begin{array}{c} P = \{\} \\ N = \{\texttt{A},\texttt{B}\} \\ A = \{\oplus(\texttt{A},\texttt{B}), \oplus(\texttt{B},\texttt{A})\} \end{array}$$

Indeed, there are exactly three complete models h_1, h_2, *and* h_3:

$$\begin{array}{lll} h_1(\texttt{A}) = + & h_2(\texttt{A}) = 0 & h_3(\texttt{A}) = - \\ h_1(\texttt{B}) = + & h_2(\texttt{B}) = 0 & h_3(\texttt{B}) = - \end{array}$$

Coherence is also not necessary for a unique complete model.

Example 2.11 *Consider the following example. Here,* (N,A) *is not coherent.*

$$\begin{array}{c} P = \{\texttt{A}(+), \texttt{B}(+)\} \\ N = \{\texttt{A},\texttt{B},\texttt{C}\} \\ A = \{\otimes(\texttt{A},\texttt{B}), \oplus(\texttt{B},\texttt{C}), \oplus(\texttt{C},\texttt{A})\} \end{array}$$

Here (P,N,A) *does have a unique complete model* h_1.

$$h_1(\texttt{A}) = + \quad h_1(\texttt{B}) = + \quad h_1(\texttt{C}) = +$$

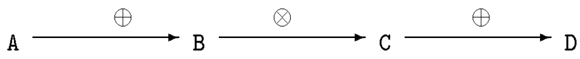

Figure 2: A linear causal map.

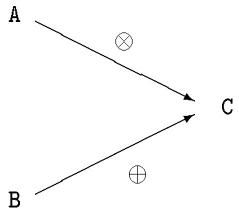

Figure 3: A converging causal map.

Example 2.12 *The following causal map is not coherent.*

$$P = \{\}$$
$$N = \{\mathtt{A}, \mathtt{B}\}$$
$$A = \{\otimes(\mathtt{A}, \mathtt{B}), \oplus(\mathtt{B}, \mathtt{A})\}$$

Here (P, N, A) does have a unique complete model h_1, where $h_1(\mathtt{A}) = 0$ and $h_1(\mathtt{B}) = 0$.

2.3 Reasoning with causal maps

We first restrict consideration to causal maps that contain no confluences. This includes linear causal maps, for example Figure 2, and diverging causal maps, for example Figure 4. In this case, since there are no heads or tails, this reasoning is straightforward. Assuming the causal map (N, A) is coherent and the news report P is consistent, we take the transitive closure of the causal map $\Sigma(N, A)$. Then any event inferred by the $\models$ relation can be obtained by the following inference rules:

If $\alpha(+) \in P$, and $\oplus(\alpha, \beta) \in \Sigma(N, A)$, then $\beta(+)$ is an inference.
If $\alpha(-) \in P$, and $\oplus(\alpha, \beta) \in \Sigma(N, A)$, then $\beta(-)$ is an inference.
If $\alpha(+) \in P$, and $\otimes(\alpha, \beta) \in \Sigma(N, A)$, then $\beta(-)$ is an inference.
If $\alpha(-) \in P$, and $\otimes(\alpha, \beta) \in \Sigma(N, A)$, then $\beta(+)$ is an inference.

Unfortunately, causal maps without confluences are too restricted for many situations, so we now consider more general causal maps, such as in Example 2.13 below.

Example 2.13 *Consider the causal map that is captured in Figure 5. We have the following knowledge about a sector: (1) The aircraft market is increasing; (2) The cost of aircraft development is increasing; and (3) The effect of the second observation is more pronounced than the first. From this, we would like to derive the following: The unit cost of each aircraft produced is increasing, and*

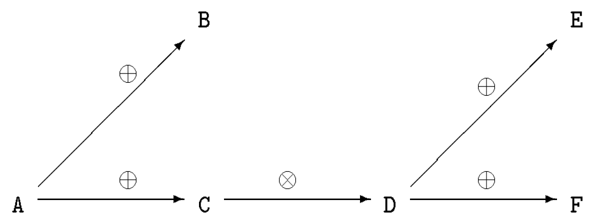

Figure 4: A diverging causal map.

therefore there is a downward pressure on profitability for each aircraft maufacturer. Now suppose we obtain the news that a particular industrial company with no aircraft manufacturing interests has just bought an aircraft manufacturer. What ramification can we derive?

In the unrestricted case, we need to address the problem of reasoning with confluences. Consider Figure 1. If both `interest-rate` and `inflation` increase, then does `exchange-rate` increase, decrease or remain the same? Two possible options are: (1) introduce "unknown" value, and extend the theory to propagate "unknown"; and (2) introduce extra information to resolve the conflict for each head. We adopt the second option. For this, we define the notion of a resolution table.

Definition 2.16 *A* **resolution table** *for a head γ is a table, denoted T_γ where for each tail α in the confluence of γ, there is a column in T_γ labelled with α. The last column in the table is labelled with γ. So for any combination of event values for the tails, we can determine the event value for the head.*

Consider the resolution table in Table 1 for the confluence in Figure 6. To use this, we find the row in the resolution table that has the correct event values for the tails of `exchange-rate`, and then read the event value for `exchange-rate`. For example, suppose the tail `interest-rate` has event value $+$, the tail `trade-surplus` has event value $-$, and the tail `inflation` has event value $+$, then we read the event value for the head `exchange-rate`. We now need to generalize the notion of a causal map to formalize the use of resolution tables.

Definition 2.17 *A* **tabulated causal map** *denoted (N, A, T) is a causal map (N, A) and a set of resolution tables T such that for each head γ in (N, A), there is a resolution table for γ, denoted T_γ, in T.*

Definition 2.18 *Let T_γ be a resolution table for a head γ. T_γ is* **complete**, *if there is a row T_γ for every combination of truth-values for the elements of $Tail(\gamma)$, otherwise the table is incomplete. T_γ is* **consistent**, *if every combination of truth-values for the elements of $Tail(\gamma)$, there is at most one row in T_γ, otherwise T_γ is inconsistent.*

If T_γ be a a complete and consistent resolution table, and $|T_\gamma| = n$, then there are $n+1$ columns, and 3^n rows, in T_γ.

Now, we provide a semantics for tabulated causal maps.

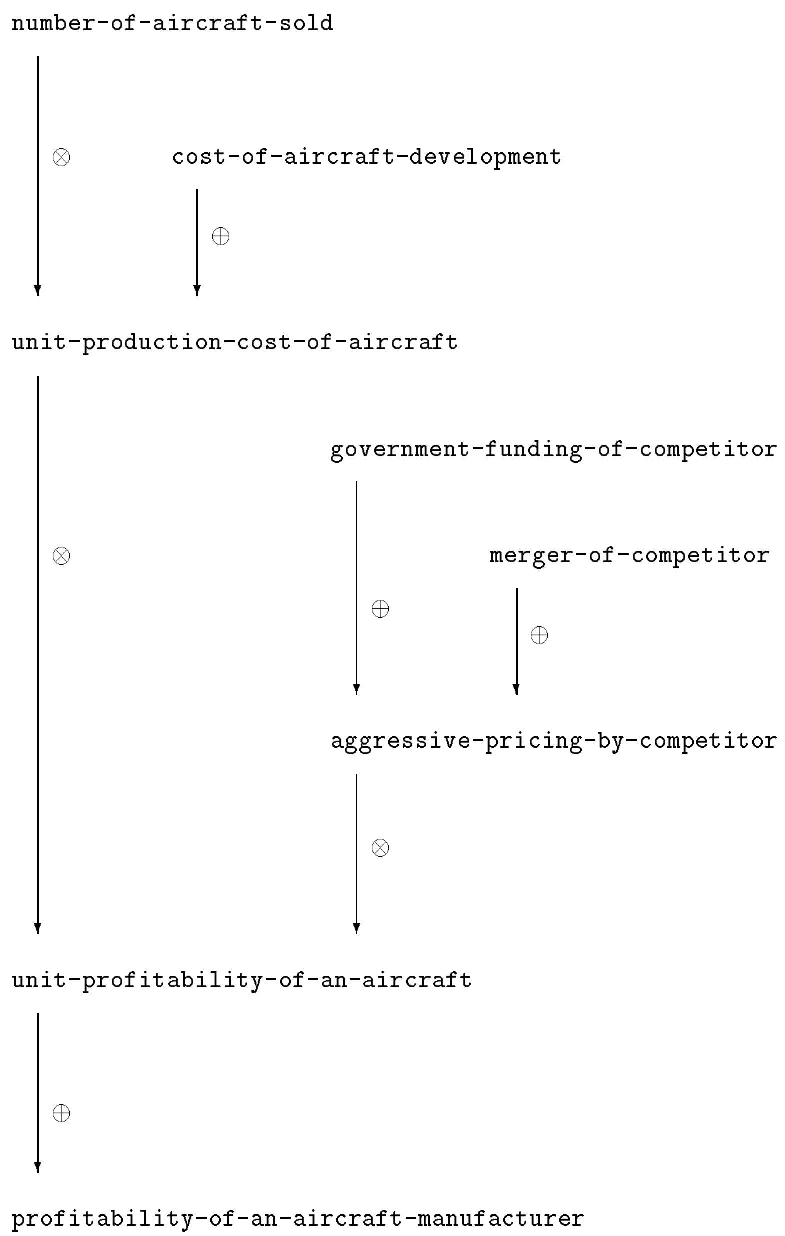

Figure 5: A causal map showing the like ramifications of news on the profitability of an aircraft manufacturer

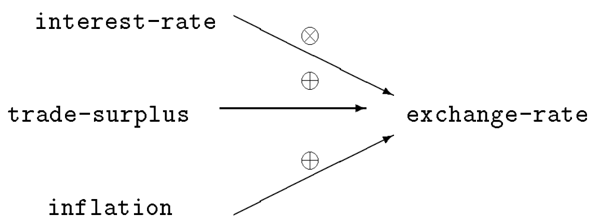

Figure 6: A confluence for `exchange-rate`.

`interest-rate`	`trade-surplus`	`inflation`	`exchange-rate`
+	+	+	+
+	+	0	+
+	+	−	+
+	0	+	0
+	0	0	+
+	0	−	+
+	−	+	0
+	−	0	+
+	−	−	+
0	+	+	0
0	+	0	0
0	+	−	+
0	0	+	−
0	0	0	0
0	0	−	+
0	−	+	−
0	−	0	−
0	−	−	0
−	+	+	−
−	+	0	−
−	+	−	−
−	0	+	−
−	0	0	−
−	0	−	−
−	−	+	−
−	−	0	−
−	−	−	−

Table 1: A resolution table for the confluence of in Figure 6, where the head is `exchange-rate` and $Tail($`exchange-rate`$) = \{$`interest-rate,trade-surplus,inflation`$\}$.

Definition 2.19 *For a tabulated causal map (N, A, T), a news report P, and an interpretation h, h is a model for (P, N, A, T) if and only if for the $\alpha \in N$ the following ten conditions hold.*

1. *If $\alpha(+) \in P$, then $h(\alpha) = +$.*
2. *If $\alpha(0) \in P$, then $h(\alpha) = 0$.*
3. *If $\alpha(-) \in P$, then $h(\alpha) = -$.*
4. *If $h(\alpha) = +$, and $\oplus(\alpha, \gamma) \in A$, and γ is not a head, and $\gamma(0) \notin P$, and $\gamma(-) \notin P$, then $h(\gamma) = +$.*
5. *If $h(\alpha) = -$, and $\oplus(\alpha, \gamma) \in A$, and γ is not a head, and $\gamma(0) \notin P$, and $\gamma(+) \notin P$, then $h(\gamma) = -$.*
6. *If $h(\alpha) = +$, and $\otimes(\alpha, \gamma) \in A$, and γ is not a head, and $\gamma(+) \notin P$, and $\gamma(0) \notin P$, then $h(\gamma) = -$.*
7. *If $h(\alpha) = -$, and $\otimes(\alpha, \gamma) \in A$, and γ is not a head, and $\gamma(0) \notin P$, and $\gamma(-) \notin P$, then $h(\gamma) = +$.*
8. *If $h(\alpha) = 0$, and $(\oplus(\alpha, \gamma) \in A$, or $\otimes(\alpha, \gamma) \in A)$, and γ is not a head, and $\gamma(+) \notin P$, and $\gamma(-) \notin P$, then $h(\gamma) = 0$.*
9. *If γ is a head, and $\gamma(+) \notin P$, and $\gamma(0) \notin P$, and $\gamma(-) \notin P$, and $Tail(\gamma) = \{\alpha_1, ..., \alpha_n\}$, and T_γ is the resolution table for γ, and there is a row $(x_1, ..., x_n, x_{n+1}) \in T_\gamma$, where x_1 is the event value for the event class α_1, and x_n is the event value for the event class α_n, and $h(\alpha_1) = x_1$ and ... and $h(\alpha_n) = x_n$, then $h(\gamma) = x_{n+1}$.*
10. *If $\alpha(+) \notin P$, and $\alpha(-) \notin P$, and there is no β such that $(\oplus(\beta, \alpha) \in A$ or $\otimes(\beta, \alpha) \in A)$, then $h(\alpha) = 0$.*

Note, since tabulated causal maps have resolution tables, the notion of rebutted is not required in Definition 2.19.

Definition 2.20 *For a tabulated causal map (N, A, T), and a news report P, the entailment relation $\models$ is defined as follows,*

1. *$(P, N, A, T) \models \alpha(+)$ iff for all h, if h is a model for (P,N,A), then $h(\alpha) = +$*
2. *$(P, N, A, T) \models \alpha(0)$ iff for all h, if h is a model for (P,N,A), then $h(\alpha) = 0$*
3. *$(P, N, A, T) \models \alpha(-)$ iff for all h, if h is a model for (P,N,A), then $h(\alpha) = -$*

Definition 2.21 *For a tabulated causal map (N, A, T), and a news report P, if there is no h such that h is a model for (P, N, A, T), then (P, N, A, T) is inconsistent. (P, N, A, T) is consistent iff it is not consistent.*

Definition 2.22 *Let h be a model for (P, N, A, T). h is a* **complete model** *for (P, N, A, T) iff for all event classes α, $\alpha \in N$ iff α is in the domain of h. h is a* **unique complete model** *iff h is a complete model (P, N, A, T) and for all g, if g is a complete model for (P, N, A, T), then $h = g$.*

Proposition 2.5 *For a tabulated causal map (N, A, T), and a news report P, if (N, A) is coherent, and each $T_\gamma \in T$ is complete and consistent, and P is not contradictory, then there is a unique complete model for (P, N, A, T).*

Proof: For each $\alpha \in N$, exactly one of the following applies: (1) there is an event in P with event classes α; or (2) α is a consequent but not a head; or (3) α is a head and there is a resolution table T_α. In any of these cases, there is a unique truth value of α, given (P, N, A, T). Since this hold for all α, there is a unique complete model for (P, N, A, T).

If all the resolution tables are consistent and complete in a tabulated causal map, the representation and reasoning is expressible in classical propositional logic.

Definition 2.23 *For a tabulated causal map (N, A, T), the function Π translates (N, A, T) into a set of classical propositional formulae such that:*

$$\begin{aligned}
&\Pi(N, A, T) = \\
&\quad \{\alpha_1(x_1) \wedge .. \wedge \alpha_n(x_n) \rightarrow \alpha_{n+1}(x_{n+1}) \mid (x_1, ..., x_n, x_{n+1}) \in T_\gamma \textit{ and } T_\gamma \in T\} \\
&\quad\quad \cup \{\alpha(+) \rightarrow \beta(+), \alpha(-) \rightarrow \beta(-), \alpha(0) \rightarrow \beta(0) \mid \oplus(\alpha, \beta) \in A \textit{ and } \beta \textit{ is not a head}\} \\
&\quad\quad \cup \{\alpha(+) \rightarrow \beta(-), \alpha(-) \rightarrow \beta(+), \alpha(0) \rightarrow \beta(0) \mid \otimes(\alpha, \beta) \in A \textit{ and } \beta \textit{ is not a head}\}
\end{aligned}$$

Definition 2.24 *For a tabulated causal map (N, A, T), and news report P, the function Φ translates (N, A, T) into a set of classical propositional formulae such that*

$$\begin{aligned}
&\Phi(P, N, A, T) = \\
&\quad P \cup \{(\alpha_1(x_1) \wedge ... \wedge \alpha_n(x_n) \rightarrow \alpha_{x+1}(x_{n+1})) \in \Pi(N, A, T) \mid \\
&\quad\quad \alpha_{n+1}(+) \notin P \textit{ and } \alpha_{n+1}(0) \notin P \textit{ and } \alpha_{n+1}(-) \notin P\} \\
&\quad \cup \{\alpha(0) \mid \alpha(+) \notin P \textit{ and } \alpha(-) \notin P \\
&\quad\quad \textit{and there is no } \beta \textit{ such that } (\oplus(\beta, \alpha) \in A \textit{ or } \otimes(\beta, \alpha) \in A)\}
\end{aligned}$$

Proposition 2.6 *For a tabulated causal map (N, A, T), where (N, A) is coherent and each $T_\gamma \in T$ is complete and consistent, a news report P, and an event ϕ,*

$$(P, N, A, T) \models \phi \textit{ iff } \Phi(P, N, A, T) \vdash \phi$$

where $\vdash$ is the classical propositional logic consequence relation.

Proof: ($\Rightarrow$) Suppose $(P, N, A, T) \models \alpha(+)$ for some event class α (event values 0 and $-$ can be handled similarly). Either $\alpha(+) \in P$ or $\alpha(+) \notin P$ hold. Suppose $\alpha(+) \in P$ holds. By Definition 2.24, $\alpha(+) \in \Phi(P, N, A, T)$ and so by definition of $\vdash$, $\Phi(P, N, A, T) \vdash \alpha(+)$ holds. Now suppose $\alpha(+) \notin P$. So either α is a head or α is not a head. Suppose α is a head. By Definition 2.19, there is a resolution table T_α for the head α, and a row $(x_1, .., x_n, +) \in T_\alpha$, and by recursion $(P, N, A, T) \models \alpha_1(x_1)$ and ... and $(P, N, A, T) \models \alpha_n(x_n)$. So by Definition 2.24, $\exists \alpha_1(x_1) \wedge .. \wedge \alpha_n(x_n) \rightarrow \alpha(+) \in \Phi(P, N, A, T)$ and by recursion, $\Phi(P, N, A, T) \vdash \alpha_1(x_1)$ and ... and $\Phi(P, N, A, T) \vdash \alpha_n(x_n)$. So by the definition of $\vdash$, $\Phi(P, N, A, T) \vdash \alpha(+)$ holds. Now suppose α is not a head. By Definition 2.19, either $\oplus(\beta, \gamma) \in A$ or $\otimes(\beta, \gamma) \in A$. Suppose $\oplus(\beta, \gamma) \in A$ ($\otimes(\beta, \gamma) \in A$ can be handled similarly), then by recursion $(P, N, A, T) \models \beta(+)$. So by Definition 2.24, $\beta(+) \rightarrow \alpha(+) \in \Phi(P, N, A, T)$ and by recursion $\Phi(P, N, A, T) \vdash \beta(+)$. So by the Definition of $\vdash$, $\Phi(P, N, A, T) \vdash \alpha(+)$. The proof of ($\Leftarrow$) follows similarly.

Proposition 2.7 *For a tabulated causal map (N, A, T), where (N, A) is coherent and each $T_\gamma \in T$ is complete and consistent, a news report P, and an event ϕ, the decision of whether $(P, N, A, T) \models \phi$ holds can be determined in polynomial time.*

Proof: (1) Clearly calculating $\Phi(P, N, A, T)$ from (P, N, A, T) can be determined in polynomial time. (2) Calculating whether $\Phi(P, N, A, T) \vdash \phi$ holds can be determined in polynomial time since

A	B	C
+	+	+
+	0	+
+	−	+
0	+	+
0	0	0
0	−	0
−	0	−
−	−	−

Table 2: An incomplete resolution table: There is a row missing where `A` has truth value − and `B` has truth value +.

A	B	C
+	+	+
+	0	+
+	−	+
+	−	0
0	+	+
0	0	0
0	−	0
−	+	−
−	0	−
−	−	−

Table 3: An inconsistent resolution table: There are two rows where `A` has truth value + and `B` has truth value −, and `C` has different truth-values in them.

$\Phi(P, N, A, T)$ is a set of definite clauses and, by [9], there is a linear-time algorithm for theorem proving with definite clauses. Therefore the time taken in (1) plus (2) is polynomial, and it is an upper limit on the time taken on the decision of whether $(P, N, A, T) \models \phi$ holds.

Whilst we argue that causal maps are lucid and that they constitute a useful route to developing a knowledgebase for ramification analysis, there are some requirements that need to be addressed:

1. handling incompleteness in news reports.
2. using a more concise representation for resolution tables.
3. handling incompleteness in resolution tables.
4. using automated reasoning with tabulated causal maps.
5. increasing the expressibility by having the facility to add further conditions on causal relationships and confluences.

We address these requirements in the next section by using default logic.

3 Incorporating causal maps into default logic

A graphical representation can offer a lucid abstraction of a complex system. However, for ramification analysis, we want to automate the reasoning with causal maps. To do this we translate causal maps into knowledgebases of default logic. Essentially, we present a method to translate any general causal map into a default theory. This means we can then exploit a default theory as an efficient representation of a causal map and use default logic technology to automate reasoning with a causal map.

3.1 Reminder on default logic

As a basis of representing default knowledge we employ **default logic** originally proposed by Reiter [24]. Default logic is one of the best known and most widely studied formalisations of default reasoning [4, 5, 11, 1]. Furthermore, it offers a very expressive and lucid language. In default logic, knowledge is represented as a *default theory*, which consists of a set of first-order formulae and a set of *default rules* for representing default information. Default rules are of the following form, where α, β and γ are first-order (classical) formulae,

$$\frac{\alpha : \beta}{\gamma}$$

The inference rules are those of classical logic plus a special mechanism to deal with default rules: Basically, if α is inferred, and $\neg\beta$ cannot be inferred, then infer γ. For this, α is called the pre-condition, β is called the justification, and γ is called the consequent.

The set of formulae that are derivable from a default theory is called an extension. Each extension is a set of classical formulae. There may be more than one extension per default theory.

Default logic extends classical logic. Hence, all classical inferences from the classical information in a default theory are derivable (if there is an extension). The default theory then augments these classical inferences by default inferences derivable using the default rules.

The methods for obtaining an extension from a default theory is given in Definition 3.1. This definition introduces the operator Γ that indicates what conclusions are to be associated with a given set E of formulae, where E is some set of classical formulae. In other words, this definition determines whether E is an extension of the default theory.

Definition 3.1 *Let (D, W) be a default theory, where E is a set of classical formulae, D is a set of default rules and W is a set of classical formulae. Let Th be the function that for a set of formulae returns the set of classical consequences of those formulae. For this, $\Gamma(E)$ is the smallest set of classical formulae such that the following three conditions are satisfied.*

1. $W \subseteq \Gamma(E)$
2. $\Gamma(E) = Th(\Gamma(E))$
3. *For each default in D, where α is the pre-condition, β is the justification, and γ is the consequent, the following holds:*

$$\textit{if } \alpha \in \Gamma(E) \textit{ and } \neg\beta \notin E \textit{ then } \gamma \in \Gamma(E)$$

Once $\Gamma(E)$ has been identified, E is an extension of (D, W) iff $E = \Gamma(E)$.

We can view E as the set of formulae for which we are ensuring consistency with the justification of each default rule that we are attempting to apply. We can view $\Gamma(E)$ as the set of conclusions of a default theory: It contains W, it is closed under classical consequence, and for each default that is applicable (i.e. the precondition is in $\Gamma(E)$ and the justification is satisfiable with E), then the consequent is in $\Gamma(E)$. We ask for the smallest $\Gamma(E)$ to ensure that each default rule that is applied is grounded. This means that it is not the case that one or more default rules are self-supporting. For example, a single default rule is self-supporting if the pre-condition is satisfied using the consequent. The test $E = \Gamma(E)$ ensures that the set of formulae for which the justifications are checked for consistency coincides with the set of conclusions of the default theory.

So, if E is an extension, then the first condition ensures that the set of classical formulae W is also in the extension, the second condition ensures the extension is closed under classical consequence, and the third condition ensures that for each default rule, if the pre-condition is in the extension, and the justification is consistent with the extension, then the consequent is in the extension.

Example 3.1 *Let D be the following set of default rules:*

$$\frac{\texttt{company(x)} \wedge \texttt{loss(x,y)} \wedge \texttt{turnover(x,z)} \wedge \texttt{y} < \texttt{z} : \neg\texttt{new-venture(x)}}{\neg\texttt{financially-sound(x)}}$$

$$\frac{\texttt{company(x)} \wedge \texttt{debt(x,y)} \wedge \texttt{capitalization(x,z)} \wedge \texttt{z} < \texttt{y} : \neg\texttt{takeover-target(x)}}{\neg\texttt{financially-sound(x)}}$$

$$\frac{\texttt{company(x)} \wedge \texttt{sector(x,y)} \wedge \texttt{growthsector(x,z)} : \texttt{financially-sound(x)}}{\texttt{good-buy(x)}}$$

$$\frac{\texttt{company(x)} \wedge \texttt{sector(x,y)} \wedge \texttt{sector-many-mergers(x,z)} : \texttt{financially-sound(x)}}{\texttt{good-buy(x)}}$$

If W is

{`company(Talk)`, `sector(Talk,telecoms)`, `growthsector(telecoms)`, `new-venture(Talk)`}

there is one extension from (D, W) which is $Th(W \cup \{$`good-buy(Talk)`$\})$.

We have chosen default logic as the formalism for ramification analysis because it is one of the most well-explored formalisms for default knowledge. There is a range of useful variants, and inferencing technology is being developed (for a review see [25]). Last, but not least, default logic offers a natural and straightforward route for developing ramification analysis knowledgebases.

3.2 Translating causal maps into default logic

In section 2.3, we identified a number of requirements for using causal maps with resolution tables. In this section, we address the requirements for proof theory for reasoning with causal maps, for supporting automated reasoning, and for handling incomplete resolution tables. We do this by translating each tabulated causal map into a default theory.

Definition 3.2 *The* **map translation** *is a function Ψ that takes a coherent tabulated causal map (N, A, T), and a news report P, and returns the smallest default theory (D, W), satisfying the following conditions. Note, we do not assume that the elements of T are complete or consistent.*

1. *If $\phi \in P$, then $\phi \in W$.*
2. *For every event class α, if $\alpha(+) \notin P$, and $\alpha(-) \notin P$, and there is no β such that $(\oplus(\beta, \alpha) \in A$ or $\otimes(\beta, \alpha) \in A)$, then $\alpha(0) \in W$.*
3. *For every $\oplus(\alpha, \beta) \in A$, such that β is not a head, there are the following default rules in D.*

$$\frac{\alpha(+) : \beta(+)}{\beta(+)} \qquad \frac{\alpha(-) : \beta(-)}{\beta(-)} \qquad \frac{\alpha(0) : \beta(0)}{\beta(0)}$$

4. *For every $\otimes(\alpha, \beta) \in A$, such that β is not a head, there are the following default rules in D.*

$$\frac{\alpha(-) : \beta(+)}{\beta(+)} \qquad \frac{\alpha(+) : \beta(-)}{\beta(-)} \qquad \frac{\alpha(0) : \beta(0)}{\beta(0)}$$

5. *For each head γ, where T_γ is the resolution table for γ, and for each tuple $(x_1, .., x_n, x_{n+1}) \in T_\gamma$, where the event class for x_1 is α_1 and ,..,and the event class for x_n is α_n, there is the following default rule in D.*

$$\frac{\alpha_1(x_1) \wedge ... \wedge \alpha_n(x_n) : \gamma(x_{n+1})}{\gamma(x_{n+1})}$$

6. *For every event class α, the following formulae are in W:*

$$\alpha(+) \rightarrow \neg\alpha(0) \wedge \neg\alpha(-)$$
$$\alpha(0) \rightarrow \neg\alpha(+) \wedge \neg\alpha(-)$$
$$\alpha(-) \rightarrow \neg\alpha(+) \wedge \neg\alpha(0)$$

The resulting default theory is a normal default theory.

Example 3.2 *From Table 2, $N = \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$, and $A = \{\otimes(\mathtt{A}, \mathtt{C}), \oplus(\mathtt{B}, \mathtt{C})\}$, we have the default theory where W is*

$$\begin{array}{lll} \mathtt{A}(+) \rightarrow \neg\mathtt{A}(0) \wedge \neg\mathtt{A}(-), & \mathtt{A}(0) \rightarrow \neg\mathtt{A}(+) \wedge \neg\mathtt{A}(-), & \mathtt{A}(-) \rightarrow \neg\mathtt{A}(+) \wedge \neg\mathtt{A}(0) \\ \mathtt{B}(+) \rightarrow \neg\mathtt{B}(0) \wedge \neg\mathtt{B}(-), & \mathtt{B}(0) \rightarrow \neg\mathtt{B}(+) \wedge \neg\mathtt{B}(-), & \mathtt{B}(-) \rightarrow \neg\mathtt{B}(+) \wedge \neg\mathtt{B}(0) \\ \mathtt{C}(+) \rightarrow \neg\mathtt{C}(0) \wedge \neg\mathtt{C}(-), & \mathtt{C}(0) \rightarrow \neg\mathtt{C}(+) \wedge \neg\mathtt{C}(-), & \mathtt{C}(-) \rightarrow \neg\mathtt{C}(+) \wedge \neg\mathtt{C}(0) \end{array}$$

and D is,

$$\frac{\mathtt{A}(+) \wedge \mathtt{B}(+) : \mathtt{C}(+)}{\mathtt{C}(+)}, \frac{\mathtt{A}(+) \wedge \mathtt{B}(0) : \mathtt{C}(+)}{\mathtt{C}(+)}, \frac{\mathtt{A}(+) \wedge \mathtt{B}(-) : \mathtt{C}(+)}{\mathtt{C}(+)}$$

$$\frac{\mathtt{A}(0) \wedge \mathtt{B}(+) : \mathtt{C}(+)}{\mathtt{C}(+)}, \frac{\mathtt{A}(0) \wedge \mathtt{B}(0) : \mathtt{C}(0)}{\mathtt{C}(0)}, \frac{\mathtt{A}(0) \wedge \mathtt{B}(-) : \mathtt{C}(0)}{\mathtt{C}(0)}$$

$$\frac{\mathtt{A}(-) \wedge \mathtt{B}(0) : \mathtt{C}(-)}{\mathtt{C}(-)}, \frac{\mathtt{A}(-) \wedge \mathtt{B}(-) : \mathtt{C}(-)}{\mathtt{C}(-)}$$

Example 3.3 *From Table 3, $N = \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$, and $A = \{\otimes(\mathtt{A}, \mathtt{C}), \oplus(\mathtt{B}, \mathtt{C})\}$, we have the default theory given in Example 3.2 together with the following two default rules.*

$$\frac{\mathtt{A}(-) \wedge \mathtt{B}(+) : \mathtt{C}(-)}{\mathtt{C}(-)}, \frac{\mathtt{A}(+) \wedge \mathtt{B}(-) : \mathtt{C}(0)}{\mathtt{C}(0)}$$

Now resolution tables are represented as default rules, we can economize on the the number and form of rules, by taking a logically equivalent but more economical set of rules. For example, if we have the following three rules:

$$\frac{\mathtt{A}(+) \wedge \mathtt{B}(+) : \mathtt{C}(+)}{\mathtt{C}(+)}, \frac{\mathtt{A}(+) \wedge \mathtt{B}(0) : \mathtt{C}(+)}{\mathtt{C}(+)}, \frac{\mathtt{A}(+) \wedge \mathtt{B}(-) : \mathtt{C}(+)}{\mathtt{C}(+)}$$

We can represent these rules by the logically equivalent rule:

$$\frac{\mathtt{A}(+) : \mathtt{C}(+)}{\mathtt{C}(+)}$$

Example 3.4 *Consider the tabulated causal map, where the causal map is given in Figure 1, and the resolution table in Table 1.*

$$\frac{\texttt{interest-rate(-)} : \texttt{exchange-rate(-)}}{\texttt{exchange-rate(-)}}$$

$$\frac{\texttt{interest-rate(+)} \wedge \texttt{trade-surplus(+)} : \texttt{exchange-rate(+)}}{\texttt{exchange-rate(+)}}$$

$$\frac{\texttt{interest-rate(+)} \wedge (\texttt{inflation(0)} \vee \texttt{inflation(-)}) : \texttt{exchange-rate(+)}}{\texttt{exchange-rate(+)}}$$

$$\frac{\texttt{interest-rate(+)} \wedge \texttt{inflation(+)} \wedge (\texttt{trade-surplus(0)} \vee \texttt{trade-surplus(-)}) : \texttt{exchange-rate(0)}}{\texttt{exchange-rate(0)}}$$

$$\frac{\texttt{interest-rate(0)} \wedge \texttt{inflation(+)} \wedge \texttt{trade-surplus(+)} : \texttt{exchange-rate(0)}}{\texttt{exchange-rate(0)}}$$

$$\frac{\texttt{interest-rate(0)} \wedge \texttt{inflation(0)} \wedge \neg\texttt{trade-surplus(-)} : \texttt{exchange-rate(0)}}{\texttt{exchange-rate(0)}}$$

$$\frac{\texttt{interest-rate(0)} \wedge \neg\texttt{inflation(-)} \wedge \texttt{trade-surplus(-)} : \texttt{exchange-rate(-)}}{\texttt{exchange-rate(-)}}$$

$$\frac{\texttt{interest-rate(0)} \wedge \texttt{inflation(-)} \wedge \texttt{trade-surplus(-)} : \texttt{exchange-rate(0)}}{\texttt{exchange-rate(0)}}$$

$$\frac{\texttt{interest-rate(0)} \wedge \texttt{inflation(-)} \wedge \neg\texttt{trade-surplus(-)} : \texttt{exchange-rate(+)}}{\texttt{exchange-rate(+)}}$$

$$\frac{\texttt{interest-rate(0)} \wedge \texttt{inflation(+)} \wedge \texttt{trade-surplus(0)} : \texttt{exchange-rate(-)}}{\texttt{exchange-rate(-)}}$$

Here we consider some of the properties of causal maps in default logic.

Proposition 3.1 *If h is a model of (P, N, A, T) and $\Psi(P, N, A, T) = (D, W)$, there is an E such that E is an extension of (D, W), and $h(\alpha) = x$ iff $\alpha(x) \in E$.*

Proof: ($\Rightarrow$) Suppose $h(\alpha) = +$. The argument is the same for $h(\alpha) = 0$, and $h(\alpha) = -$. $h(\alpha) = +$ holds as a result of Condition 1, 4, 6 or 9 of Definition 2.19. If Condition 1, then $\alpha(+) \in P$. Hence, $\alpha(+) \in W$, and so $\alpha(+) \in E$. If Condition 4, then $\exists \oplus(\beta, \alpha) \in A$, and α is not a head and $h(\beta) = +$ and $\alpha(0) \notin P$ and $\alpha(-) \notin P$. Hence, there is a default rule in D:

$$\frac{\beta(+) : \alpha(+)}{\alpha(+)}$$

and $\alpha(0) \notin W$ and $\alpha(-) \notin W$, and by recursion $\beta(+) \in E$, and so $\alpha(+) \in E$. If Condition 6, then the argument is the same as for the previous case. If Condition 9, then there is a default rule in D,

$$\frac{\alpha_1(x_1) \wedge ... \wedge \alpha_n(x_n) : \gamma(x_{n+1})}{\gamma(x_{n+1})}$$

where $\alpha_1(x_1), ..., \alpha_n(x_n) \in E$ by recursion, and so $\alpha(+) \in E$. The argument for ($\Leftarrow$) follows similarly.

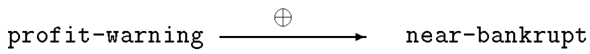

Figure 7: A confluence for `near-bankrupt`.

Proposition 3.2 *If* $\Psi(P, N, A, T) = (D, W)$ *and* E *is an extension of* (D, W)*, there is an* h *such that* h *is a model of* (P, N, A, T) *and* $h(\alpha) = x$ *iff* $\alpha(x) \in E$.

Proof: Analogous to Proposition 3.1.

Proposition 3.3 *For a tabulated causal map* (N, A, T)*, where* (N, A) *is coherent and each* $T_\gamma \in T$ *is consistent and complete, a non-contradictory news report* P*, and an event* ϕ*,*

$$(P, N, A, T) \models \phi \textit{ iff } \quad \exists! E \textit{ such that } E \textit{ is an extension of } (D, W) \textit{ and } \Phi(P, N, A, T) = (D, W) \textit{ and } \phi \in E$$

Proof: ($\Rightarrow$) Suppose ϕ is $\alpha(+)$ (the case for $\alpha(0)$ and $\alpha(-)$ follows similarly). From Proposition 2.5, for a tabulated causal map (N, A, T), where (N, A) is coherent and each $T_\gamma \in T$ is consistent and complete, and a non-contradictory news report P, there is a unique h such that h is a complete model of (P, N, A, T), and $h(\alpha) = +$. Therefore, by Proposition 3.1 and Proposition 3.2 there is a unique E such that E is an extension of (D, W), and $\phi \in E$. ($\Leftarrow$) Follows similarly.

So by translating tabulated casusal maps into default logic, we can handle incomplete and inconsistent resolution tables. An inconsistent resolution table is translated into a default theory with conflicting defaults which if applied would result in multiple extensions. An incomplete resolution table requires further default rules. If $(x_1, .., x_n, x_{n+1})$ is a row missing from a resolution table, where x_{n+1} is a null value, then one option is to have three defaults — the first default has consequent $\gamma(+)$, the second $\gamma(0)$, and the third $\gamma(-)$. We consider more sophisticated options in the next section.

3.3 Adding further constraints

In the previous section, we addressed some of the requirements raised in Section 2.3. Here we address further requirements raised in Section 2.3. In particular, increasing the expressibility of causal maps by having the facility to add further conditions on causal relationships and confluences.

For example, suppose we have an investment in the fictional regional airline Happy Air, and we receive the news, that due to the unexpected sharp rise in fuel prices, Happy Air has issued a profits warning. Also suppose that you have no other knowledge that would lead you to believe that Happy Air is going bankrupt. In addition, suppose that when airlines issue profits warnings due to increased costs, they do not usually go bankrupt. From this background knowledge, we can evaluate the ramifications of this news including (1) Happy Air will go bankrupt soon, and (2) Happy Air will not go bankrupt soon. Ramification 1 is more likely than ramification 2. But the less likely ramification is more significant, as the investment in Happy Air would then be lost.

From the causal map, we have a very general statement, which can be represented by:

$$\frac{\texttt{profit-warning(+)} : \texttt{near-bankrupt(+)}}{\texttt{near-bankrupt(+)}}$$

$$\frac{\texttt{profit-warning(-)} : \texttt{near-bankrupt(-)}}{\texttt{near-bankrupt(-)}}$$

However, we may regard this default rule as too general. We can address this by adding further constraints on it, and using the new rule instead,

$$\frac{\begin{array}{l}\texttt{company(X)}\\ \wedge\ \texttt{profit-warning(+)}\\ \wedge\ \texttt{liquid-capital(X,P)}\\ \wedge\ \texttt{monthly-expenses(X,Q)}\\ \wedge\ \texttt{P} > 2 \times \texttt{Q} \qquad : \neg\texttt{unsound(X)}\end{array}}{\neg\texttt{near-bankrupt(+)}}$$

together with the following information,

$$\forall \texttt{X}\ \texttt{regional-airline(X)} \rightarrow \texttt{airline(X)}$$
$$\forall \texttt{X}\ \texttt{airline(X)} \rightarrow \texttt{company(X)}$$

Suppose, we also have the following facts.

```
profit-warning(Happy-Air)
unexpected-fuel-costs(Happy-Air)
regional-airline(Happy-Air)
liquid-capital(Happy-Air,$500K)
monthly-expenses(Happy-Air,$200K)
```

From the facts and general knowledge, we cannot drive the inference `unsound(Happy-Air)`. Hence, we can apply the default and so derive ¬`near-bankrupt(+)`.

In this way, we can view a causal map or tabulated causal map as an abstraction of a default theory, and so use it to aid explanation. Or equivalently, we can regard a causal map or tabulated map as an intermediate step in developing a default theory, and so we can regard it as an knowledge engineering technique.

4 Integration with language engineering

Language engineering is concerned with technologies for handling, understanding, and generating, information in the forms of natural langauge. It includes areas ranging information retrieval, information filtering, natural language understanding, natural language generation, and machine translation.

Of particular relevance here is the area of language engineering called information extraction. This area is concerned with transforming natural language text into a reduced and refined form,

companies	Marsam Pharmaceuticals Inc
	Schien Pharmaceuticals Inc
values	240 million dollars
	21 dollars per share
	19 dollars per share

Table 4: An instantiated `Template` for `Summary`

company-target	Marsem-Pharmaceuticals-Inc.
company-predator	Schein-Pharmaceuticals-Inc.
type-of-takeover	Friendly
value	240-million-dollars

Table 5: An instantiated `Template` for `Takeover`

highlighting pertinent points, and removing superfluous information [8, 14, 2]. In this section, we consider integrating ramification analysis with information extraction, by feeding extracted information directly into a ramification analysis knowledgebase.

As an example of an information extraction system, we consider LOLITA (Large-scale Object-based Linguistic Interactor Translator and Analyser), a system based on a general purpose natural language processing system, that has been applied in the financial domain [7, 6]. In this system, a short summary is derived from each financial news article according to specific criteria. This approach uses templates that incorporate a predefined number and type of slots and each slot can be filled with information extracted from the news article. This provides a schematic representation that is based on key types of financial activity. There are many possible templates and they can use many different types of slot. More than one template can be produced for a source article. For example, consider the templates `summary` and `takeover`. The first has slots `companies` and `values` and the second has slots `company target`, `company-predator`, `type-of-takeover`, and `value`. Below are instantiations of each of these templates derived from the same article.

Completing templates relies on the knowledgebase of the underlying natural language system LOLITA. The knowledgebase incorporates a 100,000 node semantic network. An article is processed by relating concepts in the article to the semantic network. Each template is then completed by searching for relevant information in the semantic network.

Now we can feed the output from a system, such as LOLITA, into ramification analysis. For this, we need rules to identify event classes and determine correct event values for each of these event classes. Suppose, we are interested in taking a snapshot of `interest-rate`, `trade-surplus`, and `inflation`, and want to predict the ramifications of any changes in these values on the `exchange-rate`. We take a template for these values at each point in time, say each month, and use the causal map in Figure 1.

Example 4.1 *From looking at the change from Table 6 to Table 7, we obtain the following events for month 2.*

$$\{\texttt{interest-rate(+)}, \texttt{trade-surplus(-)}, \texttt{inflation(+)}\}$$

We can now use these events with the default theory outlined in Section 3.2.

`interest-rate`	7%
`trade-surplus`	$ 3 Billion
`inflation`	4%

Table 6: An instantiated `Template` for `Economics` for `month 1`

`interest-rate`	8%
`trade-surplus`	$ 2 Billion
`inflation`	6%

Table 7: An instantiated `Template` for `Economics` for `month 2`

We can further develop this approach with a lexical knowledgebase. Lexical knowledge is increasingly important in language engineering. Of particular interest here is semantic knowledge about words. For example, if we know that some news is about `oil`, it is usually reasonable to derive that it is about `petroleum`, with exceptions such as in contexts about `cooking`. Lexical knowledge can also facilitate in the identification of synonyms, related terms, antinyms, and specializations for a word. It can also be used to identify meronymic relations, such as `engine` is `part-of` a `car`, and parts-of-speech such as relating actors with actions: For example, for the actor `terrorist` an appropriate action is `terrorism`.

Proposals for handling lexical knowledge in default logic (see for example [15, 16, 17]) could be incorporated to bridge the differences in the use of words in information extraction and the use of words in ramification analysis. For example, synonyms of the words appearing in the output from information extraction could be flagged so that the ramification analysis is less likely to be sensitive to the actual predicate symbols used for event classes.

5 Conclusions

In this paper we have provided: a discussion of the importance of ramification analysis for agents handling news, where agents need to determine likely outcomes of news and less likely, but very significant, outcomes of news; an introduction to causal maps as a formalism for representing and reasoning with ramifications of events, including proof theory and semantics; a translation of causal maps into default logic in order to exploit the theoretical and technological advantages of default logic; and a discussion of the integration of ramification analysis with natural language engineering.

Causal mapping can also be viewed as a knowledge engineering technology for default logic. Modelling real-world systems in terms of existential and directional events, and causal maps, may be intuitive to a wider audience than using default logic directly. Also, it opens the opportunity for developing graphical tools to engineering and managing this kind of knowledgebase.

In this way, causal maps provide an abstract view on what may be a complex knowledgebase, and hence, causal maps can be used as the first stage in a two stage process of engineering knowledgebases for ramification analysis. Potentially automated techniques based on machine learning (such as inductive logic programming) can be harnessed to generate causal maps from sets of examples of news reports with associated ramifications. Then by hand, the casual maps can be refined as default logic knowledgebases.

As well as facilitating knowledge acquisition, causal maps offer an appropriate way of explaining news reports and their associated ramifications. Think of an article in a newspaper such as *the Economist*. Often to make a point, or to explain particular events, an article will consider just a small proportion of the possible causes and effects of an event. Rather than try to provide very comprehensive coverage, they are selective in each article. In the same way, causal maps can be used to simplify explanations.

An informal modelling technique, that is related to causal maps, is cognitive maps. In a cognitive map, a directed labelled graph is used to capture the structure of a decision-maker's stated beliefs about a particular problem [3]. Some cognitive maps can be used for a form of ramification analysis.

Uncertainty and causality are handled in a range of other approaches. In Bayesian networks (for a review see [18]), a directed acyclic graph is used to represent causal relations between random variables. These causal relations are used to identify independence assumptions between random variables to facilitate more efficient representation and reasoning with conditional probabilities. Whilst in a sense, Bayesian networks provide a form of ramification analysis, the nature of the ramification analysis is fundamentally different from that given by causal mapping. In particular, they provide a means for looking at the ramifications of the change in random variables as dictated by the probability distribution and the axioms of probability theory.

Qualitative probabilistic networks are a qualitative form of probabilistic network [27] (for a review see [20]). Reasoning in qualitative probabilistic networks is qualitatively dictated by a qualitative probability distribution and the axioms of probability theory, and so it is also fundamentally different from reasoning with causal maps. However, in [13] there is a limited formalization of cognitive maps.

Another approach to handling uncertainty and causality is possibilistic networks. In possibilistic networks (for a review see [12]), a directed acyclic graph is used to represent casual relations between possibilistic variables. Possibilistic networks provide a form of ramification analysis, though again the nature of the ramification analysis is fundamentally different from that given by causal mapping. In particular, they provide a means for looking at the ramification of the change in possibilistic variables as dictated by the possibility distribution and the axioms of possibility theory.

Finally, causal mapping is complementary to a variety of approaches to argumentation (for examples of formalisms for argumentation see [21, 10, 22, 19, 26, 23]). Indeed, causal mapping may be viewed as a form of argumentation.

Acknowledgements

The author would like to thank the anonymous referees for some useful feedback on a previous version of this paper.

References

[1] G Antoniou. *Non-monotonic Reasoning*. MIT Press, 1997.

[2] ARPA. *Message Understanding Conference: Proceedings of the Seventh Conference*. Morgan Kaufmann, 1998.

[3] R Axelrod, editor. *Structure of Decisions: The Cognitive Maps of the Political Elites*. Princeton University Press, 1976.

[4] Ph. Besnard. *An Introduction to Default Logic*. Springer, 1989.

[5] G Brewka. *Common-sense Reasoning.* Cambridge University Press, 1991.

[6] M Costantino, R Collingham, and R Morgan. Natural language processing in finance: Automatic extraction of information from financial news articles. *The Magazine of Artificial Intelligence in Finance*, 2, 1995.

[7] M Costantino, R Collingham, and R Morgan. Information extraction in the LOLITA system using templates from financial news articles. In *Proceedings of the 18th International Conference on Information Technology Interface*, 1996.

[8] J Cowie and W Lehnert. Information extraction. *Communications of the ACM*, 39:81–91, 1996.

[9] W Dowling and J Gallier. Linear-time algorithms for testing the satisfiability of propositional horn formulae. *Journal of Logic Programming*, 3:267–284, 1984.

[10] M Elvang-Goransson and A Hunter. Argumentative logics: Reasoning from classically inconsistent information. *Data and Knowledge Engineering Journal*, 16:125–145, 1995.

[11] D. Gabbay, C. Hogger, and J. Robinson, editors. *Handbook of Logic in Artificial Intelligence and Logic Programming: Volume 3 — Non-monotonic Reasoning and Uncertainty Reasoning.* Oxford University Press, 1994.

[12] J Gebhardt and R Kruse. Background and perspectives of possibilistic graphical models. In A Hunter and S Parsons, editors, *Applications of Uncertainty Formalisms*, Lecture Notes in Computer Science. Springer, 1998.

[13] H Geffner. A formal framework for causal modelling and argumentation. In D Gabbay and H-J Ohlbach, editors, *Practical Reasoning*, number 1085 in Lecture Notes in Computer Science. Springer, 1996.

[14] R Grishman. Information extraction techniques and challenges. In M Pazienza, editor, *Information extraction.* Springer, 1997.

[15] A Hunter. Intelligent text handling using default logic. In *Proceedings of the Eighth IEEE International Conference on Tools with Artificial Intelligence (TAI'96)*, 1996.

[16] A. Hunter. Using default logic for lexical knowledge. In *Qualitative and Quantitative Practical Reasoning*, volume 1244 of *Lecture Notes in Computer Science*, pages 86–95. Springer, 1997.

[17] A Hunter. Context-dependent reasoning with lexical knowledge using default logic. Technical report, Department of Computer Science, University College London, 1998.

[18] F Jenssen. *An Introduction to Bayesian Networks.* UCL Press, 1996.

[19] S Parsons. Defining normative systems for qualitative argumentation. In D Gabbay and H-J Ohlbach, editors, *Practical Reasoning*, volume 1085 of *Lecture Notes in Computer Science*, 1996.

[20] S Parsons. *Qualitative Methods for Reasoning under Uncertainty.* MIT Press, 1999.

[21] H Prakken. An argument framework for default reasoning. In *Annals of mathematics and artificial intelligence*, volume 9, 1993.

[22] H Prakken. *Logical Tools for Modelling Legal Argument.* Kluwer, 1997.

[23] H Prakken and G Vreeswijk. Modelling argumentation in logic. In D Gabbay, editor, *Handbook of Philosophical Logic.* Kluwer, 1999.

[24] R. Reiter. Default logic. *Artificial Intelligence*, 13:81–132, 1980.

[25] T Schaub. *The Automation of Reasoning with Incomplete Information*, volume 1409 of *Lecture Notes in Artificial Intelligence*. Springer, 1998.

[26] D Vermier, E Laenens, and P Geerts. Defeasible logics. In *Handbook of Defeasible Reasoning and Uncertainty Management*, volume 3. Kluwer, 1998.

[27] M Wellman. Qualitative probabilisitc networks for planning under uncertainty. In J Lemmer and L Kanal, editors, *Uncertainty in Artificial Intelligence 2*. Elsevier, 1988.