# THE MANY USES OF 'BELIEF' IN AI

**Robert F. Hadley**

School of Computing Science, Simon Fraser University

## ABSTRACT

Within AI and the cognitively related disciplines, there exist a multiplicity of uses of 'belief'. On the face of it, these differing uses reflect differing views about the nature of an objective phenomenon called 'belief'. In this paper I distinguish six distinct ways in which 'belief' is used in AI. I shall argue that not all these uses reflect a difference of opinion about an objective feature of reality. Rather, in some cases, the differing uses reflect differing concerns with special AI applications. In other cases, however, genuine differences exist about the nature of what we pre-theoretically call belief. To an extent, the multiplicity of opinions about, and uses of 'belief', echoes the discrepant motivations of AI researchers. The relevance of this discussion for cognitive science arises from the fact that (a) many regard theoretical research within AI as a branch of cognitive science, and (b) even if theoretical AI is not cognitive science, trends within AI influence theories developed within cognitive science. It should be beneficial, therefore, to unravel the distinct uses and *motivations* surrounding 'belief', in order to discover which usages merely reflect differing pragmatic concerns, and which usages genuinely reflect divergent views about reality.

## INTRODUCTION

Within AI and the cognitively related disciplines, there exist a multiplicity of uses of 'belief'. On the face of it, these differing uses reflect differing views about the nature of an objective phenomenon called 'belief'. In this paper I distinguish six distinct ways in which 'belief' is used in AI. I shall argue that not all these uses reflect a difference of opinion about an objective feature of reality. Rather, in some cases, the differing uses reflect differing concerns with special AI applications. In other cases, however, genuine differences exist about the nature of what we pre-theoretically call belief. To an extent, the multiplicity of opinions about, and uses of 'belief', echoes the discrepant motivations of AI researchers, some of whom see themselves as simultaneously engaged in both AI and cognitive science, while others make no claims for the generality or cognitive validity of their results.

The relevance of our discussion to cognitive science is twofold. First, some of the theories (and senses) of belief described here are held by researchers who would identify themselves as cognitive scientists (whether or not they work in AI). It should be of use to these people to distinguish among several (though not necessarily all) of the current alternative views of belief. Secondly, research in AI sometimes influences the development of theories by cognitive scientists who do not regard AI as cognitive science. If AI contains conflicting theories of belief, and if (as is often the case) the motivation for these theories is a mixture of scientific and pragmatic concerns, it would be well for us to be aware of which theories are intended *primarily* as cognitive theories, and which are intended primarily as special purpose tools. Unfortunately, researchers in AI are frequently unclear about the degree to which they intend a particular theory or formalism to be taken as a cognitive model. This is especially apparent in the domain of "belief" (cf. Hadley, 1988). It is not unusual to find discussions of belief in AI which reject other treatments of 'belief' for their counter-intuitive features (e.g., logical omniscience) (cf. Levesque, 1984; Fagin and Halpern, 1985). However, when the counter-intuitive features of one's own theory are brought to light, the defense is often made that pragmatic value, rather than cognitive validity, is the issue. In what follows, I shall try to unravel some of the underlying motivations for the varying approaches to belief in AI, and to distinguish their pragmatic value from their cognitive validity.

## 1. THE SYNTACTIC THEORY

According to the 'syntactic' view of belief, beliefs are syntactic objects (sentences) which are explicitly stored in a special region of an agent's memory, often called a knowledge or belief base. Only those sentences which are explicitly stored are believed. Thus, from the mere fact that an agent believes 'Mary has a brother', it would not follow that the agent believes 'Mary has a male sibling'. In general, the syntactic approach assumes that *no* two distinct sentences express precisely the same belief.

Now, on the face of it, there is a circularity lurking in the above definition, because an appeal is made to the notion of a 'belief' base. One wonders whether the notion of a belief (or knowledge) base can be elucidated without invoking the very concept we are trying to analyse. Upon reflection, however, the circularity may be illusory. For one might attempt to distinguish a belief base from *other* regions of an agent's memory by noting that sentences which are stored in a *belief* base are taken by the agent as grounds for *action*. That is, when a sentence is stored in an agent's belief base, the agent is (usually) willing to act as though the sentence is true. If the sentence is too abstract to act upon directly, then the agent is at least willing to use the sentence as a premise in derivations which lead to action (unless the agent desires to *conceal* his/her beliefs).

If we accept the above suggestion -- that a sentence is in an agent's belief base if and only if the agent is willing (in the absence of concealment motives) to use the sentence as a premise when formulating its plans for action -- then we should accept a precondition of this suggestion, namely, that the agent is able to assign some semantic interpretation to the sentence. For an agent could not act upon the truth of the sentence unless he/she knew how to draw experientially meaningful consequences from the assumption of the sentence's truth, and this seems to presuppose at least *some* ability to interpret the terms occurring in that sentence.

Now, most AI treatments of belief do not confront the question whether the agent can interpret or "understand" (in any full-fledged sense) the sentences it "believes". No doubt this is due, at least in part, to the fact that builders of AI systems expect those systems to be used and *interpreted* by humans. That is, humans provide the semantics for the system. However, since we are here concerned with cognition generally, we cannot resort to an analysis of belief which is parasitic upon human cognition. Having said that, I must now sidestep the thorny question of how a system assigns a semantic interpretation to sentences it believes. The issue receives *some* further attention in our discussion of *intensions*, but for the most part, we must be content merely to acknowledge the existence of the problem.[1] In the remainder of our discussion, we shall assume that agents have the ability to semantically interpret sentences in their belief bases.

Apart from the difficulties just described, the "syntactic approach" has come under attack for a different, though related, reason. Many contend that the syntactic approach is hopelessly fine-grained, since it does not allow that any two syntactically distinct sentences are *necessarily* interchangeable in belief contexts. That is, the syntactic approach concedes that mutual interchangeability (or substitutivity) fails even in cases where we would intuitively judge those sentences to express the same "thought contents", proposition, or information. Now, many who reject the syntactic view would argue that one *could not* believe that Mary has a brother without believing that Mary has a male sibling. This is not to say that whoever believes 'Mary has a brother' could *describe* their belief using expressions like 'male sibling', but rather that the same belief can be *expressed* in words which may or may not be known to an agent who has the belief.

Among those who reject the purely syntactic approach to belief there is disagreement about just how fine-grained the criteria for 'same belief' should be. For example, some would insist that '*p or q*' and '*q or p*' are interchangeable in belief contexts, while others would hesitate to say that belief is closed under *any* logical transformation, however simple. Nevertheless, there are many who believe that belief is not *primarily* a relation between a sentence and an agent, but between an *abstraction* and the agent. The same abstraction is expressible by means of distinct sentences. In the following section we pursue this view of belief.

In defense of the syntactic view of belief, it should be noted that not all its advocates intend the view to be taken as a serious analysis of our ordinary concept of belief. Within AI, at least, the syntactic approach is sometimes adopted as an expedient which permits the construction of planning systems (cf. Haas, 1985, 1986). A robot who knows how to interpret and apply sentences to the world may derive useful plans based upon sentences explicitly written in its belief base. The fact that the *content* of these plans might be represented differently by other agents does not prevent the robot from using the plans it formulates.

---

[1]See, however (Hadley, 1989), and Searle's well known (1980) paper, which addresses these issues in detail.

However, the robot will be limited in its ability to communicate if it cannot conceive that another agent could represent those same plans using different words.

## 2. THE INTENSIONAL THEORY

As we have noted, a major drawback of the syntactic theory, when viewed as a serious *theory of belief*, is that it fails to account for the fact that *some* distinct sentences seem to express the same thought or belief. The notion of "thought contents" is one which has troubled philosophers for centuries. Beginning with Frege, however, substantial progress has been made in the elucidation of this concept. Nowadays, thought contents are commonly identified with intensions, propositions, and 'information expressed by a sentence'. I shall not attempt to unravel the distinctions among these notions in the space available here. Rather, I attempt to say what they have in common. This commonality, and its relationship to belief, is what I am (somewhat loosely) calling *the intensional theory.*

*Most* modern philosophers who have proposed theories of intensions, or propositions regard these as *abstract* objects which (roughly) constitute the *meanings* of sentences. Those who adopt Frege's basic approach towards these objects (e.g., Montague, 1970; Lewis, 1976) take intensions (propositions) to be *functions* which map sentences onto the possible worlds in which those sentences are true. Barwise and Perry (1983), however, take propositions to be abstract situation types whose structure is largely reflected in the syntactic structure of sentences which express those propositions. Both the Fregian-based and the situation-based view of propositions take propositions to arise as a result of *compositional* semantics. The composite, structured object expressed by a sentence arises (or is at least *picked out*) by a compositional processes involving *senses* or *sets* attached to the elementary terms occurring in the sentence. This composite, structured object (or, in the Barwise-Perry theory, a particular, situated *instance* of this object) is the appropriate object of belief. It is possible to *represent* this structured object by means of a canonical representation, and such representations are often called the *logical form of a sentence*. What is especially relevant to our present concerns is that all the propositional theories we have been considering associate sentences with such canonical logical forms, and all these theories recognize the existence of *many-to-one* mappings between symbolically distinct sentences and a given logical form. Moreover, all these theories admit the possibility, in principle, of "compilation procedures" by which a given sentence may be compiled into its logical form. It is thus possible, in principle, to ascertain whether two different sentences express the same belief by compiling the sentences into their logical forms. It is commonly recognized both by the neo-Fregian and by the situationists that such compilation processes must be sensitive to the context in which sentences are used.

Now, although the preceding discussion glosses over distinctions between the different theories which I have broadly depicted as "intensional", it has hopefully emerged that these distinct theories share a common motivation and sensitivity to the *structure* of a belief. The multiplicity of such theories partially reflects the difficulty of analysing the notion of 'thought contents'. However, the abundance of intensional theories also arises in consequence of the fact that (for the most part) proponents of these theories are attempting to describe the true nature of belief. That is, they are aiming at a cognitively accurate model rather than a special purpose construct. Within AI, the intensional stance towards belief has been adopted primarily by those concerned with cognitive fidelity and natural language (cf. Wilks & Ballim, 1987; Hadley, 1988).

An obstacle to the selection of a single 'correct' theory seems to be that there is no general consenses about the data to be explained. For example, Moore (1942) denies that the analysis of a concept is usually substitutable for an atomic term expressing that concept in a belief context. Thus, if we allow that 'male sibling' is an analysis of 'brother', Moore would deny that whoever believes that Mary has a brother also believes that Mary has a male sibling. On the other hand, Moore *could* allow that 'Tadpoles swim' and 'Polliwogs swim'. express the same belief, because 'tadpole' and 'polliwog' are not only synonymous, but are equally explicit.

Another (*prima facie*) difficulty for the 'intensional' approach arises from the fact that belief *sometimes* seems to be sensitive not only to the *structure* of a sentence, but to the identity of proper names occurring in the sentence. For example, some would argue that (a) 'Mark Twain wrote *Huckleberry Finn*' and (b) 'Sam Clemens wrote *Huckleberry Finn*' convey precisely the same proposition (or information), because the proper names 'Mark Twain' and 'Sam Clemens' denote the same individual, and so have the same

meaning (since they are *names* and not descriptions). However, it seems entirely possible that a child (say, Becky) who knew only a little about Mark Twain could believe (a) without believing (b). The syntactic theory, considered earlier, at least has the merit that it could assign a differing belief status to these two sentences.

In defense of the intensional approach, the reply could be made that the sentences in question have both a purely referential (*de re*) reading, and an opaque (*de dicto*) reading. Although 'Mark Twain' and 'Sam Clemens' both denote the same object, one could have different *concepts* or *vivid impressions* associated with these names. Thus, when Becky says 'Mark Twain wrote H. F ' she is normally *conceiving* of Mark Twain in a certain way (e.g., as a famous American author), and this *mode of conception* enters into the content of her proposition. For Becky, (a) and (b) do not convey the same *information*, and this is as it should be. A *context sensitive*, compositional semantics would *not* assign both sentences the same interpretation. To be sure, when 'Mark Twain' and 'Sam Clemens' are both being used in a purely referential mode, (a) and (b) can express the same proposition. In *that* case, however, we have no reason to suppose that Becky does not believe both (a) and (b). The fact that Becky would not assent to (b) is irrelevant when (b) is being used in a *de re* sense.

## 3. BELIEFS AS INFORMATION

In the previous section the term 'information' was used somewhat narrowly. This restricted use has been fostered by Barwise and Perry (1983), who use 'information' to denote what might also be described as a *structured state of affairs*. On their view, the structure of a state of affairs approximately mirrors the syntax of a sentence which describes that state of affairs. (Thus, information for Barwise and Perry roughly corresponds to *facts* for Wittgenstein, 1921). However, 'information' has other uses (cf. Dretske, 1981; Shannon & Weaver, 1949), and computer scientists are often concerned with the 'information' explicitly or implicitly present in a database. For many applications, one is not especially concerned about the *particular* logical structure of database information. Rather, the concern is with the 'picture of the world' that the information creates. That is, if one were to accept the information in the database as accurate, one would expect the world to be a certain way, independently of the particular syntax used to describe that world. On this use of 'information', one who knows that *if P then Q*, has the same information as one who knows that *if not Q then not P*.

Now, within AI, the practice has arisen of referring to declarative assertions in a database as 'beliefs' of the system. Given this casual use of 'belief', and given a concern for information in a broad sense, it is not surprising that we find AI researchers who regard beliefs as equivalence classes of logically equivalent sentences. We may formalize this use of 'belief' as follows:

Agent X believes S if and only if S is explicitly present in X's belief base, or S is logically equivalent to a sentence which is explicitly present.

Concerns for efficiency have lead some researchers to restrict the equivalence relation to "equivalent in a *computationally tractable* logic" (e.g., Levesque, 1984; Lakemeyer, 1987). However, such computational concerns are extraneous to the primary motivation we are now considering. If two agents have identical information encoded in syntactically different forms, that information will be identical whether or not the equivalence can be proved by a tractable algorithm. If we have chosen to identify the beliefs of a system with information in that system, then we should regard tractability as extraneous to belief as well.

Now, although the sense of 'belief' formulated above is motivated by the specialized concerns of AI researchers, it does accord with *some* of our ordinary uses of 'belief'. For example, if we have recently informed a friend that X and Y are both true, and we see that a sentence Z is an absolutely trivial consequence of X and Y, then we may reasonably conjecture that our friend will soon believe that Z. In cases such as these, we are not usually concerned with the particular *syntax* of the belief that Z, but with the information that our friend acquires. On other occasions however, a much finer-grained sense of belief seems to be at work. Thus, virtually any logic instructor will attest that a student may believe that $p \vee q$ and yet not believe that $\neg( \neg p \wedge \neg q )$.

## 4. THE LOGISTIC APPROACH TO BELIEF

Since the appearance of Levesque's "A Logic of Implicit and Explicit Belief" (1984), the prevailing approach within AI towards epistemic states has been to model belief by means of epistemic logics. Epistemic logics adopt the *modal* formalism developed by philosophical logicians who were dealing primarily with the concepts of possibility and necessity. However, they are also strongly influenced by Hintikka's (1962) application of modal logics to belief. Due to the influence of Levesque's (1984) results, together with his later arguments (1986), many have been persuaded not only that belief should be modelled via logic, but via *tractable* logics. In part this emphasis on tractability reflects a strong concern for the development of *practical* AI systems which can deliver results in feasible time (cf. Levesque, 1984; Lakemeyer, 1987). But concerns about tractability also arise for those who seek models of cognition which are at least equal in power to the cognitive abilities of humans.

We may summarize the stance towards belief currently adopted by many (though not all) AI researchers as follows:

Agent X believes sentence S if and only if S is explicitly present in X's belief base, or S is derivable, by means of a *tractable* epistemic logic, from a set of epistemic formulae corresponding to a subset of X's explicit belief base.

Now certain difficulties with the above emerge as soon as the thesis is explicitly stated. For example, the epistemic logics cited above do not address the fact that agents acquire beliefs over a period of time. Nor do they address the fact that an agent may, on occasion, validly derive a conclusion from prior beliefs, but *abandon* that conclusion because it conflicts with another of the agent's beliefs. To be sure, if the agent is rational, the conclusion will be abandoned only if the agent also discards at least one premise of the retracted conclusion. Nevertheless, agents often have inconsistent beliefs, and do not automatically "commit to" the conclusions they derive.

Now, the Levesque camp may object that there are many AI *applications* in which agents have only consistent beliefs, and never reject their own conclusions. But, this objection carries little weight. For, apart from the fact that the objection implicitly concedes that the "logistic approach" cannot provide a *general* account of belief, the objection does not begin to address the *temporal* problem. That is, current epistemic logics fail to distinguish between what an agent *now* believes, and what the agent could justifiably come to believe. Indeed, it appears that if the analysis of belief cited above is to have any plausibility, it must be construed not as an analysis of belief, but as an analysis of what an agent could *come to believe* by tractable means. This is brought home by the fact that all epistemic logics so far mentioned require agents to have an *infinite* set of (informationally non-equivalent) beliefs. For example, the logics of Levesque (1984) and Lakemeyer (1987) require that any agent who believes $p$ also believes $p \vee \neg(q \wedge r)$, and an infinity of other disjunctions. However, it seems implausible that an agent with finite resources could *at once* believe an infinity of non-equivalent propositions.

We are lead, therefore, to suppose that the motivation underlying the logistic approach is to provide logics which can characterize what an agent could justifiably come to believe by tractable means. However, the issue is complicated by the fact that the logicians cited above are greatly concerned to avoid a counter-intuitive aspect of Hintikka's epistemic logics, namely, logical omniscience. Logical omniscience is the thesis that an agent believes *all* the logical consequences of the agent's explicitly represented beliefs. Such a thesis is clearly false for finite agents, given our usual concept of belief. But, if we are liberalizing our interpretation of 'belief', and allowing that agents believe things which they are merely (logically) entitled to believe, then it is no longer clear that logical omniscience is an unacceptable doctrine. The fact that many epistemic logicians in AI find this doctrine unacceptable is puzzling, because their own logics demand a liberal interpretation of 'belief'. The situation is further complicated by the fact that creators of these logics often stress the 'intuitive aspects' of their logics.

Now, it may be that the deep reason why logical omniscience is rejected by many within AI, is not that the doctrine is counter-intuitive, but that it involves intractability. But if *tractability* is deemed essential, then it ought to be shown that under *no circumstances* may an agent arrive at new beliefs via intractable reasoning, and *no* beliefs should be stated in an intractable logic. However, it is far from clear that *purely artificial* agents can always meets their needs using only tractable logics. For example, researchers in automated planning, automated programming, and qualitative physics have not shown these domains to

be susceptible to tractable methods. Indeed, the consensus seems to be, that *if* formal logic is in fact an appropriate tool for these domains, then something at least as rich as first-order logic *with functions* will be required. And, if formal logic is *not* an appropriate tool for these domains, it is *a-fortiori* possible for agents to arrive at new beliefs (in these domains) without recourse to tractable logics. Moreover, it is known that proofs involving the principle of mathematical induction cannot even be formalized in a language as rich as first-order logic. Are we to say that those who discover theorems via mathematical induction are not acquiring *beliefs*, simply because their reasoning processes cannot be simulated by the application of tractable logics? Difficulties such as these could be dismissed by those concerned merely with restricted applications of AI, but as we have noted, the motivations of those advocating tractable logics in AI are often not clear.

## 5. BELIEFS AS 'WHAT ONE COULD RAPIDLY DISCOVER'

There is a use of 'belief', in ordinary parlance, according to which one believes not only those things which one *remembers* as true, but those things one could quickly discover. For example, if we ask a friend whether she believes that some carpenters are poets, she may pause a second, and then reply 'Yes'. In all likelihood, our friend has never considered the question, but because she can discover (or infer) the truth of a proposition so quickly, it seems not unnatural to say that she believes the proposition. Note, however, that if a considerable span of time is required for the inference, then we are reluctant to describe the proposition as something our friend believed at the time the question was posed. In certain AI applications, however, it may be reasonable to be flexible about the span of time involved. For in a given application, we may not be as much concerned with preserving the niceties of ordinary usage as we are with finding a convenient label for the things an artificial agent could discover in a given time T. Considerations such as these may have lead Fagin and Halpern (1985) to adopt a liberalization of the 'syntactic approach' described in section one.[2]

In Fagin's and Halpern's system (hereafter, the F & H approach), an agent is said to believe not only formulae explicitly present in the agent's belief base, but any formula which the agent is both 'aware of', and which is derivable from the belief base (by means of a specified logic). F & H allow 'aware of' to be interpreted in a number of ways, depending upon the application. One interpretation they suggest, however, is that an agent is aware of any formula which the agent could derive in time T. Using this extended notion of awareness, in combination with F & H's general approach, we are lead to the following characterization of belief:

Agent X believes S if and only if S is in X's belief base, or X could prove S, using the belief base and a logic L, within time T.

The above formulation permits us the option, in particular applications, of requiring L to be a tractable logic. However, we would remain more faithful to the motivation which underlies this extended use of 'believes' if we do not place a-priori restrictions upon the logic which the agent uses. For, the relevant question is whether the agent can indeed discover the truth of S in time T.[3] Moreover, we should note that the above formulation strays a fair distance from common usage, not only because it permits the interval T to be arbitrarily large, but because it counts as a belief any sentence which an agent *could* derive in time T (even though the agent may in fact never discover S because, for example, the agent is pursuing other lines of thought).

Now, it would be natural to assume that F & H's extended usage of 'belief' is intended merely as a technical convenience, and not as an analysis of 'real belief'. However, there are passages in their (1985) paper which suggest that they may be aiming at a true analysis of belief. For example, they provide 'ordinary language' examples to support their contention that awareness is a necessary component of

---

[2]I am indebted to Bill Demopolous for pointing this out.

[3]Concerning what an agent could derive in time T (where T is less than 10 seconds, say), it should be noted that there now exist surprisingly efficient, tableau-based theorem-provers for first-order logic, which can prove many difficult theorems in short order (cf. Oppacher & Suen, 1987).

belief. They then proceed to use 'awareness' in a number of senses which would be considered unusual, at least. (I have argued this in detail in Hadley, 1988). Moreover, they criticize Levesque (1984) for counter-intuitive aspects of his logic of belief, although they do not make clear which aspects of their theory they take to be intuitively appealing and which are intended merely as technical tools.

## 6. BELIEF AS DEGREES OF CONFIDENCE

We now consider a usage of 'belief' which is radically different in kind from the foregoing, and which could be viewed as a *qualification* upon each of the previous usages. I refer to Cheeseman's (1988) analysis of belief as a subjective, probability estimate of a proposition's truth. In part, Cheeseman is concerned to defend the use of probability as an AI tool for reasoning with uncertainties. However, he also suggests that we focus upon *degrees of belief* rather than upon an absolute belief/non-belief distinction. Now, there may be considerable merit in concentrating upon probability estimates when reasoning with uncertainties, but that is not our present concern.

However, it does seem true that we sometimes have degrees of belief. But, below a certain (difficult to identify) threshhold of confidence, it seems odd to describe our probability estimates as beliefs. For example, it seems strange to speak of believing a proposition which we regard as only 15% likely. To be sure, it makes sense to say that we *believe* that P is 15% likely, but here 'belief' applies to 'P is 15% likely', and not to P itself. Analogously, every assertion of the form 'P is *x*% likely' may itself become a candidate for belief. If we adopt Cheeseman's analysis of belief, then we must assign a probability estimate to 'P is 15% likely' before we can believe *this* assertion. And if we do assign a probability to this latter assertion, we can raise a similar question about whether this *last* probability estimate is believed. Clearly, if we are to avoid an infinite regress, we must eventually stop assigning probabilities to our judgements, and simply *accept* (or believe) the judgement made.

This is not to deny that we often work with probability judgements, or even to deny that we sometimes form meta-probability judgements. Rather, the point is that we must at some point, and by some mechanism, simply assign and record a probability. It does not matter how this *recording* is implemented, we may still (from a logical standpoint) regard the *record* as an entry made in the agent's belief base, which has the form 'P is *x*% likely'. Once this is recognized, we see that the suggestion that we work with probabilities does not eliminate the need for decisions about whether to accept any of the foregoing analyses of belief. For, we may still raise questions whether an agent believes only those sentences in its belief base, or must also believe equivalent sentences (or other entailed consequences). Thus, it seems plausible that a probabilistic approach to reasoning could be combined with each of the foregoing uses of belief.

## CONCLUSION

We have reviewed a number of distinct views and uses of belief. There are important interrelationships among certain of these views. For example, with the possible exception of the probabilistic view of belief, the syntactic approach occurs as an ingredient, in each of the other approaches. Moreover, the 'equivalent information', 'epistemic logic', and 'what could soon be discovered' approaches all involve the use of formal logic. Each of these approaches arises from a different motivation, and gives rise to a distinct use of 'belief'. It is not clear to what extent each of these approaches is commonly *taken* (by AI researchers) to be a *true* theory of belief, but they may be seen as compatible when taken merely as artifices for special AI applications. The 'intensional' approach, by contrast, seems to arise from a genuine concern to formulate a correct theory of belief. Its origins lie in philosophy, rather than AI, but it has found favor with some AI researchers. The probabilistic approach is distinctly a minority view within AI, but, as we have argued, it is possible to reinterpret this approach in such a way that it can be *applied* to each of the foregoing approaches. For example, one might hold that *propositions* are the appropriate *object* of belief, while insisting that in certain contexts it is pragmatically useful to believe propositions of the form 'P is 74% likely'.

In passing, it should be emphasized that the foregoing is not intended as an exhaustive list of views and uses of 'belief' in AI. Rather, my purpose has been to display at least a range of possible approaches to belief, and to show that such approaches *need not* be incompatible. Whether they are in fact incompatible depends upon their intended purpose.

## REFERENCES


Barwise, J. and Perry, J. (1983) *Situations and Attitudes*, Cambridge: Bradford Books.

Cheeseman, P. (1988) "An Inquiry Into Computer Understanding", *Computational Intelligence*, Vol. 4, pp.58-66.

Cresswell, M.J. (1985) *Structured Meanings: The Semantics of Propositional Attitudes*, MIT Press, Cambridge, MA.

Dretske, F.I. (1981) *Knowledge and the Flow of Information*, Bradford Books, Cambridge, Mass.

Fagin, R., and Halpern, J. (1985) "Belief, Awareness, and Limited Reasoning", Proceedings of IJCAI, Los Angelos, pp. 491-501.

Frege, G. (1952) "On Sense and Reference", in *Translations from the Philosophical Writings of Gottlob Frege*, Oxford.

Haas, A. (1985) "Possible Events, Actual Events, and Robots". *Computational Intelligence*, Vol. 1, pp. 59-70.

Haas, A. (1986) "A Syntactic Theory of Belief and Action", *Artificial Intelligence*, Vol. 28, 245-292.

Hadley, R.F. (1988) "Logical Omniscience, Semantics and Models of Belief", *Computational Intelligence*, Vol. 4, pp. 17-30.

Hadley, R.F. (1989, to appear) "A Default-Oriented Theory of Procedural Semantics", *Cognitive Science*, Vol. 13, pp. 107-137.

Hintikka, J. (1962) *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, Cornell University Press.

Lakemeyer, G. (1987) "Tractable Meta-Reasoning in Propositional Logics of Belief", Proceedings of IJCAI, Milan, pp. 401-408.

Levesque, H.J. (1984) "A Logic of Implicit and Explicit Belief", Proceedings of the American Association for Artificial Intelligence, Austin, TX., pp. 198-202.

Levesque, H.J. (1986) "Making Believers Out of Computers", *Artificial Intelligence*, Vol. 30, pp. 81-108.

Lewis, D. (1976) "General Semantics", in *Montague Grammar*, (ed.) Partee, B., Academic Press, New York.

Montague, R. (1970) "Universal Grammar", *Theoria* 36.

Moore, G.E. (1942) ""Reply to my critics", In *The philosophy of G.E. Moore*, (Ed.) P. Schilpp. Northwestern University Press, Evanston and Chicago.

Oppacher, F. & Suen, E. (1987) "HARP: a Tableau-Based Theorem Prover", *Journal of Automated Reasoning*, Vol. 3.

Searle, J. (1980) "Minds, Brains, and Programs", *The Behavioural and Brain Sciences*, 3, 111-169.

Wittgenstein, L. (1921) *Tractatus Logico-Philosophicus*, translation by Pears, D.F., and McGuinnes, B.F., reprinted (1963) Routledge & Kegan Paul, London.

Shannon, C.E., and Weaver, W. (1949) *The Mathematical Theory of Communication*, University of Illinois Press, Urbana.

Wilks, Y. & Ballim, A. (1987) "Multiple Agents and Heuristic Ascriptions of Belief", Proceedings of IJCAI, Milan, pp. 118-124.