Chapter 2

The Practical Logic of Cognitive Systems

> ... [T]he human brain is a highly parallel setup. It has to be.
>
> John Nash, [1954]

2.1 PLCS and Cognitive Systems

The present work is the first volume of *A Practical Logic of Cognitive Systems*. We here concentrate on the analysis of a notion which lies at the very heart of cognitive competence. The notion is relevance, and its centrality is attested to by the considerable facility with which beings like us ignore irrelevancies and 'stay on point' in the performance of our cognitive tasks. In so saying, we have a particular conception of what it is to *be* a cognitive agent, and accordingly, of how we should think of a logic of cognitive agency. Offering a rudimentary description of this logic, *PLCS*, is the principal business of the present chapter.

We wish to lay some emphasis on the fact that we are here attempting to run on two tracks concurrently. We want, of course, to get relevance right. But we also wish to develop *PLCS*, indeed, to embed the theory of relevance in it. For various reasons, both expository and tactical, we do not wade right in with the account of relevance, but rather we devote some time to describing and motivating *PLCS*. Relevance takes over in Chapter 5, and holds centre-stage for the remainder of the book. Readers who are impatient to be getting on with relevance can skip the preamble on *PLCS* and move directly to page 69.

Even so, it is possible to say now in a wholly general and informal way that information is relevant when it helps things get done. *Relevant information is information that is helpful in certain ways.*

We begin with the notion of a cognitive system. Intuitively, a cognitive system is any functioning reader of this book, or institutional agent, such as NASA, The Abductive Systems Group, or present-day neurobiology. Fundamental to the idea of a cognitive agent is that of a being or a device that processes information under conditions that qualifies the output as one or more of a class of states typified by belief restructuring and decision. In so doing, the cognitive agent exploits available cognitive assets or resources, thus facilitating the end-performance. At this stage, there is no reason to assume that cognitive agents are required to possess consciousness or that cognitive processing even by conscious agents needs always to be conscious.

We begin the account of relevance with what Hans Herzberger once called *primordial beliefs* [Herzberger, 1982, p. 133]. Primordial beliefs about something S are those held with such conviction that one is initially prepared to require of any theory of S that it formally sanction them. We say 'initially' because, as is sometimes the case, a theory of S evolves in such a way as to constitute a case for modifying the S-intuition that, so to speak, got the theory up and running in the first place. (A case in point — rather extremely so — is a theory of consciousness that ends up saying or being tempted to say that there is no such thing as consciousness. See e.g. [Dennett, 1988; Lewis, 1990].)

For us there are two primordial intuitions on which we are prepared to found a theory of relevance:

1. Cognition for beings like us is essentially and irreducibly a matter of making economical use of the requisite cognitive resources, which typically are in comparatively short supply.

2. A centrally important factor in the efficiency of cognitive processes is the comparative facility with which beings like us stay on point and evade irrelevance.

'What is wrong with irrelevance?', it might be asked. There is a twofold answer to this question: it impedes the realization of our cognitive goals, and it is *wasteful.*

Having pledged ourselves to the founding intuitions expressed by propositions (1) and (2), it is appropriate that we proceed as follows. We should first endeavour to say something about the cognitive economy in which individual human beings operate. We should then state the theory of relevance, and indicate the ways in which it facilitates the functioning of that economy.

2.2 Practical Reasoning

In one sense, all reasoning is practical.[1] All reasoning terminates in an answer to a question, a solution to a problem, a conclusion from some data, or a decision to postpone the quest until further facts are known; even aborted reasoning ('This is getting us nowhere!') produces a kind of termination.

Ordinary usage, even ordinary philosophical usage, gives little direct guidance for fixing the sense of practical reasoning. It is an expression layered with multiple meanings and suggestive of contrasts, among which are these:

ordinary, common *versus* esoteric, specialized
prudential *versus* alethic
moral *versus* factual
informal *versus* formal
precise *versus* fuzzy
conclusion is an action *versus* conclusion is a proposition
premiss is an action *versus* premiss is a proposition
goal-directed, purposive *versus* context-free
applied *versus* theoretical
concrete *versus* abstract
tolerant of incommensurabilities *versus* not

To these we add a further contrast, to which we think it prudent to take particular note of. It is the contrast of

practical *versus* strict

We illustrate with an example. In the game of (ice) hockey, a hat trick is achieved by a player scoring three consecutive goals against the opposition. (There is a counterpart achievement in cricket.) 'Consecutive' here means 'without any goal being scored between the first and the third of this triple by any of the hat-tricker's team-mates.' This is what a hat-trick is *strictly speaking*. But *in practice*, or for *all practical purposes* (including the triggering of bonus clauses in a player's contract), a hat-trick is just three goals in a game by one and the same player, never mind whether he scores them consecutively in our present sense of that term. So conceived of, practicality is resemblance enough to the real thing to be considered the

[1] There is a philosophical tradition in which a practical reason is reason for an action that involves bodily behaviour. Needless to say, not all reasoning is practical in this sense. We ourselves are disposed to think that practical reasoning in this sense hardly carves out a natural kind, so to speak. (See here, e.g., Velleman [2000]).

real thing. Thus, in one sense, 'practical' means 'approximate'. As we shall shortly see, this captures a part of our own conception of the practical.

2.3 Practical Agency

Ours is an agency view of logic. It betokens, as we said, a return to the Laws of Thought approach. On the agency view, logic is a theory of reasoning, a theory of what thinkers do and have happen to them. Correspondingly, a practical logic is a theory of what practical agents think and reflect upon, cogitate over and decide, and act. If the linguistic conception makes it necessary for the logician to say, with care, what sort of thing a language is, the agency view makes it necessary to say, with care, what sort of thing a practical agent is.

We think of practical agency as a *hierarchy* $\mathcal{H}$ of goal-directed, resource-bound entities A of various types. At the bottom of this hierarchy are individual human beings with minimal efficient access to institutionalized databases. Next up are individual human beings who operate in institutional environments — in colleges or government departments, for example, which themselves are kinds of agents. Then, too, there are teams of such people. Further up are disciplines and other corporate entities such as, again, the NASA or Italian physics in the 1930s. The hierarchy proceeds thus from the concrete to the comparatively abstract, with abstract structures being aggregations of entities lower down. Interesting as this metaphysical fact might be, it is not the dominant organizing principle of the hierarchy. The organizing principle is economic. Entities further up the hierarchy command resources, more and better, than those below are capable of.

So conceived, the hierarchy is a poset of objects partially ordered by the relation C of *commanding greater resources than.*[2]

Every agency in this hierarchy $\mathcal{H} = \langle C, A \rangle$ involves, whether by aggregation or supervenience or in some other way, the individual agent. Such agents are thus basic to any logic of agency, and it is to them that we shall concentrate our attention in the present section.

[2]We note in passing the difference of our hierarchical model from Harry Frankfurt's hierarchical model of autonomous action. On this latter conception, the behaviour that an agent makes happen in the fullest sense of that expression is that which is motivated by a desire which the agent desires to have. See Frankfurt [1988, 58–68]. But cf. Bratman [1999, 185–206].

We also note a resource-sensitive approach to cognitive agency in much of the psychological literature. See Simon [1957] and a, by now, large psychological literature ably reviewed in Stanovich [1999] and Gigerenzer and Selten [2001a].

Like all agents in the hierarchy, the individual is a performer of actions in real time. And nearly everything an individual is faced with doing, or is trying to do, can be done at the wrong time. It can be done at a time so wrong as to court equivalence with not doing it at all, or doing some opposite thing. It is not enough that an agent does the right thing, i.e., performs the right action-types. It is often essential that the right thing be done at the right time. As we look upwards at the agency-hierarchy, we see a diminishing susceptibility to exigent timeliness. No one doubts that NASA had a real deadline to meet in the 1960s, culminating in the moon shot. It might have been that the moon program would have been cancelled had that deadline not been met. Even so, individuals are exposed to myriad serious dangers, many of them mortal, that nothing 'up above' will hardly ever know on this scale; and essential to averting such dangers is doing what is required on time, directed by the right information in appropriate quantitites.

The dominant requirement of timeliness bears directly on a further constraint on individual agency. Individuals wholly fail the economist's conceit of perfect information. Agents such as these must deal with the nuisance not only of less than complete information, but with data-bases that are by turns inconsistent, uncertain, and loosely defined. To these are added the difficulties of real-time computation, limited storage capacity and less than optimal mechanisms for information-retrieval, as well as problems posed by bias and other kinds of psychological affect.

The two great scarcities that the individual must cope with are time and information. It is precisely these that institutional agents command more of, and very often vastly more of. With few (largely artificial) exceptions, the individual agent is a satisficer rather than an optimizer, a fact reflected in our distinction between the practical and the strict, and captured by the example of the hat-trick in hockey. It is also, and more centrally, on evidence in the individual's entrenched disposition to forgo truth-preservation or high levels of conditional probability in favour of rougher standards of what's plausible, which deliver the goods with requisite promptness and directness. For the most part, even seeking to be an optimizer would be tactically maladroit, if not actually harmful. The human agent is also highly sensitive to environmental cues, hence is drawn to adaptive strategies of the fast-and-frugal sort [Gigerenzer and Selten, 2001a]. The fact of the robust, continuing presence of human agents on this Earth amply attests to their effective and efficient command of scarce resources. It is a fact in which is evident the human capacity to compensate for scarcities of time and information.

We postulate that the individual agent embodies a **scare-resource compensation strategy**. Here, in rough outline and in no particular or-

der, are the compensation-factors that strike us as particularly important. But, as a quick word of preface, we must lay some emphasis on the point that, as we use the terms 'scarce-resources' and 'scarcity', we intend only a quantitatively comparative rather than a qualitatively comparative notion of scarcity. When beings like us execute our cognitive agendas, the scarcity of the resources that we draw upon is in the general case simply a matter of their being fewer and less of them than in the general case is available to institutional or theoretical agents. Less and few are are not necessarily matters to regret. The individual agent is not placed at an intrinsic disadvantage under these ordinal and cardinal comparisons, although there are particular cases in which paucity of information or time or fire-power does indeed redound to the agent negatively. In such cases, the harm done by the scarcity is the difficulty it creates for executing the tasks in question against the requisite standards of satisfactory performance. This is an affliction that can apply to agents of all types, and is not a discouragement reserved for individual or practical agents, still less for practical agents *in the general case*. In the general case, the quantitatively comparative resource-scarcities with which the practical agent must deal with are compensated for by the degree of rigour imposed by performance standards appropriate to the kind of agent an individual is.

- Human beings are natural **hasty generalizers**. It was a wise J.S. Mill who observed [Mill, 1974] that the routines of induction are not within the grasp of individuals, but rather are better-suited to the resource capacities of institutions. The received wisdom has it that hasty generalization is a fallacy, a sampling error of one sort or another. The received wisdom may be right, but if it is, individual human agency is fallacy-ridden in degrees that would startle even the traditional fallacy-theorist.[3] Bearing on this question in ways that suggest an answer different from the traditional one is the fact that the individual's hasty generalizations seem not to have served his cognitive and practical agendas all that badly. Upon reflection, in the actual cases in which a disposition towards hasty generalization plays itself out, the generalizations are approximately accurate, rather than fallacious errors, and the decisions taken on their basis are approximately sound, rather than exercises in ineptitude. Not only is the individual agent a hasty generalizer, he is a hasty generalizer who tends to get things more or less right.

- How is it possible that there be a range of cases in which projections

[3]On what we are calling the traditional account of fallacies, hasty generalization is always an error. For a contrary view see Woods [2003].

from samples are so nearly right, while at the same time qualifying as travesties of what the logic of induction requires? The empirical record amply attests to a human being's capacity for pre-inductive generalization and projection. It would appear that exercise of this capacity involves at least these following factors, some of them structural, some of them contextual. The *pre-inductive* generalizer does not generalize to universally quantified conditional propositions. Rather he generalizes to **generic** propositions. There is a world of difference between 'For all x, if x is a tiger then x is four-legged" and "Tigers are four-legged.' The former is falsified by the truth of any negative instance, whereas the latter holds true even in the light of numerous negative instances of certain kinds. We could characterize this difference by saying that universally quantified conditional statements are highly *brittle*, whereas generic statements are *elastic*. Generic propositions are essential to what is sometimes called *stereotypical* reasoning. Clearly not all stereotyped reasoning is defective.

- The elasticity of what the pre-inductive generalizer generalizes to serves the generalizer's interests in other ways, two of which are particularly important. One is that the individual agent is a **fallibilist** in (virtually) everything he thinks and does. The other is that the individual agent has the superficially opposite trait of rather high levels of accuracy in what he thinks and does when operating at the level ordained for him by the hierarchy of agency. Generalizing to generic statements is a way of having your cake and eating it too. It is a way of being right even in the face of true exceptions. It is a way of being both right and mistaken concurrently.

  Generalizing in this way also works a substantial economy into the individual's cognitive effort. It comes from the smallness of its samples and the elasticity of its generalizations. Generic inference is inference from small samples under conditions that would make it a fatally stricken induction. We see in this the idea of the affordable mistake. Generic inference is not truth-preserving. One can be wrong about whether Pussy the tiger is four-legged even though one is right in holding that tigers are four-legged. Affordable mistakes are like small infections that help train up the immune system. Just as an infant's summer sniffles is an affordable (in fact, necessary) infection, so too are the small errors of the cognitive agent which provide him evolving guidance as to the freedom and looseness with which to indulge his predilection for comparatively effortless generalizations. Baby's summer cold loops back benignly in the discouragement of more serious

illness. Affordable mistakes loop back benignly in the discouragement of serious error. We can now see that the old saw of *learning from our mistakes* has a realistic motivation. We do not learn from mistakes that kill us.

- What is it about such samples that sets them up for successful generic inference? It would appear that the record of generic inference is at its best when samples, small as unit sets though they may be, are samples of **natural kinds**. There has been a good deal of philosophical controversy about whether natural kinds actually exist; about whether the putative difference between natural kinds and conventional kinds turns on a principled metaphysical distinction. Certainly there is nothing like a settled consensus as to how the distinction should be applied. Perhaps this tells against our here using the concept in any particular theory-laden way, but it leaves it open that we introduce it as a term from unanalysed common sense (but see here [Fodor, 1998, chapter 7]). Even so, we should not disdain this literature from psychology and computer science in which concepts resembling that of natural kinds seem to be doing useful work, concepts such as *frame* [Minsky, 1975], *prototype* [Smith and Medin, 1981], and *exemplar* [Rosch, 1978]. Then, too, there is a large literature from linguistics in which the semantics of natural kind noun phrases is intimately bound up with factors of genericity [Krifka *et al.*, 1995, pp. 63–94].

  Philosophical particularities aside, the empirical record testifies to our capacity for classifying sensory stimuli in ways that reflect similarities and differences that strike us as inhering things as they really are. There is ample evidence to suggest that our classifications originate with primitive devices of *type-recognition* together with the mechanisms of fight and flight. It is significant that some of our most successful and most primitive inferences involve the recognition of something as dangerous. Generic inference is part and parcel of such strategies. Just as our capacity for recognizing natural kinds exceeds the comparatively narrow range of immediately dangerous kinds, so too does our capacity for generic inference exceed the reach of fight–flight recognition triggers. But whether in fight–flight contexts or beyond, natural kinds and generic inference are a natural pair. It is an arrangement again favouring the economic — a compensation strategy for the scarcity of time and information — but not noticeably at the cost of error. If generic inferences from natural kind samples are not *quite* right, at least they don't kill us. They don't even keep us from prospering.

- The fallibilism of generic inference is also evident in its relation to **defaults**. A default is something taken as holding, taken to be true, in the absence of indications to the contrary [Reiter, 1980]. It is closely related to and may partially be characterized by a process known as 'negation-as-failure'. Most of what passes for common knowledge is stocked with defaults, and generic inferences in turn are inferences to defaults. Default reasoning is inherently conservative and inherently defeasible. Defeasibility is the cognitive price one pays for conservatism. And the great appeal of conservatism is also economic. Conservatism is populated with defaults in the form 'X is what people have thought up to now, and still do.' Conservatism is a method of default-collection. It bids us to avoid the cost of fresh thinking, and to make do with what others have thought before us (and, experienced and remembered, too).

- Conservatism places a premium on what is already well-received.[4] On the face of it, conservatism is the **ad populum** fallacy in endemic form. Here, too, we might grant the received wisdom (and note the large irony), and concede that individual agents are notorious fallacy-mongers on a scale not dreamed of even by the traditional fallacy theorist. But as we said in our examination of a similar indictment of hasty generalization, there are factors which seem to cut across so harsh a condemnation. One is that we are, by and large, enormously well-served by the trust we place in the testimony of others. This needs to be understood. The full account, even if we could furnish it, is beyond the scope of this chapter, but certain features stand out, and should be mentioned. Popular beliefs are what Aristotle called *endoxa*. They are 'reputable opinions', the opinions of everyone or of the many or of the wise. The mere fact of popular opinion triggers an abduction problem. What best explains that p is a proposition believed by everyone? An answer, which certainly can be criticized in respect of certain particular details, but which cannot convincingly be set up for general condemnation is that p's universal acceptance is best explained by supposing that p is true or that a belief in p is reliable. What is loosely called common knowledge is an individual's (or an institution's or a society's) inventory of endoxa. What is especially striking about common knowledge is that it is acquired by an individual with little or no demonstrative effort on his own part, and with attendant economies of proportional yield.

[4] Notwithstanding the joke in which 'a Conservative is one who is enamoured of existing perils, as distinguished from the Liberal, who wishes to replace them with others.' (Ambrose Bierce, *The Devil's Dictionary*).

- It is evident therefore that individual agents depend for what they think and how they act upon the sayso of others, on the more or less uncritical and unreflective testimony of people who by and large are strangers. Here is yet another respect in which the conduct of human agents would seem to fall foul of the received opinion of fallacy theorists (let us not forget that the *endoxa* of the wise are not guaranteed to be true!). For it would appear that individual agents are programmed to commit and implement the programme on a large scale, the **ad verecundiam** fallacy. But as before, the actual record of thoughts and actions produced by such dependencies is rather good; most of what we think in such ways is not especially inaccurate and, in any case, not inaccurate enough to have made a mess of the quotidian lives of human individuals. We may suppose, therefore, that the traditional fallacies of hasty generalization, *ad populum* and *ad verecundiam* are hardly fallacious as such (e.g., when considered as an individual's strategies or components of strategies for practical action), but are fallacies only under certain conditions. We shall return to this point below.

  It has long been known that human life is dominantly social, and that individual agents find cooperation to be almost as natural as breathing. The routines of cooperation transmit to an individual nearly all of the community's common knowledge that he will ever possess. Even though the complete story has yet to be told, cooperation has received the attention of attractive and insightful theories (e.g. [Axelrod, 1984; Coady, 1992] and [Govier, 1988b]).

  There is a natural and intuitive contrast between accepting something on the sayso of others and working it out for oneself. Cross-cutting this same distinction is the further contrast between accepting something without direct evidence, or any degree of verification or demonstrative effort on the accepter's part, and accepting something only after having made or considered a case for it. The two distinctions are not equivalent, but they come together overlappingly in ways that produce for individual agents substantial further economies.

  Perhaps this is the point at which to emphasize that in our conception the individual is not the artefact of the same name championed by European thinkers of the seventeenth and eighteenth centuries. We demur from the notion (the decidedly odd notion, as we see it) that an individual's social relationships are merely contingent to his rationality. On the contrary, an individual's cognitive and decisional competence is in significant part constituted by his social relationships.

If this is right, it will matter for what we take a logic of individual cognitive and decisional agency to be. We will have more to say on this later, but will note in passing the prima facie attractions of a dialogue logic, as a formalized description of the individual agent.

Such additional economies are the output of two regularities evident in the social intercourse of agents. One has been dubbed the reason rule:

> *Reason Rule*: One party's expressed beliefs and wants are a *prima facie* reason for another party to come to have those beliefs and wants and, thereby, for those beliefs and wants to structure the range of appropriate utterances that party can contribute to the conversation. If a speaker expresses belief X, and the hearer neither believes nor disbelieves X, then the speaker's expressed belief in X is reason for the hearer to believe X and to make his or her contributions conform to that belief. [Jacobs and Jackson, 1983, 57], [1996, 103].

The reason rule reports an empirical regularity in communities of real-life discussants. Where the rule states that a person's acceptance of a proposition is reason for a second party to accept it, it is clear that 'reason' means 'is taken as reason' by the second party. Thus a descriptively adequate theory will observe the Jacobs–Jackson regularities as a matter of empirical fact. This leaves the question of whether anything good can be said for these regularities from a normative perspective. If normativity is understood as a matter of instrumental value, it would appear that the reason rule can claim some degree of normative legitimacy. Not only does it produce substantial economies of time and information, it seems in general not to overwhelm agents with massive error or inducements to do silly or destructive things. The reason rule describes a default. Like all defaults, it is defeasible. Like most defaults, it is a conserver of scare resources. And like many defaults, it seems to do comparatively little cognitive and decisional harm.

There is a corollary to the reason rule. We call it the *ad ignorantiam* rule:

> *Ad Ignorantiam Rule*: Human agents tend to accept without challenge the utterances and arguments of others except where they know or think they know or suspect that something is amiss, or when not challenging involves some cost to themselves.

Here, too, a good part of what motivates the *ad ignorantiam* rule in human affairs is economic. People don't have time to mount challenges every time someone says something or forwards a conclusion without reasons that are transparent to the addressee. Even when reasons are given, social psychologists have discovered that addressees tend not to scrutinize these reasons before accepting the conclusions they are said to endorse. Addressees tend to do one or other of two different things before weighing up proffered reasons. They tend to accept this other party's conclusions if it is something that strikes them as *plausible*. They also tend to accept the other party's conclusion if it seems to them that this is a conclusion which is within that party's competence to make — that is, if he is seen as being in a position to know what he is talking about, or if he is taken to possess the requisite expertise or authority. (See, e.g., [Petty and Cacioppo, 1986; Eagly and Chaiken, 1993; Petty *et al.*, 1981; Axsom *et al.*, 1987; O'Keefe, 1990], and the classic paper on the atmosphere effect, [Woodworth and Sells, 1935]. But see also [Jacobs *et al.*, 1985].) We see, once again, the sheer ubiquity of what traditionalists would call — overhastily in our view — the *ad verecundiam* fallacy.

- We see the individual agent as a processor of information on the basis of which, among other things, he thinks and acts. Researchers interested in the behaviour of information-processors tend to suppose that thinking and deliberate action are modes of **consciousness**. Studies in information theory suggest a different view. Consciousness has a narrow bandwidth. It processes information very slowly. The rate of processing from the five senses combined — the sensorsium, as the Mediaevals used to say — is in the neighbourhood of 11 million bits per second. For any of those seconds, something fewer than 40 bits make their way into consciousness. Consciousness therefore is highly entropic, a thermodynamically costly state for a human system to be in. At any given time there is an extraordinary quantity of information processed by the human system, which consciousness cannot gain access to. Equally, the bandwidth of language is far narrower than the bandwidth of sensation. A great deal of what we know — most in fact — we aren't able to tell one another. Our sociolinguistic intercourse is a series of exchanges whose bandwidth is 16 bits per second [Zimmermann, 1989].

  Conscious experience is dominantly linear. Human beings are notoriously ill-adept at being in multiples of conscious states at once. And time flows. Taken together these facts loosely amount to an opera-

tional definition of the linearity of consciousness. Linearity plays a role in the cognitive economy that tight money plays in the real economy. It slows things down and it simplifies them. Linearity is a suppressor of complexity; and reductions in complexity coincide with reductions in information.[5]

Psychological studies indicate that most of our waking actions are unattended by and unshaped by mental states.[6] This mindlessness of ordinary waking human behaviour is a kind of coping. Consider a case in which we are watching a short-order cook working at full blast at midday in New York. It is easy to see his behaviour as connectionist and mindless, as behaviour reflecting repertories of different skills which he draws upon concurrently and distributively, and without a jot of reflection when things are going well.

If these psychological studies are right, the received view is wrong. Conversation would just be linguistic coping. If so, the individual discussants are less often in a state of belief than many theorists suppose; and when someone is telling us, say, about the amenities of Amsterdam, though he tells us the truth, he is not transmitting any current mental state and he is not inducing new mental states in us, unless perhaps what he tells us is surprising. When we stop and think — when we put a temporary (and expensive) halt to coping — we find that in what we do in the world we are infrequently the owner of mental states, infrequently the possessor of beliefs. It is a respectable way of being mindless.

It is now evident that we must amend the claim that individual agents suffer from a scarcity of information. In so doing, however, we are able to lend appropriate emphasis to what remains true about that proposition. In pre- or subconscious states, human systems are awash in information. Consciousness serves as an aggressive suppressor of information, preserving radically small percentages of amounts available pre-consciously. To the extent that some of an individual's thinking and decision-making are subconscious, it is necessary to postulate devices that avoid the distortion, indeed the collapse, of information overload. Even at the conscious level, it is apparent that various

[5] We note in passing that the sheer paucity of information possessed by human consciousness at any given time contrasts with environments known to be fuzzy. Fuzziness, unlike probability, is unchanged by arbitrarily large increases in information.

[6] This is not a claim that everyone would endorse. Some would insist on the qualification 'conscious'. Advocates of Intentional Psychology (*IP*) tend to see such behaviour as caused by propositional attitudes, whose presence does not invariably require consciousness.

constraints are at work to inhibit or prevent informational surfeit. The conscious human thinker and actor cannot have, and could not handle if he did have, information that significantly exceeded the limitations we have been discussing. This makes the economic aspect of an agent's conscious thought and action an *ecosystemic* matter as well [Gigerenzer and Selten, 2001b, 9]. Human beings make do with slight information because this is all the information that a conscious individual can have.

Human agents make do with scarce information and scarce time. They do so in ways that make it apparent that in the general case they are disposed to settle for *comparative* accuracy and *comparative* sensibleness of action. These are not the ways of error-avoidance. They are the ways of fallibilism. Error-avoidance strategies cost time and information, except where they are trivial. The actual strategies of individual agents cannot afford the costs and, in consequence, are risky. As we now see, the propensity for risk-taking is a structural feature of consciousness itself. It might strike us initially that our fidelity to the reason rule convicts us of gullibility and that our fidelity to the *ad ignorantiam* rule shows us to be lazily irrational. These criticisms are misconceived. The reason rule and the *ad ignorantiam* rule are strategies for minimizing information overload, as is our disposition to generalize hastily.

Consciousness makes for informational niggardliness. This matters for computer simulations of human reasoning. That is, it matters that there is no way presently or foreseeably available of simulating or mechanizing consciousness. Institutional agencies do not possess consciousness in anything like the sense we have been discussing. This makes it explicable that computer simulations of human thinking fit institutional thinking better than that of an individual. This is not to say that nothing is known of how to proceed with the mechanization of an individual's conscious thinking. We know, for example, that the simulation cannot process information in quantities significantly larger than those we have been discussing here.

Consciousness is a controversial matter in contemporary cognitive science. It is widely accepted that information carries negative entropy. Against this is the claim that the concept of information is used in ways that confuse the technical and common sense meanings of that word, and that talk of information's negative entropy overlooks the fact that the systems to which thermodynamic principles apply with greatest sure-footedness are *closed*, and that human agents are not.

The complaint against the over-liberal use of the concept of information, in which even physics is an information system (Wolfram [1984]), is that it makes it impossible to explain the distinction between energy-to-energy transductions and energy-to-information transformations. Also singled out for criticism is the related view that consciousness arises from neural processes. We ourselves are not insensitive to such issues. They are in their various ways manifestations of the classical mind–body problem. We have no solution to the mind–body problem, but there is no disgrace in that. The mind–machine problem resembles the vexations of mind–body, both as to difficulty and to type. We have no solution to the mind–machine difficulty. There is no disgrace in that either.

For individual agents it is a default of central importance that most of what they experience, most of what is offered them for acceptance or action, stands in no need of scrutiny. Information-theoretic investigations take this point a step further in the suggestion that consciousness itself is a response to something disturbing or at least peculiar enough to be an interruption, a demand — so to speak — to pay attention.

Most of the information processed by an individual agent he will not attend to, and even if it is the object of his consciousness he will attend to in as little detail as the exigensies of his situation allow. Arguing is a statistically non-standard kind of practice for human agents, but even when engaged in it is characterized by incompletions and shortcuts that qualify for the name of enthymeme. The same is true of reasoning, of trying to get to the bottom of things. In the general case, the individual reasoner will deploy the fewest resources that produce a result which satisfies him. Here is further evidence that individuals display a form of rationality sometimes called 'minimal', [Cherniak, 1986],[7] or 'bounded' [Gigerenzer and Selten, 2001a]. In addition to features already discussed in this chapter, the minimal or bounded rationalist is, when he reasons at all, a non-monotonic reasoner and in ways that are mainly automatic, the successful manger of belief-sets and commitment-sets that are routinely inconsistent. Much of what makes for the inconsistency of belief-sets comes from the inconsistency of deep memory storage and further aspects of inconsistent belief-sets flow from the inefficiencies of memory retrieval.

The structure of minimal or bounded rationality shows the individual agent to be the organic realization of a non-monotonic, paraconsistent base

[7] In fact, it is better thought of as *minimalist* rationality, the rationality involved in making do with scarce resources.

logic, features which our logic must take care to embed. There is little to suggest that the strategies endorsed by classical logic and most going non-standard logics form more than a very small part of the individual agent's repertoire of cognitive and coping skills. If it is true that individuals are in matters of non-demonstrative import *pre*-inductive rather than inductive agents, the same would also appear to be the case as regards deduction. If so, human individuals are not the wet-wear for deductive logic, at least in the versions that have surfaced in serious ways in the sprawling research programmes of modern logic. There is a particularly interesting reason for this. If we ask what the value of deductive consequence is, the answer is that it is a guarantee of truth-preservation. Guaranteed truth-preservation is a guaranteed way of avoiding error.[8] But individual agents are not in the general case dedicated to error-avoidance. So for the most part the routines of deduction consequence do not serve the individual agent in the ways in which he is disposed (and programmed) to lead his cognitive and decisional life. This is not to say that agents do not perform deductive tasks even when performing on the ground level of our hierarchy. There is a huge psychological literature about such behaviour (accessibly summarized in Manktelow [1999]) and the point rather is that deductive thinking is so small a part of the individual's reasoning repertoire.

2.4 Practical Logics

In our description of it so far, we have left the theory of practical reasoning a fairly underdetermined affair. There is a desirable utility in such flexibility. We leave ourselves free to consider the pros and cons of extending or adapting our approach in many possible ways, and in so doing availing ourselves of the benefit of work already done and on the record. There is a lot of it, too, whether temporal logics (e.g., van Benthem [1991]), logics of action (e.g., Davidson [1980], Brand and Walton [1975], Brand [1984]), dynamic logic (e.g. van Benthem [1996], van Benthem *et al.* [2001] and Gochet [2002]), not to forget the huge literature on deontic logic, and the practical logics of the early pragmatic philosophers (e.g., Dewey [1938] and Schiller [1912]).

There are multiples of different ways of finishing a theoretical product from its relatively modest beginnings as a logic supplemented by designated resources for the treatment of action and time. This leaves the research community with multiples of chances of coming up with finished products

[8]That is, of avoiding errors not already in his database or his premiss-set or which follows from false prior information.

that receive and deserve consensus of a sort that we do not yet see much in evidence. Even so, it is an attraction of our approach that it serves the desirable end, and achieves the welcome economy, of a principled and modest shortening of the list of attributes on whose behalf the adjective 'practical' is invoked. If we return to the list developed in section 2.2 of the present chapter, it is clear that our logic sanctions some deletions.

A practical logic in our sense is not restricted to the study of reasoning about ordinary or commonplace matters. Nothing precludes the practical reasoner rushing to finish an arcane proof under press of his publisher's deadline.

A practical logic in our sense is no enemy of the alethic or truth-oriented. For example, there is a well-understood role in dialogue logic for parties to enhance their shared databases. In so doing they increase their resources for making more direct cases for various actions.

Practical logic pertains to moral reasoning but is not restricted to it. Nor does it exclude factual reasoning. (See above.)

Practical logic is no enemy of formality. Where appropriate it can involve express manipulation of logical forms; and even where reasoning is not formal in so sharply structural a way, practical logic is amenable to other grades of formal treatment. (Woods [1980], [1989], [2003, Chapter 15], van Eemeren et al. [1996]; cf. Johnson [1996, 120]).

Practical logic is not inherently about fuzzy reasoning, but can be extended to a fuzzy logic (e.g., Zadeh [1975], Chang and Lee [1975], Lee [1972], Przelecki [1976] and Hájek [1998]) or to a logic of vagueness (e.g. Tye [1990], Williamson [1994]) in those cases in which reasoning requires attending to in a more or less direct way the fuzziness of terms or, to fuzzy states of affairs. There are those who argue that practical reasoning is inherently fuzzy in just this sense. In our view this is an open question. (See, e.g., Woods [2000].)

Practical logic subsumes but is not restricted to what Aristotle calls practical syllogisms. The same is true for the adaptation of the same idea in Gabbay and Woods [1999]. In a practical logic of the kind under review, a move in a dialogue always occasions an action by the other party, even though his action needn't be the action, if any, implied or suggested by his *vis-à-vis* premisses. For example, one party may say to the other: 'So, you see, you ought to mow the lawn now.' One way for the second party to react to that move is to start mowing the lawn. This is an explicit action that will also serve as implicit acceptance of his interlocutor's claim. Or he might reply, 'Yes, I really should be mowing the lawn,' which is explicit acceptance and intimation of an action yet to be taken. A third answer is 'Like hell!' which is an explicit (and emphatic) rejection. A fourth is

phoning a friend to arrange for a golf game, which is explicitly not mowing the lawn and implicit rejection of the argument that called for it.

Neither do we think that practical logic should be reserved for reasoning involving incommensurabilities. Incommensurability is ambiguous (Gray [2000]). In its most basic sense, reasoning from incommensurabilities is reasoning of a *pluralistic* kind. It is illustrated by the following schema.

1. Harry and Sarah value both friendship and patriotism.

2. Friendship and patriotism though different, and sometimes behaviourally non-co-satisfiable, are incomparable values.

3. In circumstances K, Harry opted for friendship and Sarah for patriotism.

4. Both acted rightly. Period.

It is true that normative reasoning is often occasion for judgements of incommensurability, but this is also sometimes true of scientific thinking. Pluralism abounds in logic, for example. And paraconsistent logics have been purpose-built to accommodate incommensurabilities (in the form of outright inconsistencies) whether in set theory or quantum mechanics (Priest [1998], and Brown [1993]). However, the incommensurability view of practicality intersects with our own conception, in the following way. Sometimes when faced with an incommensurability or an inconsistency, the practical (i.e., individual) agent has no realistic option but to let it be. He may lack the resources to adjust his database for consistency, which puts him in a situation in which he must think or act *in spite* of inconsistency. On the other hand, the very resources that an individual agent sometimes lacks are progressively available to agents of higher type.

The only interpretation that we ourselves are able to give the applied *versus* theoretical distinction in practical logic is one of the following inequivalent pair. First is the distinction between reasoning in a fully interpreted as opposed to a merely semi-interpreted vocabulary. To achieve its generality economically, a practical logic may operate with a semi-interpreted object language. But it will also have the means of giving its theorems full interpretations. (This is tricky. No such procedure will preserve formal invalidity. See here [Woods, 2003, chapter 15].) The second way of drawing our present distinction is to see it as an instance of a particular way of construing the descriptive-normative distinction. In a widely accepted view of this latter, the task of finding a descriptive application of a normative theory is a matter of (a) finding the discrepancies between them, and (b) accounting for the descriptive deviations as approximations to the ideal

conditions, full compliance with which would qualify as normatively perfect performance.

Unless we are mistaken, the sense we have proposed to give our logic offers guidance on the applicability of other distinctions appropriated by those intent on giving 'practical' some principled meaning. The purported distinction between concrete and abstract is handled by what we have said about the applied-theoretic distinction. Also there covered is the distinction between unregimented language and canonical notation. The distinction between a natural logic and an artificial logic can be captured by the distinction just mentioned. Alternatively it is the distinction between the psychologically real and the psychologically ideal, which we have already discussed.

There is also an intuitive distinction between tasks whose performance requires little or no tutelage and those whose performance require specialized technical information. Cutting across this distinction, but in ways that produce some degree of overlap, is the contrast between ordinary and esoteric subject matters. If we wanted the distinction between practical and theoretical logics to be constrained by these contrasts, they would push in somewhat different directions; and formal logics such as first order quantification theory would elude classification altogether. We ourselves see little appeal in the first of these proposed criteria. A logic that attempted to give some insight into what goes on when an individual attempts to solve the Four Colour Problem is as much a practical logic as any that attempts to elucidate an agent's choice of breakfast cereal. Neither are we persuaded that, for our purposes here, there is any abiding value in the contrast between the ordinary and everyday and (say) the business of quantum non-locality in physics. A more fruitful way of drawing the contrast between a practical and theoretical logic is by piggy-backing on our distinction between a practical and a theoretical agent. The value of so doing (apart from the naturalness of the concurrence) is that it is very much less necessary to discredit a logic for its failure to model realistically actual human behaviour. Most mainstream logic since 1879, and most direct rivals of it, are subject to this failure. They fail for the most part because their strategies are too complex for the computational capacities of human individuals or, because their latitude in other respects (e.g., monotonicity) exceeds actual human reach. True, some mitigation of these misrepresentations can be found in the notion of idealization; but idealization is a more fraught device than is usually recognized (one cannot idealize at will). Even so, many of these logics, which fail as principled descriptions of what human individuals are capable of, succeed or come closer to succeeding as formalized accounts of what *institutional* agents are capable of. So a decision to regulate the dis-

tinction between practical and theoretical logics in this way has the virtue, even on an idealized agent approach to logic, of saving much of what fails as a practical logic as what succeeds as a theoretical logic.

We have already said that we find ourselves somewhat vexed by the descriptive–normative distinction in logic. As we bring this section to a close, it would be helpful if we could briefly shed some light on our reservation.

2.4.1 The Method of Intuitions

There is a considerable body of opinion in the century and a quarter since 1879 that a logician's job is axiomatization and that axioms are what the logician finds to be most intuitive. Much the same view can be found among logicians who favour natural deduction approaches. Here, too, one's choice of structural and operational rules is seen as a matter of what strikes the theorist as most intuitively correct. Much the same *modus operandi* is evident in other disciplines, especially abstract disciplines that lack — in any direct way anyhow — empirical checkpoints. In philosophy this approach is the heart and soul of conceptual analysis in the manner of G.E. Moore and an entire generation which fell under his influence.

The method of analytic intuitions raises a fundamental methodological question. Given that an intuition is what the theorist antecedently believes, and that a fundamental intuition is what he believes utterly, is there any good reason to suppose that intuitions are *epistemically privileged*? Is there any reason to suppose that what the theorist believes utterly qualifies as knowledge? If the answer is Yes, the essential methods of conceptual analysis are confirmed. If the answer is No, the methodology of the abstract sciences must take this into account.

One attraction of the method of analytic intuitions in logic is that it secures a comfortable purchase on the shelf of normativity. It allows for it to be the case that a human being *should* reason in such-and-such a way, if the logician-theorist's intuitions lend support to a rule or a theorem to the same effect. But shorne of the comforts of the method of analytic intuitions, the normatively minded logician will find less desired normativity a lot more difficult to get a sure grip on. It may be that such a theorist would be well-served in taking the following approach.

> First, he might try to make this account conform closely to how in the general case practical agents actually perform under the conditions the theory takes note of.
>
> Secondly, he might also try to take note of what in actual practice is regarded as mistakes or errors.

If he does both these things, we will say that his account is *descriptively adequate*. The sixty-four dollar question is whether:

> the theorist obtains a serviceable standard of normativity by putting it that a practical agent performs as he should if his performance conforms to what his fellows do and is not marred by mistakes in the sense of a paragraph ago. The answer is that we propose is strongly in the affirmative.

There is an ancient way of characterizing the practical. It is to be found in the contrast between Practical and Theoretical Reason, between *phronesis* and *epistēmē*. Perhaps we now have the wherewithal to characterize this contrast in ways that would be found credible by present-day readers. Accordingly, we repeat our proposal that Practical Reason be thought of as a repertoire of skills characteristic of the lower strata in the hierarchy of agency, that Theoretical Reason be thought of as sets of skills characteristic of higher up, and that the contrast be seen as a matter of degree — a matter of how low down and how high up the agent in question chances to be. Here is a suggestion which preserves the truth that *all* reasoning is goal-directed, that all reasoning portends *some* kind of action. But it allows us to cross-cut this universality with considerations of indigenous import, in which Practical Reason is characterized by features of the *agent* whose reasoning it is.

It is also well to emphasize that we are taking the agency view of logic, as opposed to the disembodied linguistic view. The distinctions we have been tracking and the exclusions we have been proposing, have been transacted within the tent of agency logic. Agency logic is the natural home of practical logic, and offers reasonable accommodation to one reasonable conception of theoretical logic. However, it is not our view that the linguistic conception of logic should be rejected. There is nothing good to be said for the idea that we should say no to recursion theory, model theory, proof theory and set theory. This is a book about the practical turn in logic. It obliges us to give sense to what is practical and to give some idea as to where the idea of the practical is best pursued by logical theory. In the end, it is this question which we bring to the distinction between the agency and linguistic conceptions of logic. And, with respect to the matters that concern us here, it is our view that an agency logic is a natural home for practical reasoning and that embodied linguistic logic is not. But saying so is a long way from pleading the exclusion of linguistic logic. We shall amply attest to this assurance when, in Part III of this book, we produce formal models of relevance. (So we aren't looking for a fight with champions of mainstream post-Frege logic!)

2.5 Allied Disciplines

In absorbing the dialogical approach to practical reasoning, we are free to engage — to appropriate or adapt — a large research literature. Dialogue logics come in a variety of stripes, some of the most interesting of which are Hamblin [1970], Lorenzen and Lorenz [1978], Barth and Krabbe [1982], Carlsen [1982], MacKenzie [1990], Walton and Krabbe [1995], Girle [1993], [1996], [1997], and Gabbay and Woods [2001] and [2001d]. A bounty of rich resources also arises from developments in cognitive science, AI and linguistics.

We take it as obvious that, irrespective of how we finally settle the question of the normative–descriptive distinction for theories of practical reasoning, it would be a mistake to ignore developments in these allied disciplines. For example, consider the impact of psychology. The psychological studies to date have concentrated on deductive, and probabilistic and inductive reasoning, with somewhat less attention given to decisional and causal reasoning. There is no simple dominant paradigm at present; in fact, there are at least four main approaches that are currently in contention. These are the *mental models* account (e.g., Johnson-Laird and Byrne [1991]), *mental logics* (e.g., Rips [1994]), *rational analysis and information gain* (e.g., Chater and Oaksford [1999], Oakford, Chater, Grainger and Larkin [1997]), and *domain specific reasoning schemas* (e.g., Evans and Over [1996]). Notwithstanding these theoretical and methodological differences, experimental evidence bears on the business of practical reasoning in two especially telling ways. One is that human beings do indeed seem disposed to commit fallacies, that is, errors of reasoning which are widely and cross-culturally made, easy to make and attractive, and difficult to correct. (Woods [1992]). A second point is that human reasoning performance seems to improve, that is, to commit fewer fallacies, when the reasoning in question is set in a deontic-context (Cheng and Holyoak [1985]). 'Deontic' here means directed to or productive of an action, which is the core sense of our notion of practicality. Since our *PLCS* is already moored in deontic and prudential contexts, a mature theory which is an extension of it must try to explain what is and what isn't a fallacy in a deontic environment or in a practical reasoning task, and why theoretical reasoning should be more prone to fallacies than practical reasoning. It is entirely possible that some of this difference lies in the fact that one and the same strategy might be a reasoning error in a non-practical context of reasoning, and yet be an error-free strategy deontically. (Gabbay and Woods [1999], [2004a].)

A practical logic should also incorporate important developments in the AI sector. It should exploit the fact that human reasoning is non-monotonic

and that non-monotonic structures have been investigated by AI researchers (e.g., Geffner [1992] and Pereira [2002]). Human reasoners are also adept at recognizing and manipulating defaults. A default is something taken as true provisionally or, as is said, in default of information to the contrary (Reiter [1980]). Default reasoning introduces into the business of human inference some extraordinary economies, which a practical logic must take pains with. For reasoning is good not only when it produces the right answer, but when it produces it on time. As a related development from linguistics, generic inference discloses its thinking to default reasoning. Generic claims are generalizations of a particularly elastic kind. Like 'Tigers are four-legged,' they tolerate true negative-instances (Carlson and Pelletier [1995]). They also seem triggered by very small samples, as we have seen. The two features are linked. Somehow human beings are rigged for what classically would be seen as hasty generalization fallacies in precisely these cases in which the reasoner is not generalizing to a universally quantified conditional (which is as brittle as a generic generalization is elastic), but rather to a generalization certain negative instances of which happen not to matter.

It is easy to see how default reasoning and generic inference touch on the classical fallacy of hasty generalization, and necessitate a substantial reconsideration of its traditional analysis. Other forms of default reasoning pertain in the same way to the classical fallacy *argumentum ad ignorantiam.* The basic structure of the fallacy is the (invalid) argument form:

1. It is not known that P
2. Therefore not P.

On the standard analysis, *ad ignoratiam* arguments are not only deductively invalid, but wholly implausible as well. But as studies of autoepistemic reasoning show (e.g.,) there are non-deductive exceptions to so harsh a verdict, as witness:

1. If there were a Department meeting today, I would know about it.
2. But in fact I know nothing of any such meeting.
3. So, it can reasonably be supposed that there'll be no meeting.

Here is further occasion for a mature theory of practical reasoning to winnow out the mistakes in classical accounts of fallacious reasoning (concerning which see Gabbay and Woods [2005]).

2.6 Psychologism

In our conception of a practical, agent-oriented, resource-based logic, we have not honoured every stricture against psychologism. Critics of, for example, the logic of discovery, those who think it a misbegotten enterprise as such, are drawn to the idea that accounts of how people entertain and select hypotheses, form and deploy conjectures, and more generally how they think things up, are a matter for psychology. Underlying this view is something like the following argument. Let **K** be a class of cognitive actions. Then if **K** possesses an etiology (i.e., a causal ancestry), this precludes the question of the performing or disperforming the **K**-action for good or bad reasons. If there *were* a logic of **K**-action it would be an enquiry into when **K**-actions are performed rationally, that is, for the right reasons. Hence there can be no logic of **K**.

Against this Donald Davidson is widely taken as having shown that far from reasons for actions precluding their having causes, reasons *are* causes, or more carefully, *having a reason* for an action is construable as a cause of it. ([Davidson, 1963]. See also [Pietroski, 2000] to the same effect.)[9]

We ourselves are inclined to emphasize a substantial body of work in reliabilist and other forms of causal epistemology. In its most basic form, a subject performs a cognitive action rationally when his performance of it was induced by causal mechanisms that are functioning reliably, that are functioning as they should.

We would do well, even so, to take brief note of a possible objection. If the aspects of cognition in which a logician could be expected to take an interest are often a matter of being in the right psychological state, and if such states are sometimes the output of causal mechanisms unattended by either attention or effort on the agent's part, how can this be squared with our view of logic as a principled description of (aspects of) what a logical agent *does*? Our answer is that just as we deny that there is an inherent incompatibility between reasons and causes, neither do we find any essential incompatibility between being in a causally induced mental state in whose attainment the agent played no intentional role and being the subject of admissible answers to questions such as 'What is X *doing*?' (answer: 'He is thinking that P'), and 'What was X *doing* that he came to be in state S?' (answer: 'He was looking at Harry's Corot print'). In a quite general way, whenever there is something that an agent is doing, there are constituent happenings, not all of which qualify to be described as what X is doing,

[9] Another approach to the reasons–causes issue is that of *agent causation*, skillfully developed in [O'Connor, 2001]. While we do not adopt this view here, we recognize it as an attractive alternative.

which might nevertheless enter into the description of what *does* qualify for the designation 'what X is doing'.

The idea of logic as a theory of rational performance runs into a different, though related, objection. The trouble with such a view of logic, it is said, is that it commits us to *psychologism*, and psychologism is false.

Anti-psychologism is not a single, stable thesis. It is at least three pairwise inequivalent propositions.

1. In one sense, it is the case made by the argument we have just reexamined and rejected.

2. In another sense, it is the view that although logic deals with the canons of right reasoning, no law of logic is contradicted by any psychological law or psychological fact.

3. In a third and more emphatic sense, it is the view that logic has nothing whatever to do with how people do reason or should.[10]

Having dealt with anti-psychologism in the first sense, it remains to say something about the other two. Sense number two need not detain us long. It is a view of anti-psychologism which is accepted by logicians who take a traditionally normative view of logic. On this view, psychology is purely descriptive, and logic is purely prescriptive. Hence the laws of logic remain true even in the face of massive misperformance on the ground. On the other hand, those who plump for reliabilist theories of rational performance will reject anti-psychologism in its present sense, just as they reject it in sense number one.

This leaves the third conception, the idea that logic has nothing to do, normatively or descriptively, with how human beings — or other kinds of cognitive agents, if any — think and reason. It is a view with an oddly old-fashioned ring to it, suggesting a position which simply has been overtaken by events of the past quarter century, referred to collectively by the

[10]It is interesting that the case which Frege actually pressed against psychological methods in logic are not transparently present in the trio of interpretations currently in review. In Frege [1884] and subsequent works, Frege's resistance was twofold, as was mentioned in the Preface of this book. First, if psychological methods were engaged in such a way as to make mathematics an experimental science, then those methods should be eschewed or anyhow not deployed in such ways. Second, if psychological methods were engaged in such a way that mathematics lost its intersubjective character, then psychological methods should be either abandoned or not employed in such ways. It bears on the present point that whereas Boole was a psychologicist about logic, and whereas Frege was a critic of Boole, Frege never criticized Boole for his psychologism. Logic for Boole is not a matter of how people *actually* think but rather is a normative account of the correct use of reasoning [Boole, 1854, pp. 4 and 32].

founding editor of the *Journal of Logic and Computation* as 'the new logic'. He writes:

> Let me conclude by explaining our perception of the meaning of the word 'Logic' in the title of this Journal. We do not mean 'Logic' as it is now. We mean 'Logic', as it will be, as a result of the interaction with computing. It covers the new stage of the evolution in logic. It is the new logic we are thinking of.
>
> [Gabbay, 1990]

Twelve years on, the editor's prediction has been met with considerable confirmation, and then some. The buds of the early 1980s have in numerous instances achieved full flower. Non-monotonic logics, default logics, labelled deductive systems, fibring logics, multidimensional, multimodal and substructural logics are now better established and methodologically more self-aware than they were even a decade ago. Intensive re-examinations of fragments of classical logic have produced fresh insights, including at times, decision procedures for and equivalency with non-classical systems. Perhaps the most impressive achievement of the new logic as arising in the past decade or so has been the effective negotiation of research partnerships with fallacy theory, the logic of natural language reasoning and argumentation theory.[11]

The new logic, the logic born of the application of the procedural sophistication of mathematical logic to the project of informal logic, has triggered the very rapprochement that mathematical logic was not structured to deliver or to seek. The new logic, whatever its multifarious differences of mission and detail, has sought for mathematically describable models of what human agents actually do in real-life situations when they cogitated, reflected, calculated and decided. Here was an approach that would in an essential way take what mathematical logic would see as inert context into the theory itself, where it would be directly engaged by the ensuing formalisms.

If psychologism is the view that logic has something to do with how beings like us think and reason, then we are psychologicists. But we are psychologicists of an ecumenical bent which counsels the theoretical rapprochement of logic more narrowly conceived with cognitive science and computer science. It is an approach to logic which leaves it an open research programme as to whether there might be a satisfactory logic of discovery.

In so saying, we do not place ourselves squarely in or squarely out of the ambit of our interpretations of psychologism (save the first). In particular,

[11] Attested to, for example, by the Netherlands Royal Academy Conference in Logic and Argumentation in 1995, and the two Bonn Conferences in Practical Reasoning in 1996 and 1997, and the De Morgan Conference on Logic, held in London annually since 1999.

we have not expressly declared ourselves on what might be called *Boole's question.* Is our approach one in which how people do reason is ignored in favour of how they should reason? Our answer at this stage is somewhat equivocal, but it is the best we can do for now: *we have doubts about the purported exclusiveness of this very distinction.*

2.6.1 Issues in Cognitive Science

The psychologism of our approach to logic places us in a nettle of contentious and unresolved issues in the philosophy of psychology and cognitive science. Exposure to these issues would be nothing if not tactically maladroit except for the various psychological indispensabilities to the laws of thought approach to logic. We do not have the wherewithal to settle the contentions that such a conception lands us in. But we would do well, even so, to try to situate ourselves in the midst of these entanglements. Like it or not, psychology, especially cognitive psychology, is a part of our project, and we meet with psychology as we find it, warts and all.

Cognitive science has taken on two principal tasks. One is to give a mentalistic description of the laws under which cognition occurs (and is largely successful). The other is to give an account of the mechanisms by which these laws function without drawing upon the lexicon of mental terms and expressions.

For the better part of a generation, it has been widely assumed by cognitive scientists that this latter account will prove to be a computational one. The still dominant view is that the cognizer's mind operates as a linear symbol processor, by which mental symbols are transformed by virtue of the syntactic character of those symbols. Against this, is the view that the practical agent is a parallel distributed processor, many whose operations are parallel rather than linearly connected, and non-symbolic or pre-linguistic. Their difference of opinion has yet to be resolved. We ourselves lean to a PDP approach if only because of its clear affinity to our fast and frugal conception of individual agency.

Either way, however, further assumptions are granted and further problems are met with. Whether on the standard computational or the PDP approach there is general agreement about the modularity of mind (see, e.g., Fodor [1975]) and disagreement as to whether the mind is comprehensively modular or whether central cognition (hypothesis formation, belief revision and the various other routines of practical reasoning) can be satisfactorily modelled in computationally symbol-processing terms. We see in this a natural concurrence between the modular and standard computational approaches. Part of the promise of PDP theories is that it disrupts this rough

equivalence and frees up the question of the modularity of central cognition from strictly symbolic assumptions.

Another matter on which virtually all are agreed is the importance of a distinction between automaticity and control in matters of cognitive attention. Here, too, there are disagreements. There are those who hold that automatic processing does not require attention, whereas central processing is effortful and subject to voluntary control (Schneider *et al.* [1984]). Others (e.g., Kahneman and Treisman [1984]) distinguish between early-selection (or filtering) models of attention and late-selective models, both of which appear to be automatic and yet the second of which requires attention. Bearing on this question is the further issue of at what stage does information processing take on a semantic character. A good many cognitive scientists are of the view that semantic processing and control go hand in hand, leaving no room for automatic-belief revision. But here too the evidence of semantic processing of information lodged in unattended channels. (See Treisman [1960] for the classic paper; also Treisman [1964], Corteen and Wood [1972] and von Wright *et al.* [1975]. For doubts see Dawson and Schell [1982] and Treisman *et al.* [1974].)

Among philosphers of mind, Fodor is perhaps best known for his insistence on a limitedly modular analysis of cognitive systems (Fodor [1975] and [1983]). Central cognition, he says, is holistic in design and operation, and, as such, slips entirely out of the ambit of cognitive psychology (see also Fodor [2000]). Fodor argues for the holism of central processing from the holism of science. Since holism requires comprehensive surveys of knowledge-bases (or belief-sets), and such surveys are computationally intractable, Fodor infers the computational intractability of central cognition if it *had* a requisitely computational structure. But central cognition actually occurs, so it cannot, he concludes, be computationally structured.

Our own view is that the holism of central cognition does not follow from the fact (if it is a fact) that science is holistic. There is room therefore for a non-holistic orientation in investigations of central cognition. Two such enquiries stand out. In the one, an attempt is made to link central cognition to local problem-solving heuristics that are cued automatically. In the other, evolutionary psychologists are drawn to modularist explanations on the basis of the highly structured complexity of the cognitive agent's brain. Since an entirely holistic central cognitive system, while highly complex, couldn't have anything like this same degree of structure, evolutionists conclude that it is more plausible to model the actual complexity of central cognition on the structured complexity of the cognizer's brain.

We find ourselves floating on the choppy seas of these interesting and interconnected disagreements. (These are nicely reviewed in Botterill and

Carruthers [1999].) If they have not yet been brought to successful resolution by psychologists, how much less the imperative of definitive pronouncement by logicians. Still, the practical logic of cognitive systems carries some expressly psychological assumptions, which are caught in the cross-hairs of these rivalries. To some extent, therefore, we find ourselves pitched on one or other side of these issues. Like any psychologically real account of cognition, the computational aspects must be made compatible with the plain fact of computational tractability (indeed of low-time, high pay-off set-ups quite generally). Both PDP and comprehensively modular approaches show promise here. A psychologically realistic account of cognition must also leave room for subconscious (and possibly pre-linguistic) and largely automatic cognitive operations. Here, too, the psychological literature on attention (e.g., Parasuraman and Davies [1984]) is, even though equivocal, helpful in setting the relevant parameters. If, for example, automatic processing is not always completely non-attentional, and yet if some even non-attentional processing can be said to have a semantic character, there is room for the idea that the avoidance of irrelevance is a centrally important component of cognitive success which is achieved automatically.

Consciousness is tied to a family of cognitively significant issues. This is reflected in the less than perfect concurrence among the following pairs of contrasts.

1. conscious *versus* unconscious processing
2. controlled *versus* automatic processing
3. attentive *versus* inattentive processing
4. voluntary *versus* involuntary processing
5. linguistic *versus* non-linguistic processing
6. semantic *versus* non-semantic processing
7. surface *versus* depth processing

What is striking about this septet of contrasts is not that they admit of large intersections on each side, but rather that their concurrence is approximate at best. For one thing, 'tasks are never wholly automatic or attentive, and are always accomplished by mixtures of automatic and attentive processes' [Shiffrin, 1997, p. 50]. For another, 'depth of processing does not provide a promising vehicle for distinguishing consciousness from unconsciousness (just as depth of processing should not be used as a criterial attribute for distinguishing automatic processes ...' [Shiffrin, 1997, p. 58]. Indeed '[s]ometimes parallel processing produces an advantage for automatic processing, but not always .... Thoughts high in consciousness

often seem serial, probably because they are associated with language, but at other times consciousness seems parallel ...' [Shiffrin, 1997, p. 62].

In what follows, these and other such matters will arise from time to time. If, when this happens, we judge ourselves to have something useful to say, we shall propose it. Otherwise we shall attempt to negotiate our way past.