Rationality in human nonmonotonic inference

Rui Da Silva Neves, Jean-François Bonnefon and Eric Raufaste

Université de Toulouse-Le Mirail
5 allées Antonio Machado
31058 Toulouse Cedex, France
Email: {neves, bonnefon, raufaste}@univ-tlse2.fr

Abstract

This article tests human inference rationality when dealing with default rules. To study human rationality, psychologists currently use classical models of logic or probability theory as normative models for evaluating human ability to reason rationally. Our position is that this approach is convincing, but only manages to capture a specific case of inferential ability with little regard to conditions of everyday reasoning. We propose that the most general case to be considered is inference with imperfect knowledge – in the present case restricted to uncertain knowledge – and that a natural framework for testing the rationality of plausible reasoning is System P. This system provides rational postulates for nonmonotonic inference. The semantic of the nonmonotonic inference is given by a possibilistic constraint introduced by Dubois and Prade (1991). This constraint states that a rule $p \rightarrow q$ is a plausible rule if « the degree to which $p \wedge q$ is possible » is greater than « the degree to which $p \wedge \neg q$ is possible ». Given the choice of this constraint, we study two supplementary postulates of rationality. Eighty-eight subjects participated in an experiment whose results confirm – provided that reflexivity and left logical equivalence would be tested in a further experiment – the rationality of human nonmonotonic inference according to the rational postulates of System P.

Introduction

Uncertainty is a constitutive aspect of human cognition. Several consequences derive from this: human knowledge is inconsistent to a certain degree; human inference is nonmonotonic; and some previously accepted beliefs must in some occasions be revised in order to adapt our behavior to an environment which is dynamic by nature. This article deals with nonmonotonicity in human logical inference as a consequence of the use of default rules. Nonmonotonicity is obvious when we are reasoning about default rules as in the following example. Consider the following plausible rule: “If one does not brush one’s teeth every day, then usually one gets severe cavities”, and the fact: “Eva does not brush her teeth every day”. From these premises, most people conclude: “Eva suffers from severe cavities”. But given the additional fact: “the water drunk everyday by Eva contains much chlorine”, most people informed that chlorine has a positive effect on tooth health will continue to believe that Eva rarely brushes her teeth, but will judge possible that Eva might not have severe cavities. This conclusion seems rational, but it goes against the rules of classical logic. Following these rules one would continue to believe that Eva has severe cavities, that is, one would reason monotonically. In spite of the obvious frequency of this sort of nonmonotonic inference in daily life, psychologists have devoted little attention to nonmonotonic human reasoning and inference.

Since the beginning of the 2Oth century, psychologists have been concerned with human reasoning and rationality. They have been developing a systematic research program in order to test the agreement between human reasoning and classical logic on the one hand and classical theory of probability on the other hand. Classical logic has furnished the norms for assessing human deductive ability. This project mainly consisted in submitting various logical arguments (with connectives, categorical premises and relational premises) to human subjects and asking them (i) to judge the validity of some valid and invalid conclusions, (ii) to select the correct conclusion in a pool of given valid and invalid conclusions, or (iii) to generate the correct conclusion. Classical theory of probability has furnished for its part the norms for assessing human ability to reason with uncertain knowledge. The very important mass of accumulated empirical data exhibited that people succeed in some classes of problems and fail in others (Evans 1982, 1990; Kahneman, Slovic and Tversky 1982). Three positions have been currently adopted given the systematic deviations of human reasoning from normative models. These positions are expressed relatedly to the question of human rationality.

The concept of rationality has not yet received an uncontroversial definition. In psychology as in social sciences, two main senses have been attributed to rationality. The first one is that humans are rational if in deciding what to do they engage in a logically correct reasoning. The second is that humans are rational if their behavior is optimal in achieving their goals.

The first interpretation is closed to Newel’s (1982) view. Indeed, Newel proposed that the laws of human behavior depend on a “principle of rationality”. This principle stated that “if an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select that action” (p. 102). The rational level in cognition proposed by Newell supposes that the human knowledge is logically related to human goals. The second interpretation is closed to the position in economics and has been defended by Anderson (1990) in the field of cognitive psychology. According to Anderson "the cognitive system operates at all times to optimize the adaptation of the behavior of the organism" (p. 28). In this sense, rationality is a matter of human's behavior adaptation to environment (adaptive rationality), not a matter of logical relation between knowledge and goals (normative rationality). Moreover, in the Anderson's view, human's rationality depends on person's experience rather on human knowledge. As a consequence, we can expect normative rationality with people who have a direct experience of normative rules given some normative model. Furthermore, people not familiar with such normative models would be rational under the conditions where adaptive rationality fit normative rationality. What are these conditions is an open question.

Psychological conclusions about human rationality in reasoning have been mainly formulated according to the normative sense considered by Newell.

A first position adopted given the systematic deviations of human reasoning from normative models is that people is fundamentally irrational when reasoning. This position is classically attributed to Tversky and Kahneman (1974). These authors defend the idea that human reasoning is driven by heuristics that produce suitable conclusions in certain contexts but lead to recurrent biases in others. Such a position has been defended in philosophy by Stitch (1985).

A radically opposite position have been defended by the philosopher L. Jonathan Cohen (1981). In the Cohen's line, psychological experiments cannot be used to demonstrate human irrationality. An important argument advanced by Cohen (see Stich 1985 and 1988 for a discussion of the main arguments provided by Cohen) concerns the competence/performance distinction introduced by Chomsky (1965) in the context of linguistic theory. Chomsky described the structure of language as a formal system of rules sufficient for allowing every well-formed grammatical utterance. The competence in the Chomskian sense is a linguistic notion, not a psychological one. However, one often considers that human language competence is governed by a set of internalized rules that form a mentally represented grammar which interacts with others systems like the attentional system, the motivational system, mnemonic systems… These interactions override the competence system and explain why utterances are sometimes grammatically incorrect. By analogy, Cohen supposes that human reasoning depends on a system of internalized rules of reasoning that forms human competence and grounds human rationality. Irrationality cannot be experimentally demonstrated because experimental data capture only the performance level. Pollock (1987) has adopted a similar point of view in proposing a human rational architecture in order to introduce a theory of warrants based on defeasible reasons. In Pollock's view, reasoning is also guided by internalized rules that constitute procedural knowledge and form a production system. This system is a competence system of the kind supposed by Cohen. The production rules are the rules for "correct reasoning", and "what we "should" do is what the rules of our production system tell us to do, but we do not always conform to those rules because they are embedded in a larger system that can override them" (p. 483). A second important argument advanced by Cohen (1981) is that reports of subject' errors depend on the normative system adopted by the experimenter. Cohen has pointed out that subjects can reason rationally according to alternative systems of logic or probability. This argument has been the focus of recent works in cognitive psychology. Indeed, new accounts of human deductive performance have proposed that one reason for which human reasoning diverges from classical logic is that classical logic fails to capture some basic properties of human knowledge, human inference and human cognitive limitations (Oaksford and Chater 1994, 1995; Chater, Redington, Kakisa, and Oaksford 1999; for similar arguments against classical probabilistic models of reasoning under uncertainty, see Gigerenzer and Goldstein 1996 and Todd and Gigerenzer 1999).

A third position adopted by psychologists given subject failures and successes in reasoning tasks, is that people have a bounded rationality in the line of Simon's proposal (1955). That is, reasoning ability would be bounded by the computational cost of some reasoning problems. For example, playing – under normal time pressure - a perfect game in chess is intractable for the human mind (as for any known artificial system). So, people would be rational when dealing with tasks that does not override their mnemonic and computational capacities, and would only "tend" to rationality in other cases. This limitations have for consequence that "people are rational in principle, but fallible in practice" (Johnson-Laird and Byrne 1991, p. 19). Similar considerations have led Gigerenzer and Goldstein (1996) to assess that classical probability theory should be rejected as the norms of human reasoning under uncertainty because their intractability in a human cognitive device. Instead, for Gigerenzer and Goldstein, psychologists should search for "fast and frugal algorithms" in the manner of Simon's satisficing model (1955).

The position adopted in this paper is consistent with the view that classical logic and classical probability theory do not capture the essential properties of human interaction with the environment. In that way, a species fitted solely with these two theories would surely be doomed to quickly disappear. So, we agree with the idea that alternative norms are necessary to evaluate human rationality, and instead of searching for particular algorithms that would describe and allow to measure human performance in nonmonotonic reasoning, we adopt – along with pioneering psychologists - the view that formal mathematical models must provide criteria to assess human rationality. However, the underlying concepts and assumptions of these models must in return exhibit some psychological plausibility.

For two decades, in the field of artificial intelligence, an important effort has been devoted to the production of rational models of nonmonotonic reasoning. Within this line of research, a coherent set of properties known as "System P" has been proposed by Kraus, Lehman and Magidor (1990) and has been recognized as the minimal set of rational postulates shared by formal models of nonmonotonic inference (Benferhat, Dubois and Prade 1997). Because of its status in the AI community, System P appears as a natural framework for testing of human rationality when reasoning with default rules. However, other properties have been proposed that extend System P in a controversial manner. The a priori benefit from the empirical test of the rationality of human nonmonotonic inference can also concern AI in the sense that mathematical intuitions could gain or lose support in the light of our empirical data.

In the next section, we will describe the general framework adopted in our empirical study. The subsequent section will present the experimental apparatus and our main results.

Formal apparatus and previous empirical findings

The properties of rational inference

Nonmonotonic reasoning must satisfy certain properties in order to be logically sound, that is, in order for a rational agent to infer reliable and relevant knowledge or beliefs from some uncertain, imprecise and/or incomplete information. Kraus, Lehmann and Magidor (1990) have proposed six basic properties that form the core of System P and that are all satisfied by a preferential consequence relationship introduced by these authors. This preferential consequence relationship states that α preferentially entails β if one prefers the conclusion β to non β in the context α. For the presentation of these properties, we adopt the following notational conventions. A plausible rule "if α then preferentially β" is noted $\alpha \rightarrow \beta$. The logical equivalence is noted $\alpha \Leftrightarrow \beta$, and the material implication is noted $\alpha \Rightarrow \beta$. Other classical propositional connectives will be denoted by $\neg$, $\wedge$, and $\vee$.

Reflexivity: $\alpha \rightarrow \alpha$

This property has essentially been introduced for the sake of consistency with classical logic. It expresses that in the context α, we prefer the conclusion α to $\neg\alpha$.

Left logical equivalence (LLE): $>\alpha \Leftrightarrow \beta \,;\, \alpha \rightarrow \gamma$

$\beta \rightarrow \gamma$

This property expresses that if α is logically equivalent to β (and this equivalence is a logical consequence in the sense of classical logic; noted >), then, the plausible consequences γ of α are also plausible consequences of β.

Right weakening (RW): $> \alpha \Rightarrow \beta \,;\, \gamma \rightarrow \alpha$

$\gamma \rightarrow \beta$

If we accept α as a plausible consequence of γ, then the classical consequences of α must be accepted as plausible.

Cut : $\alpha \wedge \beta \rightarrow \gamma \,;\, \alpha \rightarrow \beta$

$\alpha \rightarrow \gamma$

If we accept γ as a plausible consequence of α and β, then γ is a consequence of α without β if β is itself a plausible consequence of α. This property is exhibited when the normal cases of α are sufficient to observe γ, and consequently when β carry little more information than α.

Cautious monotony (CM): $\alpha \rightarrow \beta \,;\, \alpha \rightarrow \gamma$

$\alpha \wedge \beta \rightarrow \gamma$

This property allows to conclude γ in the context where α and β are considered if γ and β are known as plausible consequences of α, that is, models of α are normally associated with models of β.

Or: $\alpha \rightarrow \gamma \,;\, \beta \rightarrow \gamma$

$\alpha \vee \beta \rightarrow \gamma$

This property means that when γ is a preferential consequence of both α and β, γ is a preferential consequence of the disjunction of α and β.

These six properties form the System P. Another interesting property that can be derived from the previous one is the "And" property.

And: $\alpha \rightarrow \beta \,;\, \alpha \rightarrow \gamma$

$\alpha \rightarrow \beta \wedge \gamma$

This property means that when α is known on the one hand to preferentially entail β, and on the other hand to preferentially entail γ, then α can be said to preferentially entail β and γ together.

Analyses of the inference defined by these properties show that an inference equipped with these properties is too cautious, because it does not always produce the desirable conclusions. Kraus and his colleagues have consequently proposed three others properties. Of particular interest is the Rational Monotony (first introduce by Makinson 1989) that states that:

Rational Monotony (RM): $\alpha \rightarrow \gamma \,;\, \neg(\alpha \rightarrow \neg\beta)$

$\alpha \wedge \beta \rightarrow \gamma$

This property means that if we accept the consequences γ of a premise α, and if this premise does not have for consequence $\neg\beta$, then the consequences of α are not retracted when β is added to α.

The possibilistic inference

When Kraus, Lehmann and Magidor wrote their seminal paper presenting System P in the early 1990's, few formal models appeared rational under the properties of System P. The models cited by these authors were Delgrande's first-order conditional logic and Pearl and Geffner's probabilistic semantics based on Adams's logic of conditionals (1975). Some other models have been proposed in the last decade of the century. For example, Pearl (1990) have proposed the System Z in a probabilistic framework; Benferhat, Dubois and Prade (1992) have proposed a possibilistic inference relation; Dubois and Prade (1994) have proposed the logic of conditional objects… (see Benferhat, Dubois and Prade 1997). To our knowledge, despite the obvious interest of all of these models for an understanding of human rational inference in plausible reasoning, only the possibilistic logic first introduced by Dubois and Prade (1987) has been studied from a psychological point of view and has exhibited psychological plausibility (see Raufaste and Da Silva Neves 1998).

In possibility theory, a conditioning notion closed to Bayesian conditioning has been defined by Dubois and Prade (1987) which states that the degree to which one can be certain to observe β in the context α is totally determined by the degree to which it is possible to observe α and β on the one hand, and the degree to which it is possible to observe α and $\neg\beta$ on the other hand. Formally, β is said a possibilistic consequence of α when $\Pi(\alpha \wedge\beta)>\Pi(\alpha \wedge\neg\beta)$. The degree of conditional necessity of α knowing β $(N(\beta/\alpha))$ is positive iff this constraint is satisfied. Dubois and Prade (1991) and Benferhat, Dubois and Prade (1992) have shown that this consequence relation satisfies LLE, RW, CUT, CM, OR, RM, and AND, but does not satisfy reflexivity (intead, they introduce a restricted definition of reflexivity - the Nihil ex Absurdo property - that states that nothing can be deduced from a contradiction; this property is not studied in the present preliminary research).

Empirical findings

To our knowledge, no empirical study devoted to the test of the rational postulates presented before has been published in the cognitive science literature, and very few studies have focused explicitly on the study of nonmonotonic reasoning or inference. Results about nonmonotonic inference, can be divided into two sets of studies: (i) studies of human reasoning from uncertain knowledge, and (ii) studies of factors that affect nonmonotonic reasoning.

With regard to the first set of studies, an indirect way to evaluate the strength of the relation embedded in conditional statements is to consider endorsement rates (i.e., the percentages of subjects who accepted the conclusion of a given argument) of the valid conditional syllogisms. George (1995), Politzer and George (1992), and Stevenson and Over (1995) have shown a statistically correlated transmission of beliefs from premises to conclusions of conditional arguments (Modus Ponens and Modus Tollens). In addition, Politzer and George (1992), have shown that this transmission depends on the conformity of the presented scenario to their normality in the real world, and that the suggestion of counterexample situations to the conditional rule decreased the Modus Ponens endorsement rate. Byrne (1989) and Cummins, Lubart, Alksnis and Rist (1991) found results that support the idea that the perceived certainty of a conditional rule varies as a function of factors related to the situations that falsify it (their number, their accessibility or their relevance in context).

A second set of results refers explicitly to belief revision and nonmonotonic reasoning. Works on Bayesian inference have shown that subjects are conservative (Edwards 1968; Slovic, Fischhoff and Lichtenstein 1977), that is, prior probabilities are generally overweighted compared to new information. Other results have exhibited that people often neglect the a priori probability of events (Kahneman and Tversky 1973). The "base-rate neglect" can be avoided with causal interpretations of conditional probabilities (Ajzen 1977). In a logical approach to belief revision, Elio and Pelletier (1994) have shown that people generally accept to revise in the sense of falsity an hypothesis initially believed as true, but they tend to refuse the contrary. Nonmonotonic reasoning appears to be very sensitive to contextual and content factors (Byrne 1989; Elio 1997; Elio and Pelletier 1994, 1997). For example, Elio and Pelletier found that human default reasoning is influenced by the specificity of the given information, its similarity, the size of considered categories… However, according to these authors, nonmonotonic reasoning by human agents depends not only on contexts and contents but also on syntactic cues.

Objectives and predictions

The results presented before are consistent - in a human adaptive perspective - with the permanent need for a behavioral adaptation to new information. But in this same perspective, behavioral regularities are also necessary conditions for rational behavior. Our main objective in the following experimental study is to test human inference rationality from the rational postulates described in the previous section. In addition, our results would permit to gain insight into the relevance of these postulates from an AI perspective. Indeed, if the postulates of System P are generally accepted by the AI community, the two other ones are only partially accepted.

Our predictions are based on the following rationale. If subjects' inferences from default rules are rational, these inferences must at least satisfy System P properties. Each property is viewed like an inference rule in Lehmann and Magidor' fashion (1992). In the general case, a property is judged as satisfied if there is a significant difference between (1) the frequency with which the default rules serving as its premises and conclusion are judged plausible and (2) the frequency with which the default rules serving as its premises are judged plausible but not the default rule serving as its conclusion. Put in other words, we want to know if the acceptation of the conjunction of the premises is preferentially associated with the acceptation of the conclusion rather than with the non-acceptation of the conclusion. A default rule of the form $\alpha \rightarrow \beta$ is judged plausible if subject judgments are such that $\Pi(\alpha\wedge\beta) > \Pi(\alpha\wedge\neg\beta)$. In order to test RW, we conclude that $\alpha \Rightarrow \beta$ if $\alpha \rightarrow \beta$ and $\Pi(\alpha\wedge\neg\beta) = 0$. In order to test LLE, we conclude that $\alpha \Leftrightarrow \beta$ if $\alpha \Rightarrow \beta$ and if $\beta \Rightarrow \alpha$.

Experiment

A pre-experiment involving thirty two French subjects – students in psychology and native speakers - was made with the objective to determine a set of rules endorsed by our population as plausible rules. For each rule $\alpha \rightarrow \beta$, (e.g. "no minor has the right to vote"), subjects were asked to evaluate $\Pi(\alpha\wedge\beta)$ and $\Pi(\alpha\wedge\neg\beta)$ (that is the possibilities that "an individual is a minor and has the right to vote" and "an individual is a minor and has no right to vote").

Subjects

Eighty eight first-year psychology students - all native speakers in French - in a first course of cognitive psychology at the University of Toulouse-Le Mirail took part in this study. None of them had received any logical training.

Material

The material consisted of forty six questions involving an unknown character designed by a surname (e.g. Mathilde B.; Simon A. …). Subjects were asked to imagine that these characters were randomly selected from a phone book and had to respond to questions of the form: "To which degree do you judge possible that Simon A. is a vegetarian and enjoys bullfights ?" and "To which degree do you judge possible that Simon A. is a vegetarian and does not enjoy bullfights ?". For example, in order to test Right Weakening, subjects are asked, on the one hand: "To which degree do you judge possible that Christophe C. is entering high school and is minor?", "To which degree do you judge possible that Christophe C. is entering high school and is not minor?"; and on the other hand: "To which degree do you judge possible that Christophe C. is minor and has the right to vote?", "To which degree do you judge possible that Christophe C. is minor and does not have the right to vote?"; subsequently subjects are asked "To which degree do you judge possible that Christophe C. is entering high school and has the right to vote?" and "To which degree do you judge possible that Christophe C. is entering high school and does not have the right to vote?". (The material involved in the test of the other postulates is available in appendices). Subjects were invited to express their judgements of possibility on an axis with no graduation.

The studied postulates of rationality are the six properties of System P except Reflexivity. In addition, we have studied the Rational Monotony and the And properties.

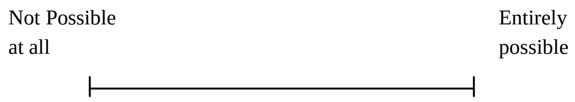

Figure 1: Axis for the subjects' responses.

Procedure

The experiment was administrated to eighty eight subjects during an introductory course of cognitive psychology. The questions were presented in a randomly selected order and in reverse order. Subjects were informed that in responding, they could (i) check the point of the axis that best matched their judgement, or (ii) draw an ellipsoid around the portion of the axis that contained their answer, if they had only an imprecise idea of the localisation of this answer.

Experimental results

The experimenter first applied an eleven-graduation grid to the response axis and then encoded the subject's response as an interval [Πinf, Πsup]. If the answer was an ellipsoid, the lower and upper values of the interval were given by the graduations that were the closest to the intersection point between the ellipsoid and the axis. In the case where an intersection point was equidistant from two graduations, the retained value was the lowest one if the lowest part of the interval was concerned, and the upper one in the other case. If the answer was a single check, the closest graduation provided both the lower and upper values of the interval. If the mark was equidistant from two graduations, the left one gave the lower value and the right one gave the upper one. In order to test the postulates of rationality, for each postulate, we constructed a contingency table on the model of Table 1.

		Conclusion acceptation	
		Yes	No
Premises acceptation	Yes	A	B
	No	C	D

Table 1: Model of the contingency tables constructed for the evaluation of the degree of association between the premises' acceptability (yes/no) and the conclusion's acceptability (yes/no) for each postulate. The cells A, B, C and D represent the number of subjects that accept or not the premises and the conclusion.

We wanted to know whether the acceptance of the conjunction of the premises was preferentially associated with the acceptance of the conclusion rather than the non-acceptance of the conclusion. The measure of the degree of association between the premises' conjunction and the conclusion was computed by the Phi (ϕ) coefficient (see Siegel and Castellan 1988). For each property, the statistical hypothesis H0 states that there is no significant association between the premises and the conclusion. H0 is rejected in favour of H1 (there is a significant association) if the probability of obtaining a value as large as the observed Khi2 is equal to or less than .05 as usually done in Psychology. Consider the Cautious monotony (CM) example:

$$\alpha \rightarrow \beta \; ; \; \alpha \rightarrow \gamma$$
$$\text{------------------}$$
$$\alpha \wedge \beta \rightarrow \gamma$$

CM is satisfied if people conclude preferentially γ in the context α and β given that people conclude preferentially β in the context α and γ in the same context α. In order to test CM, we have affected each of the 88 subjects to one cell of the contingency table according to the rules summarized on table 2.

Premise acceptation	Conclusion acceptation: Yes	Conclusion acceptation: Not
Yes	$\Pi(\alpha\wedge\beta)>\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)>\Pi(\alpha\wedge\neg\gamma)$ and $\Pi(\alpha\wedge\beta\wedge\gamma)>\Pi(\alpha\wedge\beta\wedge\neg\gamma)$	$\Pi(\alpha\wedge\beta)>\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)>\Pi(\alpha\wedge\neg\gamma)$ and $\Pi(\alpha\wedge\beta\wedge\gamma)\leq\Pi(\alpha\wedge\beta\wedge\neg\gamma)$
No	$(\Pi(\alpha\wedge\beta)\leq\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)\leq\Pi(\alpha\wedge\neg\gamma))$ or $(\Pi(\alpha\wedge\beta)>\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)\leq\Pi(\alpha\wedge\neg\gamma))$ or $(\Pi(\alpha\wedge\beta)\leq\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)>\Pi(\alpha\wedge\neg\gamma))$ and $\Pi(\alpha\wedge\beta\wedge\gamma)>\Pi(\alpha\wedge\beta\wedge\neg\gamma)$	$(\Pi(\alpha\wedge\beta)\leq\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)\leq\Pi(\alpha\wedge\neg\gamma))$ or $(\Pi(\alpha\wedge\beta)>\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)\leq\Pi(\alpha\wedge\neg\gamma))$ or $(\Pi(\alpha\wedge\beta)\leq\Pi(\alpha\wedge\neg\beta)$ and $\Pi(\alpha\wedge\gamma)>\Pi(\alpha\wedge\neg\gamma))$ and $\Pi(\alpha\wedge\beta\wedge\gamma)\leq\Pi(\alpha\wedge\beta\wedge\neg\gamma)$

Table 2: Model of subject's distribution in the four cells of the contingency table given subject's responses for the Cautious monotony postulate of rationality.

The cell A represents the number of subjects that accepted the two premises rules of the CM postulate and its conclusion rule (remember that a rule is accepted as plausible if $\Pi(\alpha\wedge\beta)>\Pi(\alpha\wedge\neg\beta)$). This cell has a positive contribution to the hypothesis of a positive association between premises and conclusion acceptation. The cell D is another cell with a positive contribution to the association hypothesis. It represents the number of subjects that did not accept the conclusion $\Pi(\alpha\wedge\beta\wedge\gamma) \leq \Pi(\alpha\wedge\beta\wedge\neg\gamma)$ and that accepted only one of the two premises or none of the premises. The two others cells have a negative contribution to the association hypothesis. It is obvious for the B cell (where the two premises are accepted and not the conclusion), but problematic for the C cell. Indeed, obviously, we can conclude nothing about premises and conclusion association given the case where premises are not accepted and conclusion is accepted because the conclusion acceptation can be based on other premises. Our decision to affect this case to the C cell is based on the fact that in considering this case as a negative one, our statistical test of the association hypothesis becomes stronger.

For each postulate of rationality, three contingency tables have been constructed (see the appendix): a table with the lower value of subjects' ellipsoids (or mark) response (min table), a table with the mean computed from the lower and upper values (mean table), and a table with the upper value (max table). Results are summarized in Table 3. This table shows that for all tested postulates except LLE, under all the conditions (min, mean and max), the acceptance of the conjunction of the premises is associated with the conclusion acceptance with a probability of rejecting H0 by error less than .05.

	Min		Mean		Max	
	ϕ	sig.	ϕ	sig.	ϕ	sig.
RW	.26	.016	.22	.034 .	.26	.014
CUT	.45	.000	.32	.003	.33	.002
CM	.31	.004	.29	.006	.32	.002
OR	.54	.000	.49	.000	.39	.000
AND	.56	.000	.55	.000	.57	.000
RM	.49	.000	.36	.000	.36	.000
LLE	-	-	-	-	-	-

Table 3: Values of ϕ and significance (sig.) of the association between the conjunction of the premises and the conclusion for each postulate of rationality under the three conditions min, mean and max (N = 88 for each computation).

The examination of the LLE table (table 4 in appendices) shows that the vast majority of subjects (over or equal to 79.5%) have rejected both the conjunction of premises and the conclusion (cell D) and that none of them has accepted both (cell A). This is the reason why no ϕ value has been computed. The comparison of the ϕ degrees of association shows the following order in postulates' degree of association: AND > OR > RM > CUT > CM > RW. Any generalization from this order would be doubtful anyway because of the independence of the material used for testing each postulate. Another result is that there is no effect from the contingency table model (min, mean and max). It exhibits a strong correlation between min, max and the mean values of min and max. At a finer degree of analysis, the marks made by subjects are distributed along the entire response axis whatever the acceptation or the reject of the conclusion rules given the acceptation of the premises rules conjunction. Our study was not dedicated to the comparison between formal possibilistic and probabilistic models of nonmonotonic reasoning, but these results associated with the degrees of association founded for RM suggest that Adam's conditional logic might not be a satisfying candidate as a descriptive model of human reasoning.

Conclusion

Our results have shown that all the postulates of rationality for which a statistical test has been computed are satisfied by our subjects' inferences based on their own beliefs. Consequently, to the extent that Left Logical Equivalence and reflexivity would be also studied in further experiments, human inference appears to be rational when System P is taken as a norm of rationality. Moreover, the good association found with Rational Monotony confirms the relevance of the possibilistic logic as a normative framework for the study of human inference and reasoning. It does not exclude however other probabilistic approaches. A comparison between possibilistic and probabilistic approaches will represent our next effort on this line of research. Furthermore, these preliminary results encourage us to engage in the testing of other important postulates pointed in the AI literature, and more generally to argue for a close collaboration between psychologists and AI researchers.

Acknowledgments

This research was supported by granted number 90N35/005 of the CNRS-"GIS Science de la Cognition" and by grant number 99001558 of "Conseil régional de la région midi pyrénées". This paper has benefit from fruitful discussions with Salem Benferhat, Didier Dubois and Henri Prade.

Appendices

Content of the rules involved in the test of the rational postulates.

Right Weakening

With « entering high school → being a minor »

And « being a minor ⇒ not having the right to vote »

Do subjects endorse « entering high school → not having the right to vote » ?

Cut

With « being a smoker → having some light, most of the time »

And « being a smoker and having some light most of the time → rarely asking for light »

Do subjects endorse « being a smoker → rarely asking for light » ?

Cautious Monotony

With « being a lawyer → not having one's hair dyed in red »

And « being a lawyer → having a large income »

Do subjects endorse « being a lawyer and having a large income → not having one's hair dyed in red » ?

Rational Monotony

With « non (being a lawyer → not speaking Italian) »

And « being a lawyer → having a large income »

Do subjects endorse « being a lawyer and speaking Italian → having a large income » ?

OR

With « eating a lot of chocolate, of candies → having cavities »

And « rarely brushing one's teeth → having cavities »

Do subjects endorse « eating a lot of chocolate, of candies or rarely brushing one's teeth (or both) → having cavities » ?

AND

With « being a lawyer → not having one's hair dyed in red »

And « being a lawyer → having a large income »

Do subjects endorse « being a lawyer → not having one's hair dyed in red and having a large income » ?

Left Logical Equivalence

With « being a vegetarian ≡ not eating meat »

And « being a vegetarian → not enjoying bullfights »

Do subjects endorse « not eating meat → not enjoying bullfights » ?

Or

50 46 50	18 21 20
2 0 2	18 21 16

Right Weakening

39 35 38	13 13 11
18 19 22	18 21 17

Cautious Monotony

57 58 59	19 18 18
4 4 4	8 8 7

And

60 60 61	1 13 2
15 2 13	12 13 12

Cut

25 23 24	23 28 27
4 5 6	36 32 31

Rational Monotony

49 44 47	8 16 10
12 10 15	19 18 16

Left Logical Equivalence

0 0 0	8 7 8
10 8 8	70 73 72

Table 4: Contingency tables obtained for each postulate of rationality under the min, mean and max condition. In each cell, the upper left value is the number of responses in the min condition; the upper right value is the number of responses in the max condition, and the lower value is the value obtained in the mean condition. In the tables, the upper left cell represents the cases of premises conjunction and conclusion acceptance. The upper right represents the premises acceptance and conclusion non acceptance. The lower left represents premises non acceptance and conclusion acceptance. The lower right cell represents both premises and conclusion non acceptance.

References

Adams, E. W. 1975. *The Logic of Conditionnals*. D. Reidel, Dordrecht, Holland.

Anderson J.R. (1990*). The adaptative Character of Thought*. LEA. Hove and London (UK).

Ajzen, I. 1977. Intuitives theories of events and the effect of base-rate information in prediction. *Journal of Personnality and Social Psychology*, 35, 303-314.

Benferhat, S., Dubois, D., and Prade, H. 1992. Representing Default rules in possibilistic logic. In *Proc. Of the 3rd Inter. Conf. On Principles of Knowledge Representation and Reasoning KR'92.* Cambridge, MA. Oct. 26-29, 673-684.

Benferhat, S., Dubois, D., Prade, H 1997. Nonmonotonic Reasoning, Conditional Objects and Possibility Theory, *Artificial Intelligence*, vol. 92, pp259-276.

Byrne, R.M.J. 1989. Suppressing valid inferences with conditionals. *Cognition*, 31, 61-83.

Chater, N., Redington, M., Kakisa, R. and Oaksford, M. 1999. Rationality the Fast and Frugal way. In *Proceedings of the 21th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Lawrence Erlbaum, pp96-101.

Chomsky, N. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Cohen, L. J. 1981. Can human irrationality be experimentally demonstrated ? *The Behavioral and Brain Sciences*, 4, 317-370.

Cummins, D.D., Lubart, T., Alksnis, O., and Rist, R. 1991. Conditional reasoning and causation. *Memory and Cognition*, 19, 274-282.

Dubois, D. and Prade, H. 1987. *Théorie des possibilités*. Masson, Paris.

Dubois, D. and Prade, H. 1991. Possibilistic logic, preferential models, non-monotonicity and related issues. In Proceedings of the 12th International Joint Conference on Artificial Intelligence IJCAI'91, Sydney, Australia, 419-424.

Dubois, D. and Prade, H. 1994. Conditional objects as non-monotonic consequence relationships. *IEEE Trans. on Systems. Man and Cybernetics.* 24 12 1724-1740.

Edwards, W. 1968. Conservatism in human information processing. In B. Kleinmutz Ed., *Formal representation of Human judgment: 3rd Annual Symposium on Cognition,* April 67. New-York: John Wiley and Sons, Inc, pp. 17-52.

Elio, R. 1997. What to believe when inferences are contradicted. The impact of knowledge type and inference rule. In *Proc. of the 9th Annual Conference of the Cognitive Science Society, 211-216.* Hillsdale, NJ. Lawrence Erlbaum Associates.

Elio, R., and Pelletier, F.J. 1994. The effect of syntactic form on simple belief revisions and updates. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society,* Hillsdale, NJ: Lawrence Erlbaum, pp260-265.

Elio, R., and Pelletier, F.J. 1997. Belief change as propositional update. *Cognitive Science*, 21 4, 419-460.

Evans, St. J. 1989. *Bias in human reasoning. Causes and Consequences.* Lawrence Erlbaum Associated, Publishers. Hove and London (UK).

George, C. 1995. The endorsement of the premises : Assumption-based or belief-based reasoning. *British Journal of Psychology*, 86, 93-111.

Gigerenzer, G. and Goldstein, D. G. 1996. Reasoning, the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.

Johnson-Laird, P. N., Byrne, R. M . J. (1991). *Deduction.* Hillsdale, NJ : Erlbaum.

Pearl, J. 1990. System Z: A natural ordering of defaults with tractable applications to default reasoning. In *Proc. 3rd Conf. On Theoritical Aspects of Reasoning About Knowledge TARK'90*, Morgan and Kaufman, 121-135.

Kahneman, D., Slovic, P., and Tversky, A. 1982. *Judgement under uncertainty : Heuristics and biases.* Cambridge: Cambridge University Press.

Kraus, S., Lehmann, D., and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44, 167-207.

Makinson, D. 1989. General theory of cumulative inference, In M. Reinfrank, J. De Kleer, M.L.

Oaksford,M. and Chater, N. 1994. A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608-631.

Oaksford,M. and Chater, N. 1995. Theories of reasoning and the computational explanation of everyday inference. *Thinking and Reasoning*, 1, 121-152.

Politzer, G., & George, C. 1992. Reasoning under uncertainty and nonmonotonicity : A pragmatic account. *XXVe Congrès international de psychologie*, Bruxelles, 19-24 juillet 1992.

Pollock, J.L. 1987. Defeasible reasoning. *Cognitive Science*, 11, 481-518.

Raufaste, E., & Da Silva Neves, R. 1998. Empirical evaluation of possibility theory in human radiological diagnosis. In H. Prade Ed., *Proceedings of the 13th Biennal Conference on Artificial Intelligence, ECAI'98* pp. 124-128. London: John Wiley & Sons.

Simon, H.A. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118.

Slovic, P., Fishhoff, B. and Lichtenstein, S. 1977. Behavioral decision theory. *Annual Review of Psychology*, 228, 1-39.

Stevenson, R.M., and Over, D.E. 1995. Deduction from uncertain premises. *The Quarterly Journal of Experimental Psychology*, 48A, 613-643.
Stich, S. 1985. Could man be an irrational animal? *Synthese*, 64, 115-135.
Stich, S. 1988. Reflective equilibrium, analytic epistemology and the problem of cognitive diversity. *Synthese* 74, 391-413.
Todd, P. M. and Gigerenzer, G. 1999. *Simple Heuristics that Make Us Smart*. Oxford University Press.
Tversky, A. and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.