Logique & Analyse **133-134** (1991), 41-55

ON LOGICAL OMNISCIENCE

Jacques DUBUCS

Introduction.

The status of belief in contemporary cognitive science is a widely disputed issue. On the one hand, cognitive science may be viewed as common-sense psychology under a new guise: most of the time, the explanations of intelligent behavior persist in giving a comfortable place to folk notions as *beliefs, desires, knowledge, plans* and the like. On the other hand, it is by no way clear in what sense such propositional attitudes may be attributed to *information-processing systems*: the categories of the folk-psychology cannot be applied in a computational setting without deep modifications. Epistemic logic (broadly conceived, i.e. encompassing doxastic logic) has been designed from the start as a formalization of the common-sense notions of knowledge and belief, fully overlooking the details of the actual way of the processing of information. The main symptom of this immaculate conception of knowledge and belief is the well-known *problem of logical omniscience*: knowers and believers are assumed to be impeccable logicians, able to draw all the logical consequences of what they know or believe. Such an epistemic logic provides, as it stands, a satisfactory formalization of the standard notion of *rational behavior*: logical omniscience fits well with a *normative* approach of knowledge and belief. But so far as our task is to describe realistic information-processing believers, logical omniscience is to be rejected, for computation is just the kind of activity that logically omniscient agents have no use for. So we have to build such a logic X that

$$\frac{\text{Epistemic Logic}}{\text{Folk Psychology}} = \frac{\text{X}}{\text{Cognitive Science}}$$

According to a widely shared opinion, X can be obtained so to speak by substracting logical omniscience from standard epistemic logic, that is by introducing in it some additional machinery intended to block such inferences as '$B_a p$; p implies; thus $B_a q$'. There is a number of such proposals. For example one can allow "impossible" possible worlds beside the standard ones, or define sieves ("awareness" operators) intended to restrict the corpus of the formulas mastered by the agents. But it can be argued that such local accomodations of the standard epistemic logic, whatever the heuristic value they may have, are only attempts to remedy the symptoms of a more serious disease. For so far as the modelling of actual cognizers is concerned, the possible worlds approach of epistemic logic suffers two major defects: its *ontology* is too abstract (actual agents have nothing to do with propositions), and its *taxonomy* is too coarse (agents with limited deductive power are crucially sensible to some differences between intensionnaly equivalent sentences). Thus what is nedded is both an ontology of particulars and a hyper-intensional taxonomy. I shall try to show that these requirements can only be satisfied by moving towards a proof-theoretical approach of cognition.

1. *Intentional states as normalizing behavior.*

Belief finds sense in both contexts of human action and discourse. Let us begin with action. Action is not explainable by beliefs alone, but by beliefs and desires together (or, in terms more familiar to the economists, by probabilities and utilities together). Both components are linked by the *pragmatic principle*: humans undertake just the actions that, according to their beliefs, will lead to the satisfaction of their desires. This principle is appositely expressible in the possible-worlds format.

1.1. *Possible-worlds formalization of the pragmatic principle.*

One's beliefs delineate a part of the set W of possible worlds, namely the part B containing the worlds compatible with the content of these beliefs. Similarly, one's desires delineate the subset D of the worlds in which they are satisfied. Now an action may be viewed as a transformation A of W: A(w) is intended to be the world that results from doing the action A in the world w. An action A may be termed *optimal* for an agent X iff it transforms each world compatible with X's beliefs into a world compatible with X's desires, i.e. iff

$$\text{(O)} \quad \forall w \in W\, [w \in B \rightarrow A(w) \in D]$$

The pragmatic principle affirms that human action is always optimal in this sense. The common sense *explanation* of behavior along this principle is thus at the same time a *normalization* of it: according to this view, the cognitive states (beliefs, desires) that are the reasons of an action are just the states with respect to which the action appears as optimal ([1]).

This kind of normalizing explanation does not require that one identifies, among the possible worlds, the "true" one, i.e. the world in which the action actually takes place. Nor does it require that one identifies among them the most appropriate world to the objective interest of the agent. It is a *solipsistic* explanation, that refers neither to the reality that the agent is prepared to transform, nor to the goals he would be well advised to pursue.

1.2. *Doxastic duals, beliefs as theoretical constructs.*

The aim of the normalizing explanation of behavior is to specify the cognitive state responsible for the actions of an agent. But there is no guarantee that an explanation guided by the pragmatic principle should converge towards a well-defined state. For the task of the theorist (given the actions, specify the states relatively to which these actions are optimal) is by no way of the same kind as the task of the agent he considers (given the states, find the optimal actions). The problem that the agent is confronted with is sometimes difficult but always well-defined. It consists, given two subsets B and D of W, in calculating an element A of W^W satisfying

(O') $A(B) \subset D$.

([1]) The current theory of economical behavior works basically along the same line. It supposes that the agents are guided by the principle of maximization of expected utility, which asserts that between several possible actions, one has to choose the action A for which the term

$$\sum_{w \in W} pr(w) .\ u(A(w))$$

takes the greatest value, where pr(w) and u(w) are the subjective probability and utility attached to the eventuality w. But this assumption is a simple generalization of the pragmatic principle to the case of partial beliefs and graded desires. For if beliefs and desires are perfectly categoric, i.e. if the range of values of pr and u is the pair {0,1} instead of the whole interval [0,1], then these functions may be considered as the respective characteristic functions of B and D, in such a way that the expected utility of A reduces to

$$\sum_{w \in B} u(A(w)),$$

term which attains its maximum when (O) is satisfied.

The theorist's problem (given A, specify two subsets B and D satisfying (O')) has generally no unique solution. Normalizing psychology, that has *two* degrees of freedom, is under-determined by the evidence of behavior: how could it distinguish between the individual who presses the brake pedal because his beliefs about driving are correct and he wants to brake, and his "cognitive dual" who does the same because he wants to accelerate and he has false beliefs about the role of the pedal ? Given that two contradictory beliefs may be held to be responsible (in conjunction with some appropriate desires) for the same overt behavior (2), the beliefs attributed to an agent by virtue of the pragmatic principle are not to be taken realistically, as genuine distinguishable states of an internal system causing behavior, but rather as theoretical constructs posited on mere instrumental grounds, for the intelligibility they procure. The possible worlds semantics provides a suitable account for these normalizing beliefs (N-beliefs).

1.3 *Effability of N-beliefs.*

N-beliefs are not supposed to be fully expressible in any language. Recall that they have been defined purely in terms of agenthood and choice, in such a way that entertaining a N-belief requires nothing more than the ability to distinguish between several eventualities and between several possible actions, and to evaluate the result (even erroneously apprehended) of each of them: creatures that cannot outwardly speak can however be N-believers. Therefore the relevant notion of possible world is not the language-dependant notion of a maximal consistent set of sentences, but a primitive (non-linguistic) notion of eventuality.

Now belief ascriptions, whether empirical or normalizing, are of course couched in the attributor's language, say L . Thus standard normalizing explanation of behavior specifies the cognitive state of the agent by means of the list $P_B(L)$ of the sentences of L that are true in B, together with the corresponding list $P_D(L)$ for D. For each sentence p in $P_B(L)$ (resp. $P_D(L)$), the agent is said to believe (resp. to desire) that p. In other words, the agent

(2) Let pr and u {0,1}-valued (degenerate) functions of probability and utility respectively, which define together the cognitive state (pr,u) of an agent. The action A is optimal w.r.t. the state (pr,u) iff pr $\leq$ u$\circ$A. Then there are many other states w.r.t. which A is optimal. E.g. A is optimal w.r.t. (1-pr, u') for any function u' satisfying u'$\circ$A $\leq$ 1-pr: each individual possesses a "doxastic dual" whose beliefs are exactly opposite to his, but whose (optimal) behaviors are indiscernible from his...

is said to believe that p iff $B \subset I(p)$, where $I(p) = \{w \in W; w \models p\}$ (the *proposition* that p). The whole corpus of the sentences "X believes that p", with $p \in P_B(L)$, is equivalent to the statement

$$B \subset \bigcap_{p \in P_B(L)} I(p).$$

This corpus provides an approximation *from above* of the set of the eventualities that are to be described: the *identity* of B and

$$\bigcap_{p \in P_B(L)} I(p)$$

cannot be guaranteed, for there is no reason for supposing that the theorist's language is rich enough to describe each relevant class of *possibilia* by a conjunction of sentences in L (straight cardinality arguments show that ineffability is even the rule (Lewis 1973, 90)).

Therefore the optimality notion actually used by the theorist, namely

$$\text{(O'')} \quad A\Big(\bigcap_{p \in P_B(L)} I(p)\Big) \subset \bigcap_{p \in P_B(L)} I(p),$$

is language-dependant, and as such deeply different from the absolute notion (O').

It is well-known that the necessity of representing N-beliefs in some background language is often responsible for apparent irrationality in behavior: the less rich the belief-ascription language, the less rational the behavior. For if L' extends L, then

$$B \subset \bigcap_{p \in P_B(L')} I(p) \subset \bigcap_{p \in P_B(L)} I(p):$$

the class of the doxastically accessible worlds is wider when described in L than when described in L' , and some actions then appear erroneously sub-optimal ([3]).

([3]) Anthropologists know that the very ground of most imputations of irrationality or "prelogicity" lies in the poverty of the theorist's language. As Lévi-Strauss (1969, Chap I) notices, you do not have the slightest chance of understanding the behavior of the Wik Muntan if you are not able to distinguish *Eucalyptus Papuana* from *Eucalyptus Tetrodonta*... Notice however that the poverty of the ascription language is not able to introduce *systematic* biases in the evaluation of behavioral rationality. For its impact on the description of agent's desires has the opposite effect: the more poor the language, the more numerous the worlds

1.4. *N-beliefs, L-beliefs, and logical omniscience.*

N-beliefs, when expressed in some ascription language L, have an interesting particularity: the corpus $P_B(L)$ - that contains any sentence p such that the agent believe that p - is deductively closed. In other words, $P_B(L)$

(i) contains any sentence deducible from p if it contains 'p'
(ii) contains any logical or mathematical truth
(iii) contains any sentence if it contains two contradictory sentences

This property of N-beliefs is by no way anomalous. For N-belief has merely to do with (subjectively) optimal behavior. And from this viewpoint it is not a sensible question whether hen's belief that eggs are to be kept at 25° on Celsius scale is or is not the same belief as its belief that they are to be kept as 77° on Fahrenheit scale: if ϕ and ψ are necessarily equivalent sentences in L, they describe the same set of *possibilia*, therefore N-belief that ϕ is *obviously* the same as N-belief that ψ.

Behavior is however not the only way towards belief. A thirsty man who believes that water slakes the thirst will drink that liquid if he comes upon it. But he will also *assert* that water slakes the thirst, if asked. Let us call linguistic beliefs (L-beliefs) the beliefs that are ascribed to an agent on grounds of his public assertions. Contrasting with N-beliefs, whose relata are unstructured propositions, L-beliefs have to do with sentences.

The *problem* of logical omniscience arises when L-beliefs are confronted with (reports of) N-beliefs. There is no possible *behavioral* evidence that you do not believe theorem: nobody can act as if 1729 was not the first integer descriptible in two different ways as a sum of two cubes. On the other hand there may be some *linguistic* evidence that you do not believe a theorem: there is very likely only one man (Ramanujan, according to Hardy (1940)) who has spontaneously asserted, in answer to an incidental question, the sentence "1729 is the first integer...". From which it may be concluded that Ramanujan knew more than others about number theory: logico-mathematical competence is testified by linguistic utterances, not by outer behavior.

N-believers are logically omniscient, and (ordinary) L-believers are not so. But some N-believers, namely those who are also speakers, are L-

compatible with desires, and the more successfull the actions !

believers too. What way out ? No one compelling recognition. We have a choice of strategies, that we have to consider systematically. In order to get them, let us try to suppose that an agent X does not *N-believe* some logical truth. Let us also suppose, to simplify, that the language mastered by the agent is just the ascription-language L . We obtain a contradiction, for the following assertions are not consistent together:

(1) There is such a sentence p of L that
(1a) $W \subset I(p)$ (p is true in any possible world)
(1b) X does not N-believe that p

(2) X N-believe that p iff $B \subset I(p)$ (definition of reported N-beliefs)

(3) $B \subset W$ (by definition of B)

For let p as in (1a). Then $W \subset I(p)$. Then $B \subset I(p)$, by (3). Then X N-believes that p, by (2). Contradiction with (1b).

Among (1), (2) and (3) some assertion is to be rejected. The choice of this assertion defines the space of all the possible solutions to the paradox of logical omniscience.

2. *The space of the solutions to the paradox.*

2.1. *Dismissing L-Beliefs.*

The paradox arises from a conflict between two kinds of belief-ascriptions: ascriptions grounded in overt behavior, and ascriptions grounded in linguistic utterances. But for familiar reasons the last ones may be considered as more controversible: the process of recovering the "true" cognitive state responsible for the utterance is afflicted by indeterminations of various kind, even when the believer and the ascriptor share the same public language. Thus linguistic evidence, which only testifies against logical omniscience of the agents, is then to be very cautiously considered. According to this strategy, the premise (1b) of the paradox is the intruder: it may be that the agents actually are perfect logicians, and that we have the contrary impression only because we do not understand what they say.

As a theorem, the sentence

(A) $12^3 + 1^3 = 1729$

is true in any possible world, thus in any world compatible with one's beliefs. It is therefore impossible to explain that X affirms

(B) $12^3 + 1^3 = 1728$

by saying that X believes that $12^3 + 1^3 = 1728$. The only explanation consists in supposing that in X's mouth '1728' refers to 1729, i.e. that the proposition that X expresses by means of (B) is just the same as the propositon that *we* express by means of (A). X does not ignore a necessary truth, but a contingent one, namely that we do not use '1728' in order to refer to 1729. So the logical competence of the individuals is preserved at the price of their linguistic competence. They appear as perfect logicians, as soon as we charitably renounce to give an homophonic translation of their utterances.

But this solution suffers a serious difficulty. If to ignore a mathematical truth is the same as to ignore a contingent truth relative to the use of symbolism in it, then mathematical truths themselves are to be equated with contingent truths about conventional matters. Among the premises of the paradox, what this strategy actually recuses is not (1b), but rather (1a), i.e. the existence of genuine necessary truths.

2.2. *Introducing impossible possible worlds.*

Fermat's "last theorem" F is yet undecided. It may be true, but it may be false too. If it is true, it cannot be false, and if it is false, it cannot be true. Suppose it true, and then necessarily true. The sentence "it may be true, but it may be false too" cannot mean that it *can* be true and that it *can* be false too. "it may be false" does not mean that there is some *really* (or *metaphysically*) possible world in which it is false, but that a world in which it would be false seems possible *to us*.([4]) Such a world is possible for a doxastic point of view, but impossible from a metaphysical one. Nothing

([4]) Of course we cannot preclude F (or some other undecided Π^0_1-statement) to be actually *undecidable* in the first-order system S we use for number theory, and we can be inclined, in such a case, to express this undecidability by saying that F is true in some worlds, and $\neg$F in some others. But we have to resiste to this temptation, for there is no real symmetry between F-worlds and $\neg$F-worlds: if F is undecidable in S, then F is certainly *true* (in the standard model), for if it were false, there would be such integers a, b, c, n that $a^{n+2} + b^{n+2} = c^{n+2}$, and this equality would be obviously provable in S...

surprising, if our cognitive abilities are not great enough to detect the metaphysical impossibility of some eventualities we envisage.

According to this argumentation, that considers the third premise ($B \subset W$) of the paradox as the intruder, epistemic modalities are to be slipped from metaphysical ones. Besides really possible worlds, we have then to consider also "impossible possible worlds" where some necessary truths are false, or where some absurdities are true. Those are just the worlds tolerated by the agents who are not logically omniscient. Logical incompetence is no longer reducible to linguistic incompetence: X may ignore that p for some necessary truth p without any misunderstandnis about the meaning of p.

In the usual epistemic logic, modelled on (normal) modal logic, belief has undesirable properties: closure under valid implication, closure under believed implication, belief of valid formulas, closure under conjunction. It has been shown (Levesque (1984), Fagin-Halpern (1988)) that the introduction of non-standard possible worlds suffices to give up these properties.

But a major difficulty arises with this strategy: so far as propositional attitude reports are concerned, compositional semantics can no longer assume the principle (MTC) of the identity of meaning with truth conditions (i.e. with truth values in any standard possible world). For suppose that X believes that p and does not believe that q for some pair (p,q) of provably equivalent sentences. 'p' and 'q' have same truth conditions. But not so for "X believes that p" and "X believes that q", that have then, by (MTC), different meanings. Thus, by compositionality, 'p' and 'q' have different meanings, contradicting (MTC). But (MTC) seems to be the only way to give a clear meaning to the logical constants, namely by associating with them simple applications from W into W (negation = complementation, conjunction = intersection, a.s.o.). To sum up, impossible possible worlds make logic unclear.

2.3. *Naturalizing N-Beliefs.*

Suppose that the reasons I have given for rejecting both strategies above are convincing. There only remains a last way out, and I want to suggest that it is the right one. It consists in recusing the second premise, according to which

(2) X N-believes that p iff $B \subset I(p)$.

Or rather in suggesting that N-beliefs, ascribed on grounds of behavioral evidence modulo some rationality assumption, are to be conceived along a different line in order to cope with linguistic evidence against logical omniscience. Or even that the very notion of N-belief is to be replaced, at least as far as cognitive science —opposite to simple folk-psychology— is concerned, by a more realistic notion of belief.

The main argument is the following. N-beliefs are intended to give (together with N-desires) reasons for action. But scientific explanation is causal explanation. Thus if beliefs have an explanatory role in cognitive science, the reasons they give for the action have also to *cause* it in some sense. Let us call such beliefs C-beliefs. In the definition of the truth-conditions for C-beliefs reports, namely

(C) 'X C-believes that p' iff Δ,

Δ has to be a sentence which, if true, refers to an *actual* fact able to be the support of causal relations with X's organism. But no causal transaction is even conceivable between organisms and such abstract entities as propositions or sets of possible worlds. Thus, in view of (2), N-beliefs cannot be C-beliefs. More generally: if the truth conditions for having some kind of belief are only definable in terms of possible worlds, this kind of belief cannot have any explanatory role in cognitive science. Epistemic logic for cognitive science cannot be possible worlds epistemic logic ([5]).

Propositional attitudes cannot be held as *direct* relations between agents and propositions, under penalty to lose any explanatory role in scientific psychology. But they may be "naturalized" if one considers them as *indirect*, or better as *composed* relations. For instance the "intentional" relation to which the sentence "X N-believes that p" refers may be decomposed as follows:

(i) An innocuous, unintentional relation between X and a *material* entity S (e.g. a *token* of symbol)
(ii) A relation of "representation" between S (as *type*) and the proposition expressed by 'p'.

The first relation may be clearly taken as C-belief: given that both relata

([5]) Whether it can be useful in cognitive metapsychology, in dealing with systematization of the theorist's discourse, is of course an entirely different point. Of course, it does.

are material, causal determinations may transit by its channel. The second one, that between representations and the objects they stand for, is of course admissible from a naturalistic point of view too.

Such a representationalist theory of belief (Field, 1978) perfectly solves the question how one may believe that p without believing that q, while p entails q. For if the relata of C-beliefs are opaque or "cognitively impenetrable" representations, the inference from "X believes that $x < 3$" to "x believes that $x < 4$" is by no way more justified than the inference from "'$x<3$' is written on this page" to "'$x<4$' is written on this page". But this solution to the logical omniscience puzzle has the disadvantage of being, in some sense, *too* convincing: one explains so effectively why the consecution of the beliefs cannot always reflect the inferential relationships of their contents, that one becomes unable to explain that this parallelism can however be *frequently* observed. The risk is now of underestimating the inferential capacities of the agents by pulverizing the cognitive architecture into infinitely numerous non relational states (the belief-that-p, the belief-that-q, a.s.o.) We have then to reintroduce some dynamic among cognitive states. And we also have to define some reasonable taxonomy of them: on what conditions two representations S and S' are cognitively equivalent ? If the above analysis is correct, cognitive equivalence is certainly *more* fine-grained than mere intensional equivalence (for instance it is by no way compulsory that the differences between A&A and A, or between A&B and B&A are devoid of any cognitive relevance). But the mere efficiency of brains and computers shows that cognitive equivalence is also certainly *less* fine-grained than mere type-identity between representations. I want to briefly examine in conclusion two proposals for dealing with these questions in a logical setting.

3. *Cognizers from a representationalist point of view.*

There are basically two ways in building an epistemic logic conforming to the representationalist requirement. The first one goes by the spirit of it, the second one by the letter. I want to suggest that letter is much better than spirit.

3.1. *The quotationnal approach of belief.*

The representationalist theory forbids any propositional account of belief:

any relation towards a proposition has to be mediated by a sentence-like representation. In the same vein, Quine, who dislikes possible worlds, suggests that any opaque context is to be treated as a quotational context. According to this suggestion, "X believes that..." is to be considered as a metalinguistic predicate attached to the *name* of a sentence rather than as a sentential operator. If we take Quine at his word, we shall translate "X believes that p" by "Bx⟨p⟩", where the quotation marks are Gödel's quotes. But if we simulate the current axioms of doxastic logic in this format, we obtain a contradiction, the so-called Believer's Paradox —provided that the axiom of arithmetic are adjoined to our construction (Kaplan-Montague (1960), Thomason, 1980).

3.2. *Cognitive dissonance: an exercise in token-based logic (sketch).*

Quotational approaches of belief simply identify "representational" with "non-propositional". But the litteral meaning of the representationalist requirements also involves the distinction between types and tokens of representations: C-beliefs only have to do with tokens. A promising way of dealing with such C-beliefs seems to be provided by recent work in token-based (or: substructural) logics. I want only to give here some indication on a possible application of these logics on the well-known problem of cognitive dissonance.

A well-known explanation of the limited deductive powers of human agents lies in the compartementalization of their beliefs. The people may have several belief systems, corresponding to different sides of their personality or to their different social roles. These "frames of mind", that we can conceive as consistent each on its own, are in general weakly related to another, and their union may even fall short of global consistency. For instance we may, as subjects deeply rooted in their Ur-Lebenswelt, believe that the sun rises in the morning, together as we may, as educated adults, believe that ptolemaïc astronomy is not correct. In such a case, we may be told together to believe that the sun rises in the morning, and to disbelieve that, but one is not however entitled to attribute to us, in the same vivid sense of belief, the belief that the sun rises and does not rise in the morning.

Cognitive dissonance has received much attention from possible worlds theorists (e.g. Stalnaker, 1987, chap. 5). But this framework, as above noticed, does not seem very suitable for dealing with agents considered as information processing systems, and the very question of the contextuality of inferential processes is arguably to be considered in another perspective.

To restate the question more formally, we have to distinguish between

(i) infer A&B from the justifiable presence of A and B in the same context

(ii) infer A&B from the justifiable presence of A in a context, and the justifiable presence of B in another context.

(The first inference is supposed to be performable, but not the second one). As the formal logic accounting for such a distinction has to be compositional, the above question is ill-defined. For two different definitions of what constitutes a justification of A&B in terms of what is to be considered as a justification of A, and of what is to be considered as a justification of B cannot cohabitate together. The only solution is to suppose that we have to do with two different conjunctions in (i) and in (ii): the first one is context-sensitive, the second one is context-free.

This distinction evokes two possible formulations for the right rule of conjunction in Gentzen's sequent calculus:

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma' \vdash \Delta', B}{\Gamma, \Gamma' \vdash \Delta, \Delta', A \& B}$$

(context-free conjunction)

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \& B}$$

(context-sensitive conjunction)

The contexts (side-formulas Γ and Δ) of the rule are merely juxta posed in the first version, but they are fusioned in the second one (here the rule can apply only when the contexts are the same on both sides). Both rules can easely be shown equivalent in presence of the following *structural* rules:

$$\frac{\Gamma, A, A, \Delta \vdash \Lambda}{\Gamma, A, \Delta \vdash \Lambda}$$

(left contraction)

$$\frac{\Gamma \vdash \Delta, A, A, \Lambda}{\Gamma \vdash \Delta, A, \Lambda}$$

(right contraction)

$$\frac{\Gamma, \Delta \vdash \Lambda}{\Gamma, A, \Delta \vdash \Lambda}$$

(left weakening)

$$\frac{\Gamma \vdash \Delta, \Lambda}{\Gamma \vdash \Delta, A, \Lambda}$$

(right weakening)

$$\frac{\Gamma, A, B, \Delta \vdash \Lambda}{\Gamma, B, A, \Delta \vdash \Lambda} \qquad \frac{\Gamma \vdash \Delta, A, B, \Lambda}{\Gamma \vdash \Delta, B, A, \Lambda}$$

(left exchange) (right exchange)

Thus the very possibility to model the context-sensitivity of the inferences rests on the (partial) removing of the structural rules: in order to prevent dissonant cognizers from admitting blatant inconsistencies, we have to restrict the scope of these rules. Far from being innocuous stipulations about combinatorial manipulations on formulas, the structural rules express the whole content of the over-idealizations we do in modelling the actual way of processing information. I want to conclude this outline by some remarks on the proof-theoretical approach of the limitations of the deductive competence.

1) In absence of the contraction rule, the antecedent and the succedent of a sequent can obviously no longer be viewed as *sets* of formulas. According to the standard interpretation, they are rather *lists* or "*bags*" of formulas. But one can also see them, conforming to the above philosophical requirements on C-beliefs, as sets of *tokens*: by considering the "details" of the proofs, we can make licit a *causal* interpretation of the process of inference.

2) It is worth noticing that such an approach of the logical omniscience problem is only "implicitely" epistemic: one attempts to represent the logical rules conforming to which the actual agents reason, rather than to express the formal laws of combination of the belief ascriptions.

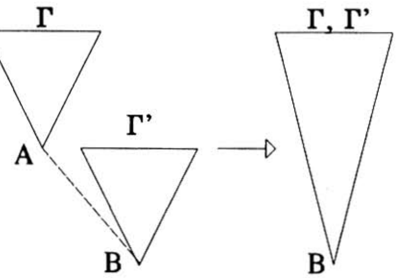
3) The (usual) cut rule

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}$$

is context-free, and it can be easely shown that it cannot be eliminated from a system containing only context-sensitive rules. There is here a noticeable difficulty, for only cut eliminability ensures the transitivity of inferences. E.g. the effect of the inference

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma' \vdash B}{\Gamma, \Gamma' \vdash B}$$

can be represented by the diagram



In a system of context-sensitive logic the only eliminable cut rule is

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta}$$

which is itself a contextual cut rule (as in a nomonotonic logic à la Gabbay), and which prevents any fusion of the inferential contexts: lack of transitivity in deduction is the price for cognitive dissonance.

Université de Paris I Sorbonne

REFERENCES

R. Fagin & J.Y. Halpern, Belief, Awareness, and Limited Reasoning, *Artificial Intelligence*, 34 (1988), 39-76.

H. Field, Mental Representation, *Erkenntnis*, XIII-1978, 9-61, repr. in N. Block (ed.), *Readings in the Philosophy of Psychology*, Harvard U.P., 1981, vol II, 78-114

G.H. Hardy, *A Mathematician Apology*, Cambridge U.P., 1940

H.J. Levesque, *A Logic of Implicit and Explicit Belief*, A.A.A.I.1984, 198-202

Cl. Lévi-Strauss, *La pensée sauvage*, Plon, 1969

D. Lewis, *Counterfactuals*, Harvard U.P., 1973

R.C. Stalnaker, *Inquiry*, M.I.T. Press, 1985