2 The Anatomy of Logic

We have seen that logic was once thought relevant to the study of cognition both as a representational format and as an inference mechanism, and that developments in psychology (Wason, increasing prominence of neural network modeling, decision theory) and in philosophy (concerns with normativity, anti-psychologism) have led to the widely shared view that logic is irrelevant to cognition. We have sketched a new view of logical reasoning, following Husserl, in which reasoning is simultaneously formal and relative to a domain. On this view, cognitive science needs to be much more attentive to *semantics* – because meaning is often not given but constructed. Indeed, we will see that subjects' behavior in reasoning tasks (e.g., Wason's selection task) is much less irrational than is commonly thought, once one takes into account that these subjects are struggling to impose a meaning on the task. It is by no means obvious to the subject that her reasoning must be based on the classical interpretation of the conditional as material implication. In fact, the interest of the standard reasoning tasks lies precisely in the window it offers on subjects' efforts to impose meaning. As a first step toward weaning the reader away from the idea that the semantics of logical expressions is given by classical logic, this chapter presents the reader with an overview of the semantic possibilities.

This chapter is organized as follows. We start from a popular conception of logical reasoning according to which, to see whether an argument is valid, one translates it into the formal language of classical logic and checks the resulting pattern for classical validity. We argue that this conception is inadequate, and oppose it to a formal version of Husserl's view, in which one distinguishes reasoning *from* an interpretation and reasoning *to* an interpretation.[1] We conceive of the latter as a form of parameter setting. To illustrate the idea, we start from the four parameter choices defining classical logic, which is appropriate for the

1. We will use the verb "to interpret" in this book in its *Oxford English Dictionary* sense of "to make out the meaning of," and the noun "interpretation" as the result of that activity, or occasionally as the process itself. This is not saying much if we do not explain what "to make out" and "meaning" mean. In fact, most of this book is concerned with explaining what is involved in "making out the meaning of," and no simple explanation can be given at this stage.

domain of classical mathematics, and by systematic variation of the parameters obtain logics which are appropriate for other domains.

2.1 How Not to Think about Logical Reasoning

In the psychology of reasoning literature one commonly finds a picture of reasoning as proceeding according to preestablished logical laws, which can be applied by anybody in any circumstances whatsoever.

It would not do to blame the psychologists for this, because it is a picture frequently promulgated in the philosophical literature. To take just one example, we see Ryle [239] characterizing logical constants (for example, *all, some, not, and, or, if*) as being indifferent to subjectmatter, or as it is sometimes callled, *topic neutral*. Characterizations such as this are related to a superficial reading of the classical definition of validity, say for a syllogism such as

All A are B.
All B are C.
Therefore, all A are C.

The validity of this schema is taken to mean something like "whatever you substitute for A, B and C, if the premises are true for the substitution, then so is the conclusion." Analyzing an argument thus consists of finding the topic-neutral expressions (the logical constants), replacing the topic-dependent expressions by variables, and checking whether a substitution that verifies the premises also verifies the conclusion. If so, one knows that the argument is correct for the particular substitution one is interested in.

This *schematic* character of inference patterns is identified with the "domain-independence" or "topic neutrality" of logic generally, and many take it to be the principal interest of logic that its laws seem independent of subject matter. In fact, however, logic is very much domaindependent in the sense that the valid schemata depend on the domain in which one reasons, *with what purpose*. We therefore view reasoning as consisting of two stages: first one has to establish the domain about which one reasons and its formal properties (what we will call "reasoning *to* an interpretation") and only after this initial step has been taken can one's reasoning be guided by formal laws (what we will call "reasoning *from* an interpretation").

2.2 Reasoning to an Interpretation as Parameter Setting

We should start with an informal illustration of what the process of interpretation involves, which falls into at least two questions – what things are actually in the domain? and also: what kinds of reasoning will be done about them? We start with the former question, which has been extensively studied in the formal semantics of natural languages. We illustrate the general distinction between the two questions with some homely examples of discourse understanding, which will then introduce a particular distinction that will figure centrally in the rest of the book.

> Once upon a time there was a butcher, a baker, and a candlestick maker. One fine morning, a body was discovered on the village green, a dagger protruding from its chest. The murderer's footprints were clearly registered in the mud. ...

Well, what follows from this discourse? For example, how many people are there? If we take the most likely approach to interpreting this discourse outside of logic class, we will assume that there are at least three people – a butcher, a baker, and a candlestick maker. There are, of course also the corpse and the murderer, but it is an active question whether these are identical with any of the former three, and who else may there be in this dire place? These questions are questions about what things, or people, or other entities are in the domain of interpretation. Mundane as these questions are, they are absolutely central to how natural language functions in progressively constructing interpretations as discourse proceeds.

It should be made clear from the outset that discourse interpretation is not at all exhausted by composing the meanings of the lexical items (as given by the dictionary) in the way dictated by the syntax of the sentences. Contextual information plays a crucial role. For instance, the question, what are the characters in this discourse? is a question about what is in the current domain of interpretation, and the answer to this question may well depend on discourse context, as we shall see. Clearly our knowledge of the dictionary plays a role in our answer to this question, but does not by itself provide the answer. Domains of natural language interpretation are often very local, as they are here. They often change sentence by sentence as the discourse proceeds. It is this sense of interpretation, rather than the dictionary-level sense, which generally occupies us here.

Suppose now we have instead a discourse that runs as follows:

> Some woman is a baker. Some woman is a butcher. Some woman is a candlestickmaker. Some person is a murderer. Some person is a corpse. All women are men.[2, 3, 4]

Now we are much more likely to entertain considerably more possibilities about how many people there are, cued perhaps by the "logical puzzle" style of the discourse. Now it becomes entirely possible that the butcher may turn out to be the baker, or one person might pursue all three professions, even before we

2. NB. The *Oxford English Dictionary* defines, under its first sense for *man*, "a human being irrespective of gender."

3. "Oh *man*!, these guys' language is archaic!" addressed to a female human is an example of the *Oxford English Dictionary*'s archaic usage hidden in modern oral vernacular English.

4. The previous two footnotes are irrelevant if this discourse is processed from a skeptical stance.

start on the problem about who is dead and who has been nasty, and just who else is in this village, if there is one.

The first discourse is likely to be understood with what we will call a *credulous* stance. As we interpret the discourse, we take our task to be to construct a model of the story which is the same as the speaker's "intended model," and we assume that we are to use whatever general and specific knowledge we have, including the assumption that the speaker is being cooperative in constructing her discourse, to help us guess which model this is. The second discourse is likely to be understood with what we will call a *skeptical* stance in which we do not use any information save the explicitly stated premises, and we are to entertain all possible arrangements of the entities that make these statements true. This stance explains already why the footnotes are completely irrelevant to this interpretation and merely designed to lead us astray.

First note that these different stances lead to quite different numbers of people in the domains of interpretations of the two texts. In the first discourse we know[5] that there is no policeman although we also know, from general knowledge, that this is likely to change rather soon. In the second we do not know whether there is a policeman, but unless we are explicitly told that there isn't (or told something which explicitly rules it out) then we still do not know, even though no policeman is ever mentioned. These "number-of-things" questions are only the most tangible tip of the iceberg of differences between the domains we get when we process with these two different stances, but they suffice for our present illustrative purposes.

Credulous and skeptical stances are good examples of the second kind of issue about interpretations – what kinds of reasoning will we do about the things in the domain? Credulous reasoning is aimed at finding ideally a single interpretation which makes the speaker's utterances true, generally at the expense of importing all sorts of stuff from our assumed mutual general knowledge. Skeptical reasoning is aimed at finding only conclusions which are true in *all* interpretations of the explicit premises. These are very different goals and require very different logics, with, for example, different syntactic structures and different concepts of validity. The differences in goals are socially important differences. In credulous understanding we accept (at least for the purposes of the discourse) the authority of the speaker for the truth of what is said. We are only at an impasse when there appears to be a contradiction which leaves us with no model of the discourse, and when this happens we try to repair our interpretation in order to restore a model. In skeptical understanding, we consider ourselves as on allfours with the speaker as regards authority for inferences, and we may well challenge what is said on the basis that a conclusion doesn't follow because we can find a single interpretation of the premises in which that conclusion is false.

5. By what is known as "closed-world reasoning," for which see section 2.3.

A good illustration of the distinction between credulous and skeptical reasoning is furnished by legal reasoning in the courtroom, of which the following is a concrete example (simplified from a case which recently gained notoriety in the Netherlands). A nurse is indicted for murdering several terminally ill patients, who all died during her shifts. No forensic evidence of foul play is found, but the public prosecutor argues that the nurse must have caused the deaths, because she was the only one present at the time of death. This is an example of "plausible" or "credulous" reasoning: an inference is drawn on the basis of data gathered and plausible causal relationships.[6]

The defense countered the prosecutor's argument with an instance of 'skeptical' reasoning, by arguing that the cause of death might as well have been malfunctioning of the morphine pumps, and contacted the manufacturer to see whether morphine pumps had had to be recalled because of malfunctioning – which indeed turned out be the case (although in the end it did not help the defendant). The move of the defence can be viewed as enlarging the class of models considered, thus getting closer to the standard notion of logical consequence where one considers all models of the premises instead of a restricted class. Here is Ryle [239,p.116] again, this time with a very pertinent remark:

> There arises, I suppose, a special pressure upon language to provide idioms of the [logical] kind, when a society reaches the stage where many matters of interest and importance to everyone have to be settled or decided by special kinds of talk. I mean, for example, when offenders have to be tried and convicted or acquitted; when treaties and contracts have to be entered into and observed or enforced; when witnesses have to be cross-examined; when legislators have to draft practicable measures and defend them against critics; when private rights and public duties have to be precisely fixed; when complicated commercial arrangements have to be made; when teachers have to set tests to their pupils; and ... when theorists have to consider in detail the strengths and weaknesses of their own and one another's theories.

We have chosen to illustrate the kinds of issues that go into deciding what domain is adopted in an interpretation with this particular distinction because it is the one that is at the center of many of the misunderstandings between experimenter and subject in the psychology of reasoning experiments. The what things are in the domain? question is always present in any process of interpretation. The what kind of reasoning are we doing? question is rather different for different distinctions.

So far we have been talking about domain in a rather loose manner, as roughly synonymous with universe of discourse. For logical purposes it is important to make a type-token distinction here. The domain mentally constructed while interpreting a discourse is a concrete instance – a token – of a general kind –

6. A note for the logically minded reader: this can be viewed as an inference where the premises are interpreted on a very restricted class of models, namely models in which no "mysterious" events happen, neither divine intervention nor unknown intruders.

the type – which determines the logical properties of the token. It is very hard to completely pin down this general notion of a type of domain itself; we will try to do so in a later chapter on evolutionary approaches to reasoning, where the notion of "domain specificity" of reasoning plays an important role. But we can at least list some examples that will be treated in this book: actions, plans and causality; contracts; norms; other people's beliefs; mathematical objects; natural laws. Slightly more formally, a domain is characterized by a set of mathematical representations, called *structures*, of the main ingredients of the domain (e.g., the objects in the domain, their relations, the events in which they participate), together with a formal language to talk and reason about these structures. The connection between structures and formal language is given by what is technically known as a definition of satisfaction: a characterization of how the formal language is interpreted on the relevant set of structures. This notion of domain is extremely general, and instead of being more precise at this point, we refer the reader to the different examples that will be given below.

The reader may wonder why language and logic should be relative to a domain: isn't there a single language – one's mother tongue – which we use to talk and reason about everything? Much of the progress in mathematical logic in the last century shows, however, that it is not useful to have a single language (with a single semantics) for talking about everything. For instance, the vague predicates that we will meet when discussing diagnostic reasoning in medicine can perhaps be represented by fuzzy logic with its continuum of truth–values (see section 2.2.3), but it would make no sense to use this semantics for classical mathematics. For another example, consider two radically different ways of doing mathematics: classical and constructive mathematics. Very roughly speaking, the difference is that the former tradition, unlike the latter, accepts the principle of bivalence: a sentence is either true or false (see section 2.2.3 for an explanation of why this principle is sometimes unwarranted). Constructive mathematics is often useful in computer science, because the results it yields have algorithmic significance, while this is not guaranteed of results in classical mathematics. It occasionally happens that the same mathematician may apply both methods, depending on the domain she is working in. So does she believe in bivalence, or doesn't she? The answer is that sometimes she does, and sometimes she doesn't, whatever is most appropriate to the domain of interest. In this sense logics are local. One might want to argue that in such cases one should adopt the weakest logic (in this case the one without bivalence) as one's generally valid logic; after all, how can a principle such as bivalence be called logical at all if it is considered to be false in some domains? One quick answer to this argument is that this "weakest logic" soon trivializes when including more domains, for example when considering also uncertain information instead of just truth and proof. One may conclude from this that logic as a system of generally valid inference principles has no role to play in actual reasoning. Another option, and the one advocated here, is to give up the idea that logic must be such a system. Clearly, however, if logic is not given, the question becomes how one comes to reason in a particular logic. The answer argued for in the book, and made explicit in chapter 11, is that mastering a particular domain essentially involves mastering its logical laws. These brief indications must suffice at this stage, and we will return to the wider issues in the concluding chapter.

We are now ready to delve into the technicalities. The approach to logic which we would like to advocate views logics from the point of view of possible syntactic and semantic choices, or what we will call parameter settings. This metaphor should not be taken too literally: we do not claim that a logic can be seen as a point in a well-behaved many-dimensional space. The use of the term parameter here is analogous to that in generative linguistics, where universal grammar is thought to give rise to concrete grammars by fixing parameters such as word order.[7] The set of parameters characterizing a logic can be divided in three subsets

1. Choice of a formal language
2. Choice of a semantics for the formal language
3. Choice of a definition of valid arguments in the language

As we shall see, different choices for the parameters may be appropriate in different domains – each domain gives rise to a notion of structure, and in principle each domain comes with its own language.[8]

To familiarize the reader with this idea, we first present classical propositional logic as resulting from four contingent assumptions, which are sometimes appropriate, sometimes not. We will then vary these assumptions to obtain a host of different logics, all useful in some context.

2.2.1 Classical Propositional Logic

The purpose of this section is to show that classical logic is inevitable once one adopts a number of parameter settings concerning syntax, meaning and truth, and logical consequence; and furthermore that these settings are open to debate. The relevant parameter settings are:

1. [*syntax*] fully recursive language: if φ, ψ are formulas, then so are $\neg\varphi$, $\varphi \rightarrow \psi$, $\varphi \vee \psi$, $\varphi \wedge \psi, \ldots$[9];

7. This is just an analogy; we are not committed to anything like UG.
8. This approach to logic was pursued in the 1980s under the heading of "model theoretic logics"; see [14].
9. This definition generates formulas like $(\varphi \rightarrow \theta) \rightarrow \psi$. The iteration of a conditional inside the antecedent of another conditional illustrated by this last formula will turn out to be a distinctive property of this language, which sets it off from the language we use to model credulous interpretation.

2. [*semantics*] truth–functionality: the truth–value of a sentence is a function of the truth–values of its components only;

(2′. as a consequence: evaluation of the truth–value can be determined in the given model (the semantics is *extensional*));

3. [*semantics*] bivalence: sentences are either true or false, with nothing in between;

4. [*consequence*] the Bolzano–Tarski notion of logical consequence[10]:

$$\alpha_1 \ldots \alpha_n / \beta \text{ is valid iff } \beta \text{ is true on all models of } \alpha_1 \ldots \alpha_n.$$

These assumptions force a unique formalization of the logical connectives *not, and, or, if ... then*, as given by the familiar truth–tables in figure 2.1.

p	$\neg p$
1	0
0	1

p	q	$p \wedge q$
1	1	1
0	0	0
1	0	0
0	1	0

p	q	$p \vee q$
1	1	1
0	0	0
1	0	1
0	1	1

p	q	$p \rightarrow q$
1	1	1
0	0	1
1	0	0
0	1	1

Figure 2.1 Truth–tables for classical logic.

It is instructive to see how our four assumptions (in conjunction with intuitive judgments about meaning) lead to the formalization of the conditional "if ... then" as the "material implication" $\rightarrow$ defined by the above truth–table.

Truth–functionality requires that if the truth–values of p, q are given, so is that of $p \rightarrow q$, and bivalence forces these to be either 0 or 1. We can see from this that if p is true and q is false, then $p \rightarrow q$ must be false; for if it were true then modus ponens ($p, p \rightarrow q \models q$) would fail. Furthermore an application of the definition of validity shows that the following argument patterns are valid : $p, q \models q$ and $p, \neg p \models q$. From this it follows from the intuitive meaning of the conditional that $q \models p \rightarrow q$ and $\neg p \models p \rightarrow q$. Indeed, one may argue for an implication $p \rightarrow q$ by assuming p and inferring from this (and given premises) that q. But this reduces $q \models p \rightarrow q$ to $p, q \models q$. The validity of $\neg p \models p \rightarrow q$ is established similarly.

The classical definition of validity is *monotonic*, that is, if $\alpha_1 \ldots \alpha_n \models \beta$, then also $\delta, \alpha_1 \ldots \alpha_n \models \beta$. It follows that the valid argument $q \models p \rightarrow q$ forces $p \rightarrow q$ to be true if p, q are true and if $\neg p, q$ are true; in addition, $\neg p \models p \rightarrow q$ forces $p \rightarrow q$ to be true if $\neg p, \neg q$ are true. We have now justified all the lines of the truth–table.

10. Whether this historical attribution is correct is debatable; see [71], also for elaborate discussion of the flaws of this particular definition of validity.

Domains to which classical logic is applicable must satisfy the four assumptions. Classical mathematics is a case in point. Here it is assumed that all statements are true or false – together with truth–functionality this gives the celebrated principle of excluded middle $p \vee \neg p$, which we will see in action later. The definition of logical consequence is a very important feature of modern mathematics: it implies that a counterexample to a theorem makes it false. Trivial as this may seem nowadays, this has not always been the case; in the eighteenth and nineteenth centuries it was not uncommon to conclude that a purported counterexample did not belong to the "domain" of the theorem, thus effectively restricting the class of models. One may consult the work of Lakatos, in particular [167], for instructive examples.[11]

Are these four assumptions in general always fulfilled? The next sections provide example domains in which various combinations of the assumptions obviously fail, and we will indicate what logics are appropriate to these domains instead.

2.2.2 Truth–Functionality without Bivalence

Why would every statement be either true or false? This depends of course very much on what you want to mean by "true" and "false." One could stipulate that "not true" is the same as false, but such a stipulation is definitely inappropriate if we consider "true" to mean "known to be true." One example of where this occurs in practice is a computerized primality test which checks whether the input $2^{1257787} - 1$ is a prime number. One could say that, while the program is running, the statement "$2^{1257787} - 1$ is a prime number" is undecided; but a decision may follow in the end, if the program halts.[12]

One possibility to formalise this idea, originated by [160] is to add a third truth–value u for "undecided" or "not known to be true and not known to be false"; u can (but need not) "evolve" toward "known to be true" or "known to be false" when more information comes in. This uniquely determines the truth–tables as given in figure 2.2.

The other three assumptions characterizing classical logic are still in force here. The resulting logic is appropriate to the domain of computable functions, and also to paradoxical sentences such as "I am false," and more generally to languages which contain their own truth predicate (such as natural language).

11. A note for the logically minded reader. In principle the language of classical mathematics is fully recursive. In practice, restrictions apply, so that particular structures, for example the reals, are described in a restricted language. One of the triumphs of mathematical logic is the use of these restrictions in language to prove positive results about the structures that the language describes.

12. It does, and the number is prime.

p	$\neg p$
1	0
0	1
u	u

p	q	$p \wedge q$
1	1	1
0	0	0
u	u	u
1	0	0
1	u	u
0	1	0
0	u	0
u	1	u
u	0	0

p	q	$p \vee q$
1	1	1
0	0	0
u	u	u
1	0	1
1	u	1
0	1	1
0	u	u
u	1	1
u	0	u

p	q	$p \rightarrow q$
1	1	1
0	0	1
u	u	u
1	0	0
1	u	u
0	1	1
0	u	1
u	1	1
u	0	u

Figure 2.2 Truth–tables for Kleene three-valued logic.

2.2.3 A Domain in which Bivalence is Truly Ridiculous

Here is an excerpt from a textbook on cancer, in a section on differential diagnosis. The reader should realise that this is the kind of text that guides a physician in her decision making. We have distinguished typographically two classes of expressions: in **boldface** vague expressions like "small," "painful," "entire," "changes," "diffuse without sharp demarcation," "feels like a tumour," . . . ; and in *italic* qualitative-probabilistic adverbs like "usually," "often," "approximately 15% of the cases," "if A maybe B," "infrequently – more often."

> **Chronic** cystic disease is *often* confused with carcinoma of the breast. It *usually* occurs in parous women with **small** breasts. It is present *most commonly* in the **upper outer quadrant** but *may* occur in other parts and eventually involve the **entire** breast. It is *often* **painful**, particularly in the pre-menstrual period, and accompanying **menstrual disturbances** are *common*. Nipple discharge, *usually* **servous**, occurs in *approximately 15% of the cases*, but there are no **changes** in the nipple itself. The lesion is **diffuse without sharp demarcation** and without fixation to the overlying skin. Multiple cysts are **firm, round and fluctuant** and *may* transilluminate *if* they contain a **clear** fluid. A **large** cyst in an area of chronic cystic disease **feels like a tumour**, but is *usually* **smoother and well-delimited**. The axillary lymph nodes are *usually* not **enlarged**. **Chronic** cystic disease *infrequently* shows **large bluish** cysts. *More often*, the cysts are multiple and **small**. (J.A. del Regato. Diagnosis, treatment and prognosis. Pages 860–861. In L.V. Ackerman (editor) *Cancer* 1970.

To find logical regularities in this domain is challenging, to put it mildly. Vague predicates have sometimes been formalized using many-valued logics, and there have been attempts to model frequency adverbs using probability theory. The reader is urged to compare the preceding piece of text with the formal systems that follow, to see whether they add to her understanding.

It is also important to be aware that in real life vagueness may be treated by being avoided. Consider the locus classicus for the rejection of logic in cognitive science: Rosch and Mervis's arguments for its inapplicability to human classificatory behavior in [233]. Classical logic represents the extension of a predicate by a set, to which things either belong or they don't. No half measures. But people classify things by shades. They represent typical members of extensions. Red is typified by the color of blood and the color of red hair is a peripheral red. There is cognitive structure here which there is not in a set. And so, argue Rosch and Mervis, logic is inapplicable.

This is a good example of a levels confusion. Rosch and Mervis are concerned with the dictionary meanings of vague words such as *red*. Logic is concerned with meaning as it occurs at the discourse level and has very little to say about the dictionary level; but the point is that it need not. Suppose we start a conversation which includes the word *red*. It is unlikely that the vagueness of this term will become critical to our mutual interpretation – we may be happy that we know how to classify all the relevant objects (perhaps three traffic lights) with regard to this term perfectly crisply. If it does become a problem then we may resort to increased precision – "by red I mean crimson lake as manufactured by Pigment Corp." – which may or may not replace the word red entirely. Practically all natural language words are vague, and they would be useless if they weren't, but if we design our discourse well, their vagueness will be well tailored to the local communicative situation.

Another way to make the same point is with reference to Marr's methodology as outlined in [183,p. 357ff], in particular chapter 7, where he conducts a dialogue with himself and asks

> *What do you feel are the most promising approaches to semantics?*

The answer is

> Probably what I call the problem of multiple descriptions of objects and the resolution of the problems of reference that multiple descriptions introduce. ... I expect that at the heart of our understanding of intelligence will lie at least one and probably several important principles about organizing and representing knowledge that in some sense capture what is important about our intellectual capabilities, [namely:]
>
> 1. The perception of an event or object must include the simultaneous computation of several different descriptions of it, that capture diverse aspects of the use, purpose, or circumstances of the event or object.
> 2. That the various descriptions referred to in 1. include coarse versions as well as fine ones. These coarse descriptions are a vital link in choosing the appropriate overall scenarios ... and in establishing correctly the roles played by the objects and actions that caused those scenarios to be chosen.

A coarse description of a vague predicate, using classical logic, may well be able to model the avoidance of the vagueness which is endemic in discourse.

Alternatively, we may move to a finer description, meet vagueness head-on, and change our logic. We give two examples.

Dealing with Vagueness: Łukasiewicz Logic

This logic also differs from classical logic in that it has a third truth–value ($\frac{1}{2}$), but this value now means "intermediate between true and false," and not "undecided, but possibly decided at some later time." The reader may verify that the truth–tables in figure 2.3 have been calculated according to the following formulas: $\neg p$ corresponds to $1-p$, $p \wedge q$ to $\min(p,q)$, $p \vee q$ to $\max(p,q)$, and $p \rightarrow q$ to $\min(1, 1+q-p)$. Once one has seen that the tables are calculated

p	$\neg p$
1	0
0	1
$\frac{1}{2}$	$\frac{1}{2}$

p	q	$p \wedge q$
1	1	1
0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
1	0	0
1	$\frac{1}{2}$	$\frac{1}{2}$
0	1	0
0	$\frac{1}{2}$	0
$\frac{1}{2}$	1	$\frac{1}{2}$
$\frac{1}{2}$	0	0

p	q	$p \vee q$
1	1	1
0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
1	0	1
1	$\frac{1}{2}$	1
0	1	1
0	$\frac{1}{2}$	$\frac{1}{2}$
$\frac{1}{2}$	1	1
$\frac{1}{2}$	0	$\frac{1}{2}$

p	q	$p \rightarrow q$
1	1	1
0	0	1
$\frac{1}{2}$	$\frac{1}{2}$	1
1	0	0
1	$\frac{1}{2}$	$\frac{1}{2}$
0	1	1
0	$\frac{1}{2}$	1
$\frac{1}{2}$	1	1
$\frac{1}{2}$	0	$\frac{1}{2}$

Figure 2.3 Łukasciewicz logic

using the above formulas, there is no reason to stop at three truth–values; one might as well take a continuum of truth–values in $[0,1]$. This system is called *fuzzy logic*. The important point to remember is that fuzzy logic is still truth–functional; in this respect it differs from our next example, probability theory, which is not.

Probability: a Many-Valued, Non-truth–Functional Semantics

One could try to represent frequency adverbs like "usually," "often" by means of probabilities. For instance, if p is the proposition that "The axillary lymph nodes are not enlarged," then 'usually(p)' could mean "the probability of p is greater than 60%," where probability is here taken in the sense of relative frequency. This idea leads to the following definition.[13]

A *probability* on a propositional language $\mathcal{L}$ is a function $P : \mathcal{L} \longrightarrow [0,1]$ satisfying

1. $P(\varphi) = 0$ if φ is a contradiction;
2. if φ and ψ are logically equivalent, $P(\varphi) = P(\psi)$;
3. if φ logically implies $\neg\psi$, then $P(\varphi \vee \psi) = P(\varphi) + P(\psi)$.

13. Note that probability is used as a semantics only. One could also try to develop "probability logics" where "the probability of p is c" is a statement of the object language.[113]

The implicit assumption underlying this definition is that the formulas in $\mathcal{L}$ satisfy the classical logical laws, so that "equivalence," "contradiction" etc. are uniquely defined. It is in fact not so easy to define probability on nonclassical logics. This will be one of criticisms when discussing recent attempts to explain logical reasoning by assuming underlying probabilistic reasoning processes: probabilistic reasoning is too much tied to classical logic to be able to encompass the wide variety of reasoning that actually occurs.

The reader may wish to show that this semantics is *not* truth–functional: the only restriction on the values of $P(\varphi \wedge \psi)$, $P(\varphi)$ and $P(\psi)$ is the a priori restriction $P(\varphi \wedge \psi) \leq P(\varphi), P(\psi)$.

Non-Truth–Functional Semantics: Intuitionistic Logic

In classical mathematics one often finds proofs which appeal to the principle of excluded middle, the syntactic analogue of bivalence. Mathematicians in the constructivist or intuitionistic tradition have pointed out that the use of this principle leads to proofs which are completely uninformative. Here is a toy example of this phenomenon.

Definition 1 *A* rational *number is one which can be written as* $\frac{p}{q}$ *for natural numbers* p, q*; an* irrational *number is one which cannot be so written.*

Suppose you want to prove:

Theorem 1 *There are irrational numbers* a, b *such that* a^b *is rational.*

PROOF. It is known that $\sqrt{2}$ is irrational. Consider $\sqrt{2}^{\sqrt{2}}$.

- If $\sqrt{2}^{\sqrt{2}}$ is rational, put $a = b = \sqrt{2}$ and we are done.
- If $\sqrt{2}^{\sqrt{2}}$ is irrational, put $a = \sqrt{2}^{\sqrt{2}}$, $b = \sqrt{2}$, then $a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^2 = 2$.

Either way you have the requisite a and b. $\square$

But what have you proved? Do you now *know* how to construct irrational a, b are such that a^b is rational? Such uninformative proofs, which do not yield concrete constructions, are typical of the use of the principle of excluded middle, which some therefore reject. If one characterises *intuitionistic logic* syntactically as classical logic *minus* the schema $\varphi \vee \neg\varphi$ (for all φ), then one can indeed show that proofs of existential statements in intuitionistic logic invariably yield concrete witnesses.

The semantics of intuitionistic logic is very different from what we have seen so far, where propositions took numbers as truth–values. The failure of the principle of excluded middle does not mean truth–functionally adding a third truth–value, one reason being that $\varphi \wedge \neg\varphi$ is still a contradiction, and inspection of the truth–tables for Kleene logic or Łukasciewicz logic shows that this is not the case there.[14] Instead we have the non-truth–functional "provability interpretation of truth," of which the following are examples

1. e.g., $\varphi \wedge \psi$ means: "I have a proof of both φ and ψ."
2. $\varphi \vee \psi$ means: "I have a proof of φ or a proof of ψ."
3. $\varphi \rightarrow \psi$ means: "I have a construction which transforms any given proof of φ into a proof of ψ."
4. special case of the previous: $\neg\varphi$ means "any attempted proof of φ leads to a contradiction."

Clearly, $\varphi \vee \neg\varphi$ is not valid on this interpretation, because it would require us to come up with a proof that φ or a proof that φ leads to a contradiction; and often one has neither.

An Intensional Logic: Deontic Logic

Deontic logic is concerned with reasoning about *norms*, i.e., what one *ought* to do; e.g., "if a person is innocent, he ought not to be convicted." We shall see in our discussion of the Wason selection task, however, that its scope extends much wider. At the syntactic level, (propositional) deontic logic consists of classical propositional logic plus the operator O, governed by the clause that if φ is a formula, so is $O\varphi$. The intuitive meaning of $O\varphi$ is "it ought to be the case that φ." Note that O cannot be a truth function, such as $\neg$; the truth of Op depends on the meaning, not the truth–value, of p. That is, both the propositions "my bike is grey" and "I don't steal" are true, but the latter ought to be the case, unlike the first.

The *intensional* semantics for deontic logic computes 'compliance'–values of formulas in a given model by referring to other models. One assumes that every model w has a "normatively perfect counterpart v"; this is formally represented by a relation: $R(w, v)$). In such a "normatively perfect" world v only what is permissible is the case; e.g., in w an innocent person may be convicted, but in v with $R(w, v)$ this same innocent person will not be convicted. We may now put $w \models Op$ if and only if for all v satisfying $R(w, v)$: $v \models p$.

To see the difference with classical logic, compare the conditionals $p \rightarrow q$ and $p \rightarrow Oq$. If in a given world, p is true and q is false, then $p \rightarrow q$ is simply false, whereas $p \rightarrow Oq$ can be true, in which case the given world is not "normatively perfect."

14. Another reason is that one would also like to get rid of the principle of double negation elimination $\neg\neg\varphi \rightarrow \varphi$.

Our last example, closed–world reasoning, is one in which the consequence relation is the main focus. Since closed–world reasoning will occupy us much throughout, it warrants a new section.

2.3 The Many Faces of Closed–World Reasoning

As we have seen above, the classical definition of validity considers *all* models of the premises. This type of validity is useful in mathematics, where the discovery of a single counterexample to a theorem is taken to imply that its derivation is flawed. But there are many examples of reasoning in daily life where one considers only a subset of the set of all models of the premises. In this section we review some examples of closed–world reasoning and their formalization.

One example is furnished by train schedules. In principle a schedule lists only positive information, and the world would still be a model of the schedule if there were more trains running than listed on the schedule. But the proper interpretation of a schedule is as a closed world – trains not listed are inferred not to exist. This is like our example of the butcher, baker, and candlestickmaker on page 21. Note that there is a difference here with the superficially similar case of a telephone directory. If a telephone number is not listed, we do not therefore conclude that the person does not have a telephone – she might after all have an unlisted number. In fact a moment's reflection suggests that such examples can be found within the "train schedule" domain. From the point of view of a prospective passenger, the inference that there is no train between two adjacently listed trains may be valid, but for a train spotter interested in trains passing through on the track, trains "not in service" may well occur between listed services. Thus, world knowledge is necessary to decide which logic is applicable.

2.3.1 Closed–World Reasoning, More Formally

Consider the Dutch database for public transportation www.9292ov.nl, which you consult for planning a trip from Amsterdam to Muiden. The database contains *facts* about trains and buses leaving at specific times, and also *rules* of the form

1. if bus 136 leaves Naarden-Bussum at 10:06, it will arrive in Muiden at 10:30
2. if train from Amsterdam CS in direction Naarden-Bussum leaves at 9:39, it will reach Naarden-Bussum at 10:00.

Backward chaining of the rules then generates a *plan* for getting from Amsterdam to Muiden.

Now suppose that www.9292ov.nl says that there is no trip that starts in Amsterdam at 9:10 and brings you to Muiden at 9:45. Then we will act as if there *is* no such trip, but why? This is an example of *closed–world reasoning*, which is appropriate if one may assume that the database lists all available positive information.

Formally, closed–world reasoning (in the version we prefer) differs from classical logic in the syntactic, semantic, and consequence parameters.

Syntactically, the occurrence of $\rightarrow$ is restricted to formulas of the form $p_1 \wedge \ldots \wedge p_n \rightarrow q$. This amounts to changing the recursive definition of the propositional language; iteration of implication is not allowed, and neither are occurrences of negation in antecedent and consequent.

Semantically, $\wedge$, $\vee$ have their customary classical interpretation, but $\rightarrow$ has a special closed–world interpretation given by

1. if all of $p_1, \ldots, p_n$ are true, then so is q;
2. if one of $p_1, \ldots, p_n$ is false and there is no other implication with q as a consequent, q is false;
3. more generally: if for $i \leq k$, $p_1^i \wedge \ldots \wedge p_{n_i}^i \rightarrow q$ are all the formulas with q in the consequent, and if for each $i \leq k$ one of $p_1^i, \ldots, p_n^i$ is false, then q is also false.

The most important technical feature of closed–world reasoning is that the associated consequence relation is *nonmonotonic*. We have encountered the monotonicity property of classical logical consequence when discussing the material implication; we repeat it here for convenience. As we have seen, the Bolzano–Tarski definition of validity of an argument $\varphi_1, \ldots, \varphi_n/\psi$ is: for all models $\mathcal{M}$ such that $\mathcal{M} \models \varphi_1, \ldots, \varphi_n$, also $\mathcal{M} \models \psi$. Given this definition, $\models$ is *monotone* in the sense that $\varphi_1, \ldots, \varphi_n \models \psi$ implies $\varphi_1, \ldots, \varphi_n, \theta \models \psi$ for any sentence θ. closed–world reasoning, however, is not monotonic in this sense: the inference from the database

> there is no trip which starts in Amsterdam CS at 9:10 and ends in Muiden before 9.45,"

licensed by closed–world reasoning, may be destroyed by additions to the database (e.g., a fast Interliner bus).

2.3.2 Unknown Preconditions

Real-world actions come with scores of preconditions which often go unnoticed. My action of switching on the light is successful only if the switch is functioning properly, the house is not cut off from electricity, the laws of electromagnetism still apply. It would be impossible to verify all those preconditions; we generally do not even check the light bulb although its failure occurs all too often. We thus have a conditional "if *turn switch* then *light on*" which does not become false the moment we turn the switch only to find that the light does not go on, as would be the case for the classical material implication. An enriched representation of the conditional as a ternary connective shows more clearly what is at issue here: "if *turn switch* and *nothing funny is going on* then *light on*." If we turn the switch but find that the light is not on, we conclude that something is amiss and start looking for that something. But – and this is the important point – in the absence of positive information to the effect that something is amiss, we assume that there is nothing funny going on. This is the closed–world assumption for reasoning with abnormalities, CWA(ab).

This phenomenon can be seen in a controlled setting in an experiment designed by Claire Hughes and James Russell ([131]), the "box task," which lends itself particularly well to a logical analysis using closed world reasoning. This task was designed for analyzing autistic behavior, to which we return in Chapter 9 below.

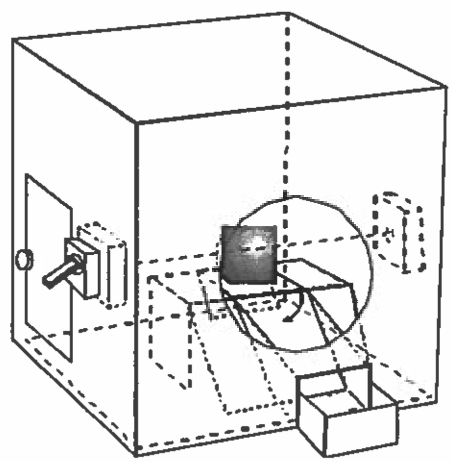

Figure 2.4 Hughes and Russell's box task. Reprinted from [236,p. 316] by permission of Dunitz.

The task is to get the marble, which is lying on the platform, inside the box. However, when the subject puts her hand through the opening, a trapdoor in the platform opens and the marble drops out of reach. This is because there is an infrared light beam behind the opening, which, when interrupted, activates the trapdoor mechanism. The switch on the left side of the box deactivates the whole mechanism, so that to get the marble you have to flip the switch first. In the standard setup, the subject is shown how manipulating the switch allows one to retrieve the marble after she has first been tripped up by the trapdoor mechanism.

A more formal analysis of the box task could go as follows. The main premise can be formulated as

(1) If you reach for the marble through the opening *and there is nothing funny going on*, you can retrieve the marble.

where the italicized conjunct is the variable, assumed to be present always, for an unknown precondition. This conjunct occasions closed world reasoning of the form

(2) I haven't *seen* anything funny.
:: There *is* nothing funny going on.

Backward chaining then leads to the plan

(3) To get the marble, put your hand through the opening.

Now a problem occurs: the marble drops out of reach before it can be retrieved. Premise (1) is not thereby declared to be false, but is now used to derive

(4) Something funny is going on.

To determine what's so funny, the information about the switch is recruited, which can be formulated as a rule "repairing" (1) as in (5a) or (5b)

(5a) If you reach for the marble, set the switch to the right position, *and there is nothing funny going on*, then you can retrieve the marble.
(5b) If the switch is in the wrong position, there is something funny going on.

Closed–world reasoning with (5b) now yields

(6) If the switch is in the wrong position, there is something funny going on, *but only then*.

Backward chaining then leads to a new plan

(7) To get the marble, set the switch to the right position and put your hand through the opening.

One interesting feature of this analysis is thus that the new plan (7) is constructed from the old one by utilising the variable for the unknown precondition.

This is not reasoning as it is usually studied in the psychology of reasoning, but it is reasoning nonetheless, with a discernible formal structure, and applicability across a wide range of domains. In fact CWA(ab) can be viewed as a definition of validity, as follows. Suppose we have an enriched conditional of the form $p \wedge \neg ab \rightarrow q$, where ab is a proposition letter indicating some abnormality. Suppose furthermore that we have as information about ab the following implications: $q_1 \rightarrow ab, \ldots, q_k \rightarrow ab$, and that this is *all* the available information about ab. Since the implication $\bot \rightarrow ab$ is always true,[15] we may include this (admittedly trivial) statement in the information available about ab.

We now want to say that, given $p, \neg q_1, \ldots, \neg q_n$, q may be concluded. This is tantamount to replacing the information about ab by the single premise

$$ab \leftrightarrow q_1 \vee \ldots \vee q_k \vee \bot,$$

and applying classical validity. Note that as a consequence of this definition, if there is no nontrivial information about ab, the right–hand side of the preceding bi-implication reduces to a falsehood (i.e., $\bot$), and the bi-implication itself to $ab \leftrightarrow \bot$, which is equivalent to $\neg ab$. In short, if there is no nontrivial information about ab, we may infer $\neg ab$. Note that although classical reasoning is used here in explaining the machinery, the closed-world inference itself is non-classical: in classical logic nothing can be concluded from the premises $p, \neg q_1, \ldots, \neg q_n$.

A famous observation can be illuminated from this point of view: Scribner's study of reasoning among the illiterate Kpelle tribe in Liberia (see [243]). Here is a sample argument given to her subjects

> All Kpelle men are rice farmers.
> Mr. Smith[16] is not a rice farmer.
> Is Mr. Smith a Kpelle man?

Subjects refused to answer the question definitively, instead giving evasive answers such as "If one knows a person, one can answer questions about him, but if one doesn't know that person, it is difficult." Scribner then went on to show that a few years of schooling in general led to the classical competence answer.

This result, like those of Luria in the 1930s (see [177]) has been taken as evidence that the illiterate subjects do not understand what is being asked of them: to answer the question solely on the basis of (an inference from) the premises given. Instead, so it is argued, they prefer to answer from personal experience, or to refrain from answering if they have no relevant experience. But this explanation presupposes that the Kpelle subject adopts the material implication as the logical form of the first premise. If, as is more plausible, he adopts a meaning of the conditional which allows exceptions (as we did in discussing the box task), he can only be charged with not applying closed–world reasoning to Mr. Smith. That is, if the Kpelle subject believes he has too little information to decide whether Mr. Smith is abnormal, he is justified in refusing to draw the modus tollens inference. On this account, what the couple of years elementary schooling teaches the child is a range of kinds of discourse in which exactly what to close the world on, and what to leave open, varies with some rather subtle contextual cues.

15. $\bot$ stands for an arbitrary contradiction, while $\top$ is a formula which is always true.
16. "Mr. Smith" is not a possible Kpelle name.

2.3.3 Causal and Counterfactual Reasoning

Counterfactual reasoning occurs when one starts from an assumption known to be false and tries to derive consequences. What would have happened if Hitler had invaded Britain? is a famous example. Such reasoning involves causal reasoning, because one needs to set up plausible chains of events. Counterfactual reasoning has been investigated in preschool children with the aim of establishing correlations with "theory of mind"

Riggs and Peterson [227] devised a "counterfactual" adaptation of the standard false belief task, in which a mother doll bakes a chocolate cake, in the process of which the chocolate moves from the fridge (its original location) to the cupboard. The question asked of the child is now

(*) Where *would* the chocolate *be* if mother hadn't baked a cake?

This question is about alternative courses of events and hence seems to use causal reasoning.

Pragmatically, the formulation of question (*) suggests it must have an answer. The answer cannot come from classical logic, starting from the description of the situation alone: classical logic compels one to ask What else could be the case? reflecting the obligation to consider all models of the data. In particular there would be models to consider in which mother eats all the chocolate, or in which the chocolate evaporates inside the fridge (an event of extremely small, but still nonzero, probability). Of course nothing of the sort happens in actual causal reasoning. There a "principle of inertia" applies, which roughly says: "things and properties remain as they are, unless there is explicit information to the contrary." This can be further spelled out as the closed–world assumption for reasoning about causality (CWA(c)):[17]

1. One assumes that only those events (affecting the entity of interest) occur which are forced to occur by the data – here the only such event is the chocolate's change of location from fridge to cupboard.

2. One also assumes that events only have those causal effects which are described by one's background theory – e.g., turning on the oven does not have a causal effect on the location of the chocolate.

3. No spontaneous changes occur, that is, every change of state or property can be attributed to the occurrence of an event with specified causal influence.

Together these principles suffice to derive an answer to (*). In fact this type of reasoning can be fully formalized in the "event calculus" originally developed in artificial intelligence (see [282] for extensive treatment and references).

17. A fully formal analysis will be given in chapter 9.

Its logical structure is similar to the one detailed in section 2.3.2 as regards properties (1) and (2), but property (3) brings in a new ingredient relating to development over time.

Formally, this can be viewed as yet another twist to the definition of validity: one now obtains a notion according to which the conclusion is evaluated at a later instant than the evaluation time of the premises. The classical definition of validity assumes that the conclusion of an argument is evaluated on models of the premises, thus validating a property like $p \models p$, that is, "on every model on which p is true, p is true." The definition of validity used in CWA(c) allows that models of the conclusion are temporal developments of models of the premises, and in this case we need no longer have $p \models p$. Suppose the models for the premises are evaluated at time t, and the models for the conclusion are temporal developments of these models considered at time $t' > t$. Clearly, even if p is true at time t, that same proposition p may be false at t'.

These considerations allow us to see the connection between closed world reasoning and planning. One feature distinguishing human planning from that of other species is the much increased capacity for offline planning. This involves mentally constructing a model, a structure representing the relevant part of the world, and computing the effect of actions in that model *over time*, taking into account likely events and the causal consequences of the actions performed. The various forms of closed–world reasoning introduced so far have to be combined here to enable the construction of the model and the computation of its development over time. What is interesting here for discussions of domain specificity is that the procedures used to construct models in offline planning can be used as well to construct models of linguistic discourse, for instance the structure of the events described by the discourse (see [282]). It is proposed in the reference cited that offline planning has been exapted[18] for the purposes of language comprehension, viewed as the ability to construct discourse models. If true, this would show an incursion of very general reasoning procedures into the purportedly domain–specific language module. Issues of modularity will crop up throughout.

2.3.4 Attribution of Beliefs and Intentions

An important step in cognitive development is the acquisition of a "theory of mind," the ability to understand that someone else may have beliefs different from one's own. A standard experimental paradigm to test theory of mind is the "false belief task," of which the following is an example (due to Wimmer and Perner [305])

> Children are first told the story:"Maxi and Mummy are in the kitchen. They put some chocolate in the fridge. Then Maxi goes away to play with his friend. Mummy decides to

18. See section 6.2.3 for a definition and discussion of the contrast between exaptation and adaptation.

bake a cake. She takes the chocolate from the fridge, makes the cake, and puts the rest of the chocolate in the cupboard. Maxi is returning now from visiting his friend and wants some chocolate." Children are then asked the test question: "Where does Maxi think the chocolate is?"

Normally developing children will be able to attribute a "false belief" to Maxi and answer "In the fridge" from around age 4 or so.

Another version[19] uses an episode from the Bob the Builder children's television series, in which Bob climbs a ladder to do some repair work on the roof of a house. While Bob is happily hammering, the series' resident gremlin Naughty Spud takes away the ladder to steal apples from a nearby apple tree. After Bob has finished his work on the roof, he makes preparations to climb down. At this point the video is stopped and the child who has been watching this episode is asked: "Where does Bob think that the ladder is?" Again, children below the cut-off age answer: "At the tree."

It is illuminating to view the reasoning leading up to these answers as an instance of closed–world reasoning. What is needed first of all is an awareness of the causal relation between perception and belief, which can be stated in the form: "if φ is true in scene S, and agent a sees S, then a comes to believe φ," where φ is a metavariable ranging over proposition letters $p, q, \ldots$. In other words, seeing is a cause of believing. Thus Maxi comes to believe that the chocolate is in the fridge. An application of the principle of inertia (cf. (3) above) yields that Maxi's belief concerning the location of the chocolate persists unless an event occurs which causes him to have a new belief, incompatible with the former. The story does not mention such an event, whence it is reasonable to assume – using 1 and 2 – that Maxi still believes that the chocolate is in the fridge when he returns from visiting his friend. Viewed in this way, attribution of belief is a special case of causal reasoning, and some correlation with performance on counterfactual reasoning tasks is to be expected. The tasks are not quite the same, however. The causal relation between perception and belief is an essential ingredient in the false belief task, absent in the counterfactual task. There are two sides to this: positively, that a belief may form after seeing something, and negatively, that there are only a few specified ways in which beliefs can form, e.g., by seeing, by being told, and by inference – this negative aspect is an application of closed–world reasoning. Children failing the false belief task could master causal reasoning generally, but fail on the aspects just mentioned. So, assimilating the reasoning involved in theory of mind tasks as a kind of defeasible reasoning potentially provides both a basis for continuity with earlier developmental or evolutionary precursors, and a basis for discontinuity – it is causal reasoning by closed–world assumptions, but causal reasoning by closed–world assumptions of a specific kind. Reasoning

19. Investigated experimentally by the van Lambalgen's students David Wood and Marian Counihan.

about minds is reasoning in a specific domain, but its characterization may be possible by a rather small extension of a logical framework for other domains.

We viewed classical logic as resulting from setting parameters for syntax, semantics, and the consequence relation. We have seen that these settings are appropriate for the domain of classical mathematics, but that they cannot claim universal validity. Other domains require different settings; e.g., closed–world reasoning about databases has only bivalence in common with classical logic. In fact, a wide range of everyday tasks involve closed–world reasoning: e.g., planning, and adapting to failures of plans during their execution, diagnosis of causes, causal reasoning itself, reasoning about mental behavior and states, interpreting speaker's intentions underlying discourse. Each of these domains leads to a logic especially suited to that domain. Reasoners have in general little trouble in selecting the logic appropriate to a domain, although, as we shall see in the course of this book, some psychiatric disorders are accompanied, and perhaps even caused, by inappropriately applied reasoning schemes.

In the following chapters, we will look at several experimental reasoning paradigms to discover evidence of parameter setting at work. It will turn out that for a subject, discovering the right parameters is often the hardest part of a laboratory reasoning task.

At this point, the psychologist reader, from our experience, is likely to be puzzled. "Subjects don't know these logics! These formalisms are just theorists' tools! All these squiggles don't happen in minds! Anyway, all you are doing is redescribing stuff which psychologists know about in terms of pragmatics!" are among typical objections, so we had perhaps better attempt to defuse them. Of course we agree that subjects don't "know these logics" just in the sense that they don't know the grammar of English, but they *do* know these logics just in the sense that they *do* know the grammar of English. Yes, they are also theorists' tools, but we take seriously the possibility that something computationally equivalent is implemented in the mind. In chapter 8 we will show that that implementation doesn't require squiggles. And yes, many of the phenomena we are describing as applying reasoning in non-standard logics have been given descriptions already by psychologists. Our claim is that those descriptions remain ad hoc until they are systematized as we are trying to do here. It is very important that the psychological reader takes us seriously when we claim that these logics are in the mind, but equally important that they realise that that claim does not bring all the baggage usually ascribed to it.