Chapter 1

Introduction

In this book, we aim to lay bare the logical foundations of tractable reasoning. We draw on Marvin Minsky’s seminal work on frames, which has been highly influential in computer science and, to a lesser extent, in cognitive science. However, only very few attempts have explored ideas about frames in logic. Such is the intended innovation of the present investigation.

In cognitive science, the idea that the human mind works in a modular fashion continues to be very popular. This idea is also referred to as the massive modularity hypothesis. The present logical investigation of modularity is directed toward two unresolved problems that have arisen in cognitive science and have yet to receive a proper solution. First, the problem of tractable reasoning, and second the problem of transmodular reasoning with a modular architecture. The first problem concerns the cognitive feasibility of global inferential reasoning: we know very little about the cognitive and logical means that allow us to draw inferences that involve larger amounts of language and memory. Global inferential reasoning appears to be intractable. Particularly challenging from a computational perspective is the operation of changing one’s beliefs in light of new epistemic input.

Ideas about a modular architecture of human cognition promise to explain how human minds cope with the challenges of computational complexity. However, cognitive scientists have not delivered a more detailed demonstration that modularity actually helps achieve tractability. This gives rise to the second problem: human reasoning is often transmodular and is therefore not confined to a small set of beliefs in a specific domain. A decomposition of reasoning into modular operations thus is needed to maintain the modularity hypothesis for central cognition.

The apparent cognitive and computational infeasibility of global inferential reasoning is a major objection to the computational theory of mind and the
language of thought hypothesis propounded by Fodor [50, 54]. But the problem is not tied to these two theoretical hypotheses: as soon as we acknowledge that human minds engage in drawing inferences, we face the problem of computational and cognitive feasibility. This problem does not result from a particular cognitive paradigm.

Why are belief changes computationally challenging, if not even computationally infeasible? Why do these computational issues matter for our understanding of human cognition? Why should a modular account of human reasoning, in terms of frames and frame concepts, help achieve tractability? In what follows, we shall outline an answer to these questions.

### 1.1 Apparent Intractability of Belief Changes

In science and everyday life, our beliefs are changing continuously. Belief changes are often initiated by a new piece of information. More specifically, we can distinguish between two types of impact that a new piece of information may have on our present beliefs: first, the new epistemic input may allow us to infer further consequences from our present beliefs. Second, we may be forced to give up some of our present beliefs because the new epistemic input is inconsistent with what we presently believe.

This simplified variant of a story by Gärdenfors [56, p. 1] exemplifies how new epistemic input may force us to give up some of our present beliefs. Oscar used to believe that he had given his wife a ring made of gold at their wedding. Later, he realises that his wedding ring has been stained by sulfuric acid. However, he remembers from high school chemistry that sulfuric acid does not stain gold. As he could not deny that the ring is stained and as he also believed that their wedding rings are made of the same material, his beliefs implied a contradiction. Hence, the new epistemic input – i.e., the observation that a certain liquid (believed to be sulfuric acid) stains his wedding ring – forces Oscar to give up some of his present beliefs. From a logical point of view, there are at least three options to regain consistency: (i) he could retract the belief that sulfuric acid stained the ring, (ii) he could call into question that sulfuric acid does not stain gold, and finally, (iii) he could give up the belief that their wedding rings are made of gold. As he paid a somewhat lower price for their wedding rings than might normally be expected, he found himself forced to accept the third option.

The logical study of belief changes has given rise to a new discipline in philosophical logic called belief revision theory. For the most part, this discipline has been founded by Alchourrón, Gärdenfors and Makinson [1]. It is for this
1.1. APPARENT INTRACTABILITY OF BELIEF CHANGES

reason also referred to as AGM, or AGM-style, belief revision theory. The original AGM theory consists in an account of the laws and the semantics of belief changes. A large number of further belief revision schemes have been devised in the wake of the AGM account. Each of these comes with a specific analysis of belief changes.

Why should it be computationally and cognitively impossible to perform belief changes in a rational way? In brief, the answer is that the amount of calculation demanded for a single belief change is likely to exceed the cognitive power of our mind. Theoretical arguments showing this come from the theory of computational complexity and related complexity results for standard, logical approaches to belief revision [83]. These results strongly suggest that belief revision is an intractable problem. In more technical terms: if $P \neq NP$ (which is quite firmly believed by most computer scientists), then the problem of rationally revising one’s beliefs is intractable. We shall explain the computational complexity of belief changes in greater detail in the course of this investigation.

Intractability of a problem means, in less technical terms, that our best computers take too much time to solve non-trivial instances of the problem, despite all the progress in the development of hardware that we have witnessed in past decades. To give an example, no efficient method has been found for determining the satisfiability of a propositional formula. For, the number of relevant valuations grows exponentially with the size of such a formula so that, in the worst case, there is an exponential growth of the number of computation steps needed to test satisfiability.

Cherniak [34, p. 93] has given a telling illustration of how devastating this type of exponential growth is for the feasibility of determining whether or not a propositional formula (or a set of such formulas) is satisfiable: suppose our super computer is able to check whether a given propositional valuation verifies the members of a set of propositions in the time that light takes to traverse the diameter of a proton. Assume, furthermore, that our propositional belief system contains just 138 logically independent propositions. If so, the estimated time from the Big Bang to the present would not suffice to go through all valuations of the system of propositions. Hence, even this super computer would not provide us with the means to test consistency of a moderately complex propositional belief system in such a manner that the test is conclusive for any possible input. We might of course be lucky and find a propositional valuation that verifies all items of the belief system after going through only a few valuations, but it is much more probable that we will be unlucky.

Exponential growth of the computation steps needed to solve a problem,
in the worst case, is considered a mark of intractability. Tractability of a problem, by contrast, means that it can be solved with a reasonable number of computation steps. To be more precise, a problem is tractable if and only if the number of computation steps needed to solve an arbitrary instance of it is bounded by a polynomial function whose argument is a parameter that characterises the syntactic size of the instance. The computation of arithmetic functions, such as addition, multiplication, etc., for natural and rational numbers is tractable because corresponding algorithms exhibit no exponential growth of the number of computation steps.

The problem of revising a belief system is computationally even harder than the problem of testing satisfiability of a propositional formula. Belief revision must therefore be considered an intractable problem, at the present state of the art. As Nebel [83, p. 121] puts it, “The general revision problem for propositional logic appears to be hopelessly infeasible from a computational point of view because they are located on the second level of the polynomial hierarchy.” It is fair to say that there is no reasonably expressive and tractable belief revision scheme to be found in the computer science literature to date.\footnote{See again Nebel [83], who has given the most comprehensive survey of the complexity results of belief revision schemes. Those belief revision schemes that yield, together with a tractable logic, a tractable determination of belief changes rest on counterintuitive assumptions about the epistemic ranking of beliefs. We shall be more explicit about this in Chapter 7.}

Why is the problem of rational belief revision even harder than testing satisfiability of a propositional formula? The runtime of algorithms that determine belief changes exactly is not only exponential, but super-exponential. That is, these algorithms require exponentially many calls of a subroutine that consists of exponentially many computation steps. Few attempts have been made toward approximate solutions to the general problem of belief revision.\footnote{See [105] and [2]. We shall discuss these proposals in Chapter 7.} And the potentialities of a modular approach to tractable reasoning and belief revision have not yet been fully exploited.

\section{1.2 Computational Theory of Mind}

The distinction between tractable and intractable problems has a certain bearing on the understanding of human cognition, as observed by Cher- niak [34], Fodor [53], Johnson-Laird [69], Stenning and van Lambalgen [98], Woods [108] and others. Human minds are far more creative than computers and for this reason are certainly distinct from the latter. But they are not
capable of executing algorithms and computations faster and more reliably than computers. As for belief revision, there is a large class of problems that appear to require formal reasoning and calculation more than creativity.

In cognitive science, the *computational theory of mind* (CTM) is a research programme that aims to exploit presumed commonalities between computers and the biological machinery underlying our cognition. It follows quite directly from CTM that any problem that is intractable for computers is intractable for human minds as well. As we shall see later on, rather weak formulations of CTM suffice to show this implication. In particular, CTM implies that we are not capable of obeying the norms of standard approaches to rational belief revision. We do not even have a clear idea of how we could approximately conform to the norms of rational belief change.

Modularity, on the other hand, appears to be a means to escape the combinatorial explosion of belief formation. This is so, however, only if we can decompose global reasoning into modular operations. For, belief formation is often a global affair in the sense that new pieces of information potentially impact a large variety of beliefs across a number of different domains. In molecular biology, for example, a new finding about the homology of two protein sequences may imply hypotheses about their functional similarity, which in turn may be relevant to our theories about specific biochemical pathways. Even in daily life it happens that new information has an impact across domains. Last winter, I met a passionate skier who sold his property at the local ski hill because of a few mediocre winters in a row and because of scientific evidence for global warming.

We are unable, in general, to tell in advance for which beliefs a new epistemic input is relevant. The modularity hypothesis, by contrast, asserts that cognitive modules work in a domain-specific way and are encapsulated. In order to maintain the modularity of central cognition, one therefore must resort to an account of interacting modular units of reasoning [33]. This idea, however, has not yet been formulated in a precise manner. We lack, in particular, something like a logical analysis of modular reasoning. Cognitive scientists have not yet been able to prove that a modular account of human cognition explains how global inferential reasoning is cognitively feasible.\textsuperscript{3}

At the same time, considerations of tractability have been used to advance the massive modularity hypothesis in cognitive science. The presumed intractability of an a-modular account of human cognition serves as one of three core arguments in favour of this hypothesis [39].

In a similar vein, Fodor [53] has observed an impasse in cognitive science: the

\textsuperscript{3}The famous work on bounded rationality by Gigerenzer [60, 59] is not concerned with modularity.
undeniable existence of global cognition, i.e., cognition that involves larger amounts of language and knowledge, is fundamentally at odds with the computational theory of mind, given the limited success of logic-oriented artificial intelligence and the difficulties of computing global inferences reliably and fast. Fodor observes that we simply have no idea how global cognition is feasible for minds with finitely bounded resources of computation and memory (see [53, Ch. 2 and 3] and [54, Ch. 4.4]).

Who is the culprit for this impasse? On the face of it, there appear to be only two candidates: CTM and the view that formal logic has some role to play in an account of human cognition. So, shall we discard CTM or some variant of it? Fodor [53] introduces CTM in the more specific sense that higher cognitive processes are classical computations, i.e., computations that consist of operations upon syntactic items. This understanding of CTM forms the core of the classical computationalist paradigm in cognitive science [19]. Broadly logic-oriented research on human cognition has been driven by this paradigm. Assuming a syntactic nature of computation appears to justify applying the computational complexity theory to human cognition in the first place. So it seems natural to discard CTM – more precisely, the syntactic formulation of CTM – in order to solve our problem. Computational complexity theory, however, pertains to human cognition quite independently of this assumption since the scope of complexity theory is governed by the Church-Turing thesis. This thesis asserts that any physically realisable computation device – whether it is based on silicon, neurones, DNA or some other technology – can be simulated by a Turing machine [14, p. 26]. Issues of computational complexity, therefore, pertain to human cognition quite independently of a commitment to a syntactic variant of CTM.

Shall we discard, then, logical approaches to human reasoning and cognition? The connectionist paradigm in cognitive science is motivated by the significant success of artificial neural networks in pattern recognition and the relatively minor success of good old-fashioned artificial intelligence, which is logic-oriented. Proponents of the connectionist paradigm have suggested that formal logic has hardly any explanatory value for human intelligence [37]. It is undeniable, however, that scientific and quotidian reasoning have a genuine propositional structure since reasoning and argumentation essentially consists in making inferential transitions from antecedently acquired or accepted propositions. It has been moreover shown that propositional reasoning can be implemented by means of neural networks [74], Stenning and van Lambalgen [98], Johnson-Laird [69] and others have successfully supplemented logical approaches to human cognition with empirical research in psychology.
One might also try to dissolve issues of tractability by emphasising the normative role of logic. Logical systems, one could argue, explicate norms of reasoning without actually describing human reasoning. But even a purely normative view of logic would not solve the problem under consideration. In order for a standard to have a normative role, it must be possible to, at least approximately, meet it. A norm that we cannot obey – neither exactly nor approximately – cannot be considered a norm in the first place. Ought implies Can. Following van Benthem [103], we view the role of logical systems in an analysis of human reasoning as partly normative and partly descriptive.

A closer look at our problem reveals more culprits to consider. There might be something wrong with the distinction between tractable and intractable problems in the theory of computational complexity. In particular, the decision to focus on worst-case scenarios in the original distinction is open to question. This line, however, is not pursued here.

Yet another option is to work on the cognitive adequacy of logical systems themselves. This is the strategy pursued here. We tackle the computational issues of belief revision using frames and frame concepts, and thus resume a research programme originated by Minsky [81]. Even though Minsky himself was rather hostile toward logic-oriented approaches to human cognition, there is a logical and set-theoretic core recognisable in his account of frames. We explicate and further develop this core using set-theoretic predicates in the tradition of Sneed [96] and Balzer et al. [17].

A note on the infamous frame problem is in order here. In an investigation of frames, one would expect to find a thorough discussion of this problem. Fodor [52, 53] makes much out of the frame problem, but is charged with not knowing ‘the frame problem from a bunch of bananas’ by Hayes [67]. In fact, the account that Fodor [52, 53] gives of the frame problem is a re-interpretation of the original frame problem as described by McCarthy and Hayes [77]. While there is some substantial connection between the original frame problem and Fodor’s re-interpretation, there is no need to discuss any variant of the frame problem in order to explain the computational challenges of dynamic, inferential reasoning. Likewise, there is no need to discuss abductive reasoning, as Fodor [53] does, for this purpose. Analysis of the computational complexity of belief revision gives us a more concise and less controversial exposition of the problem that dynamic, inferential reasoning is intractable in the setting of classical propositional logic. Minsky himself, in his seminal work on frames [81], makes no explicit reference to

\begin{enumerate}
\item Thanks to Hans Rott for this point.
\item Focusing on worst-case scenarios means that the computational complexity of a problem is determined by the maximal number of computation steps that are needed to solve any possible instance of the problem.
\end{enumerate}
the frame problem.

1.3 Frame Logic

Why may frame concepts help reduce the computational complexity of belief changes? Such concepts have a richer structure than ordinary concepts. A telling example used by Minsky [81, p. 47] is that of a child’s birthday party. Unlike an ordinary concept, this concept does not seem to apply well to a certain individual or tuples of individuals. We would not say that the concept in question applies to the birthday child, the union of birthday child and guests, or to the place where the party is given. What then are objects to which the concept of a birthday party is applied? Minsky says that it describes a situation that involves a number of different things: guests, games, presents, a birthday cake, a party meal, decor, etc. These things, normally, satisfy certain conditions: the guests are friends of the host, the games must be fun, the gifts must please the birthday child, etc.

From a logical point of view, a frame concept is a concept that applies to sequences of sets of objects as opposed to mere tuples of objects (which are not sets). Furthermore, frame concepts impose semantic constraints upon the (first-order) predicates of a small fragment of our language. For example, the guests of the birthday party are, normally, friends of the host. Frame concepts can therefore be used to interpret a piece of language in a small domain. This amounts to subdividing our global language into small sublanguages. These sublanguages, in turn, have different and yet interrelated interpretations in small subdomains. It is thus the notion of a frame concept by means of which we try to semantically explicate the notion of a cognitive module.

Modularising semantics in this way allows us to distinguish easily between intramodular and intermodular reasoning. The former type of reasoning is confined to a single module – which concerns the interpretation of a piece of language in a small domain – whereas the latter communicates information from one module to another. The distinction between intra- and intermodular reasoning gives rise to a proper logic of frames that emerges from our investigation. This logic is guided by the following two principles:

(1) Full first-order logic remains valid within the application of a frame concept.

(2) Only atomic sentences and negations thereof can be inferred from one application of a frame concept to another.
1.3. FRAME LOGIC

Restricting the scope of application of full first-order logic is key to reducing the computational complexity of global inferential reasoning. Inspired by methods of object-oriented programming, we furthermore allow compositions of frame concepts as well as compositions of compositions. Compositions, however, must be bounded in size so as to retain the tractability of frame logic. Note that object-oriented programming is a distinctive style of modular programming. It is commonly viewed as the greatest success story of Minsky’s methodology of frames in [81].

For frame logic to be devised and investigated, some formal work has to be done. First, we show how the notion of a frame concept can be formalised using set-theoretic predicates. This will allow us to give an axiomatic account of reasoning with frame concepts. Then, we merge the set-theoretic account of frames with some AGM-style belief revision scheme. The result is a belief revision theory with frame concepts. For this theory, we finally devise a truth maintenance system (TMS), i.e., an algorithm that determines how presently accepted truth values change upon new epistemic input. The TMS determines belief changes in a tractable manner.

The TMS is, furthermore, shown to serve as a powerful approximation of first-order reasoning, but it is not sound and complete with respect to finite, first-order logic. Soundness and completeness, however, can be achieved for finite frame logic. Likewise, the TMS yields only an approximate determination of belief changes, approximate with respect to standard approaches in the AGM tradition. We turn, however, the approximate nature of the TMS into a virtue by devising a belief revision scheme that mirrors the working of the TMS. This belief revision scheme is tractable for a logic with a tractable relation of logical consequence. Hence, it is tractable for frame logic. We speak of TM belief changes in order to refer to this novel belief revision scheme since it is inspired by truth maintenance.

Recall that some modification of an AGM-style approach to belief revision is necessary since a tractable logic alone does not suffice to resolve the computational issues of belief revision [83]. If we understand the notion of rational belief change in terms of AGM-style approaches, then NP ≠ P implies that there simply is no exact solution to the problem of tractable, rational belief revision. Approximations and computational simplifications of AGM-style approaches to belief revision are therefore of theoretical interest, at least from a cognitive point of view.

The semantics and the inference rules of frame logic, together with the TM belief revision scheme, are aimed at analysing our means of coping with the

---

*The belief revision scheme by Brewka [23] proved well suited for this purpose.*

*By finite frame logic, we mean frame logic confined to finite domains.*
computational and cognitive challenges of belief revision. While we do not claim that the present account is literally true in all respects, we think that this account makes significant progress toward a logical analysis of quotidian human reasoning that is cognitively plausible. Frame logic and the TM belief revision scheme are devised with the intent to pursue a major theme within the cognitive turn in logic.\footnote{Van Benthem \cite[p. 67]{103} observes that ‘modern logic is undergoing a cognitive turn’. Belief revision theory and the related dynamic epistemic logics figure prominently in this observation. Issues of computational complexity are explicitly mentioned as well (p. 74).}

The first TMS was devised by Doyle \cite{44}. For the expert reader it may be instructive to note, at this point, the differences between the present attempt at truth maintenance and Doyle’s original TMS. First, the present system is more liberal concerning the logical form of what Doyle calls justifications. Any instance of an axiom can be a justification. Second, justifications may well become retracted. Third, there is an epistemic ranking of justifications.

1.4 Modularity in Cognitive Science

Now that we have outlined the key results of our investigation, let us relate these, in somewhat greater detail, to the modularity hypothesis in cognitive science. To a great extent, this hypothesis originated from Fodor’s \textit{The Modularity of Mind} \cite{51}. There, Fodor develops a twofold thesis, which has been described as \textit{minimal} or \textit{peripheral modularity}: input and output systems of the mind, such as sensory and motor-control systems, work in a modular fashion. Central cognitive systems, by contrast, are non-modular. It is central cognitive systems that realise our capacities of explicit reasoning, belief formation and decision making. The notion of a module itself is characterised by the following properties and features in Fodor \cite{51}: (i) domain specificity, (ii) information encapsulation, (iii) mandatoriness, (iv) fast output, (v) shallow, i.e., non-conceptual output, (vi) neural localisation and (vii) innateness.

From the perspective of evolutionary psychology, more radically modular proposals have been made, also comprising central cognition. As indicated above, this work aims to show that the human mind works in a massively modular fashion (see in particular Cosmides and Tooby \cite{39}). According to this thesis, both peripheral and central cognition work with a modular architecture. Tractability is one of three arguments advanced in favour of the massive modularity hypothesis in Cosmides and Tooby \cite{39}. It is the only argument we are concerned with here.
1.4. MODULARITY IN COGNITIVE SCIENCE

Even though this is only a very rough sketch of the modularity map in cognitive science and evolutionary psychology, it is precise enough to locate our results on this map. In devising a proper logic of modular reasoning, we aim to contribute to an understanding of modularity at the level of central cognition. The modular units of reasoning characterised by frame logic share at least two important properties with the Fodorian notion of a module: they are domain-specific and work with information encapsulation. This is good news for proponents of the massive modularity hypothesis since encapsulation and domain specificity are considered most central to this hypothesis [92, p. 63].

Let us further compare our logical explanation of a module with Fodor’s notion. Elementary modules (which are not composed of other modules) have fast output insofar as they are associated with very simple patterns of inference. Furthermore, our specific notion of a module is perfectly consistent with having neural localisation. It is more than plausible to assume this property. As we are concerned with central cognition, our modules have non-shallow, conceptual output and input. Their working may or may not be mandatory. While Fodorian modules are not necessarily interactive, our logical modules are.

The present account of frames may well be viewed as a logical variant of the massive modularity hypothesis. This variant comes with a precise hypothesis about the working of information encapsulation:

(1) First-order and propositional reasoning within a module is encapsulated from any piece of extra-modular information.

(2) Information in the form of disjunctions and implications is encapsulated in the sense that it is located within a module and that it cannot directly be accessed by other modules.

(3) Only literals can be communicated between modular units of reasoning.

(4) The elementary modular units of reasoning are given by applications of generalisations.

Recall that a literal is an atomic sentence or a negation of such a sentence. By a generalisation we simply mean a universal proposition, which may or may not be strictly believed. Strict belief of a universal proposition means that we believe all instances to be true, whereas non-strict belief amounts to believing that a great deal of the instances are true. Non-strict generalisations are needed for default reasoning. As a single application of a generalisation (be it strict or non-strict) only concerns a small tuple of objects, our modules of
reasoning are highly domain-specific. Thus, we shall also be precise about
the way in which logical modules are domain-specific. Frame-logic itself is
of course domain-general.

Let us briefly motivate composition of modular units of reasoning. Recall
that any disjunction of literals can be translated into a logically equivalent
implication, and vice versa. (To give a simple example, \( p \rightarrow q \) is logi-
cally equivalent with \( \neg p \lor q \).) So, it suffices to analyse disjunctions when
analysing the complexity of reasoning with elementary modular units of rea-
soning. Now, we can distinguish between two ways of exploiting disjunctive
information. First, by direct inferences using premises in the form of literals.
A simple type of instance is given by the Disjunctive Syllogism:

\[
\begin{align*}
p \lor q, \quad \neg q & \quad \vdash p \\
\end{align*}
\]

Likewise, we can infer \( p \lor q \) from the set \( \{ p \lor q \lor r, \neg r \} \). These inferences can
easily be captured by intramodular reasoning using only elementary modules
(which are defined by a single disjunction or implication). Moreover, iterated
applications of such direct inferences from a disjunction and a literal are
captured by “communication” of literals between elementary modular units
of reasoning.

More complex, however, are inferences that involve reasoning by cases:

\[
\begin{align*}
p \lor q, \quad p \vdash r, \quad q \vdash r & \quad \vdash r \\
\end{align*}
\]

\( p \vdash r \) means that we can infer \( r \) from \( p \), possibly in a direct way using Dis-
junctive Syllogism. Reasoning by cases cannot be captured by elementary
modular units and their interaction. Hence, we need compositions of such
modular units. The composition does not consist in a simple and straight-
forward interaction of modular units, but is more demanding insofar as there
is no information encapsulation of the disjunctive information within a com-
position of elementary modules. Thus, full propositional reasoning can be
exercised within a composite module. Notably, we must limit compositions,
for otherwise intractability ensues.9

Such are the basic ideas of frame logic. The natural deduction system of
frame logic is more involved and contains frame-relative inference rules for
all logical symbols of first-order logic. The just explained distinction between
two types of exploiting disjunctive information is crucial to the present ap-
proach to modularity.

\[9\]What we describe as a direct exploitation of disjunctive formation using literals corre-
spends, of course, to an application of the inference rule Modus Ponens. It is inessential,
in the present context, in which way we present such inferences and whether we work with
disjunctions or implications.
1.4. MODULARITY IN COGNITIVE SCIENCE

Apparently, the patterns of information encapsulation and domain specificity as well as the interaction between modules can be described and summarised in a relatively simple way. Why is it then necessary to start with rather heavy tools, such as set-theoretic predicates of the Sneed formalism? With some qualifications, set-theoretic predicates and their axiomatic characterisation work as a ladder that can be thrown away once the inference rules of frame logic have been established. The natural deduction system of frame logic can be presented without set-theoretic predicates. Such predicates, however, are needed to define a proper model-theoretic semantics of frame logic. The soundness and completeness proofs for frame logic are based, therefore, on set-theoretic predicates and an axiomatic characterisation of their semantics. The set-theoretic tools, moreover, help us understand in which ways frame logic is inferentially weaker than first-order logic, while being inferentially strong enough to serve as a reasonable approximation of this logic.

Our logical variant of the massive modularity hypothesis bears some resemblance with ideas about informational modules in Samuels [91]. Informational modules are domain-specific bodies of knowledge, leading to what Samuels calls a library model of cognition. Carruthers [33, p. 34] argues that evolution would replace innate informational modules by innate computational ones. He goes on to envision that general reasoning capacities are actually realised ‘in cycles of operations of, and interactions between, existing modular systems’ (p. 35). Applied to (non-innate) informational modules, this description fits very well the working of frame logic and the TM belief revision scheme.

Carruthers [33] and Samuels [91] diverge as to what extent, if at all, the human mind has something like a central processing unit that carries out the drawing of inferences. For the present account of tractable belief revision to be implemented, some central processing may be needed to coordinate the interaction of the modular units of reasoning. But both the internal inference mechanism of an elementary modular unit of reasoning and the intermodular transmission of literals among such units is so simple and straightforward that these types of inference may well be drawn without central processing. That is, as soon as a piece of disjunctive information receives a new literal, the corresponding inference is drawn locally (where this piece of information is stored) and mandatorily in circumstances in which any inference is in fact drawn. Central processing, however, might come into play when we encounter inconsistencies and when we start composing modular units of reasoning.

Unlike the informational modules in Samuels [91], the informational modules of frame logic are not innate. Even so, the capacity to form such information...
tional modules may well be argued to be innate, in a manner comparable to how Chomsky [35] envisioned universal grammar to be innate. If one were to further argue that informational modules have their inference mechanisms hard-wired into the modules themselves, such an argument would lend support to the view that, at least, the minimal core of frame logic is innate. The minimal core is defined by frame logic without composition and confined to sentences that are literals or implications equivalent to disjunctions of literals.

Unlike Carruthers [33], we do not invoke ‘quick and dirty heuristics’ as a means to achieve tractability. To be more precise: if we understand the notion of a heuristics in opposition to a domain-general, possibly logic-oriented account of human cognition, then the present approach to tractability is not about heuristics at all. If we understand the notion of a heuristics in the weaker sense that using a heuristics means aiming at approximate solution of a given problem, then our approach certainly works with heuristics. Domain-specific heuristics may nonetheless form an important part of human cognition. But such heuristics are limited in explaining our capacity to reason about new domains of knowledge and enquiry.

Much more can be said about the present account of modularity in relation to research on modularity in cognitive science and evolutionary psychology. At this point, however, I would like to leave it to the reader to recognise and to pursue further connections. The present book is, of course, part of cognitive science insofar as tractability is an issue for philosophical logic only if we accept some form of CTM. One intent for writing the book has been to inspire further, more empirical research on modularity. For example, people working at the intersection of logic and psychology will find it easy to derive empirical predictions from the present account of modularity.

Let us briefly indicate these predictions. First, the present account of tractable belief revision predicts that human subjects have difficulties exploiting disjunctive information if this information cannot be exploited directly in the context of literals. For, composition of modular units of reasoning is computationally more demanding than iterated, straightforward inferences from disjunctions, even if the inferences to be compared have the same length. Second, the TM belief revision scheme makes distinct predictions as to how human subjects deviate from standard belief revision schemes in order to achieve tractability. Third, only very few pieces of everyday reasoning are such that reasoning by cases is unavoidable. For otherwise, this type of

\[\text{The notion of a heuristics is not clear-cut at all. See Neth and Gigerenzer [84] for a brief account and some historical references. There is some consensus that the notion of a heuristics is in opposition to a logic-oriented account of human cognition.}\]
1.5 Modularity in Knowledge Representation

Minsky’s seminal work on frames contributed, to a great extent, to the development of object-oriented programming. One must therefore wonder if modular design could be used to advance logic-oriented knowledge representation in a manner similar to the advancement of traditional, imperative programming by frames (cf. [13, p. 275]). The logic we will devise is an attempt at a logic of modular reasoning, and thus very much in line with this direction of research.

Of course, much work has been done on bringing modular design to bear on logic-oriented knowledge representation in past decades. To the best of our knowledge, however, there is no proper logic of modular reasoning that goes as far as the present proposal in exploiting ideas about frames and object-oriented design to devise a logical system of reasoning. Work on partition-based logical reasoning by Amir and McIlraith [3] is driven by similar ideas, but not taken to the development of a logic of modular reasoning with features like encapsulation and composition. Nor is this work aimed at approximating first-order reasoning by partitioning a knowledge base.

The above explained distinction between two types of exploiting disjunctive information – straightforward inferences and reasoning by cases – is likewise crucial in the logic of limited belief for reasoning with disjunctive information in Liu et al. [75] and subsequent work. Tractability is achieved, in this logic, by delimiting the level of iteration of reasoning by cases. In frame logic, by contrast, reasoning by cases can only be carried out by means of composing disjunctive information. Frame logic, therefore, differs from the logic in question as regards to the means through which reasoning by cases is constrained. Arguably, composition corresponds to a cognitive operation of putting certain pieces of information together. Thereby, we get a logical grip on the notion of a cognitive module of reasoning, which could also be important for research in knowledge representation. No doubt, it is desirable to say more about the heuristics of compositions in future work. Frame logic is not equivalent with the logic of limited belief by Liu et al. [75].

In the absence of compositions, frame-logical reasoning with sets of disjunctions of literals is equivalent to drawing inferences using just unit propagation, i.e., propagation of determinate literals. Frame-logic may therefore enable us to recognise unit propagation as a proper logic that comes with a well-defined semantics and proof theory. Such a logic is tractable and, yet, more
expressive than Horn logic insofar as we can reason explicitly with negated atoms.

The present book is written from the perspective of philosophical logic and a liberal understanding of the computational theory of mind. Yet, the truth maintenance system is expounded in such a manner that an implementation should not be exceedingly difficult for people working in the more applied areas of knowledge representation. Future research must show to what extent, if at all, the present results may be used to advance modular design in knowledge representation.11

1.6 Overview

Part I begins with an introduction to frame concepts in Chapter 2. We shall also explain, in this chapter, the notion of a set-theoretic concept. Some references to object-oriented programming may deepen our understanding of frames and will prove useful when we come to deal with the challenges of truth maintenance at a later stage of the investigation. Chapter 3 introduces the fundamental concepts and principal approaches in belief revision theory. The overall rationale of defeasible reasoning is explained in Chapter 4. There, we shall introduce an inference system of defeasible reasoning in greater detail. Also, a brief account of the truth maintenance system by Doyle [44] is given, with an explanation why this system has very interesting properties from a cognitive point of view.

Part II starts with a novel and axiomatic account of structuralist theory representation in terms of frame concepts. Chapter 6, then, introduces defeasible forms of reasoning into the structuralist approach, which paves the way for defining belief revisions in this approach. Such are the logical and set-theoretic foundations upon which the truth maintenance system will be devised in Chapter 7. Frame logic, finally, emerges from the TMS and the semantics of structuralist belief changes in Chapter 8.

11Admittedly, the label frame logic has already been used for another logic, which however is designed to reason about objects in an object-oriented language (see Kifer et al. [71]). This logic is not aimed at bringing features of object-oriented design to bear on a logical account of quotidian and scientific reasoning. Hayes [66] attempts to give an account of the logic of frames, but the result is quite different from the logic of our investigation. Apologies about any confusion arising from these ambiguities. I could not think of a better label for the logic of modular reasoning that results from the present investigation.