Dov Gabbay
John Woods

Resource-origins of Nonmonotonicity

Abstract. Formal nonmonotonic systems try to model the phenomenon that common sense reasoners are able to “jump” in their reasoning from assumptions Δ to conclusions C without their being any deductive chain from Δ to C. Such jumps are done by various mechanisms which are strongly dependent on context and knowledge of how the actual world functions. Our aim is to motivate these jump rules as inference rules designed to optimise survival in an environment with scant resources of effort and time. We begin with a general discussion and quickly move to Section 3 where we introduce five resource principles. We show that these principles lead to some well known nonmonotonic systems such as Nute's defeasible logic. We also give several examples of practical reasoning situations to illustrate our principles.

Keywords: Resource, nonmonotonic logic, psychologism, fuzzy logic, probabilistic logic, defeasible logic, fallacies.

1. Psychologism

Much is said against psychologism in logic, not all of it convincing. Boole's logic is psychologistic [2, pp. 4 and 32]. It is a logic of the ‘laws of thought’. Frege leveled some hefty criticisms against Boole, but he never criticized him for his psychologism. On the face of it, this is more than odd, since Frege's hostility to psychologism is legendary. It turns out that the psychologism against which Frege railed is a rather narrow thing, and certainly not the general idea that logic investigates the laws of thought. One thing that Frege did reject is the view — which he appeared to think is present, implicitly perhaps, in Mill's *System of Logic* [18] — that the laws of logic have the character of well-confirmed empirical generalizations of experimental science. A second thing that Frege rejected was that there is nothing more to the validity of logical laws than their intersubjective validity, that is, their acceptance by the relevent research communities. For all its notational austerity, Frege's *Begriffsschift* [5] reflected a view of the human reasoner. It is that the human reasoner cannot hit the standards of clarity and rigour required by pure mathematics by using a human language — this, too, was Peirce's view. Accordingly, what the human reasoner requires is a suitably powerful

Special Issue: Psychologism in Logic? *Edited by* **Hannes Leitgeb**

DOI: 10.1007/s11225-008-9100-2

artificial language, a language of the sort that Frege is at pains to describe in this seminal work.

This view of what the human reasoner can and cannot do embodies an insight into the cognitive psychology of beings like us. This makes Frege a psychologicist in a way not at all uncongenial to his ambitions for logic. In taking this view, Frege also implicitly endorses a resource-principle, to the effect that what counts as good reasoning for beings like us is influenced by what beings like us are able to do. We may say, then, that the *Begriffsschift* was the first contribution by a modern logician to what the present authors have called a *target-based, resource-dependent* conception of logic (*TR*-logic). However, what Frege did not do was put these assumptions to the critical test, either empirically or formally. He assumed that the reasoning sanctioned in his technical logical writings was psychologically real at least to a non-trivial degree of approximation. While there is no doubt that Frege was a resource-oriented psychologicist about logic, it may be doubted whether the assumptions that drove this view were accurate. Certainly it may be doubted that in any plausible sense of approximation the human reasoner is an approximate executor of, say, classical predicate logic. Much of logic's own development in the past half-century gives credence to these reservations. As these developments find favour in their respective research communities, we may say that the 'trouble' with predicate logic is not that it is psychologistic and not that it is resource-dependent. Rather, what is 'wrong' with predicate logic is that its psychological and resource assumptions are unrealistic in ways that call their utility as a theory of human reasoning seriously into question. In previous writings, we have suggested that, when taken as such a theory, predicate logic is descriptively thin and normatively dubious [8; 10]. In the present essay, we wish to investigate the following proposition: That the origins of nonmonotonic logics lie in the coils of dissatisfaction with the descriptive and normative inadequacies of predicate logic, especially as regards the factor of resource-dependency.

The first systematic study of logic as a distinct discipline can be found in Aristotle. He developed a system of logic as a necessary tool to facilitate, among other things, the study of the other arts and sciences, including most notably argumentation theory. In Aristotle's hands, logic was a humanities discipline. Aristotle's was the dominant position from ancient times to well into the 19th century, when logic took an evolutionary leap and developed (following Boole, De Morgan, Frege, Russell and others) into the predicate logic of today. The logic of the Frege–Russell tradition is studied almost exclusively with reference to mathematics and, to this day, bears the name 'mathematical logic'. With the introduction of mathematical logic, the general study of logic also took a 100 year detour — the main effort moved from human-oriented investigations towards applications in mathematics. Indeed, even the beginnings of the theory of computation (famously developed by Church and Turing) is still taught in some symbolic logic classes today as a piece of 'meta-mathematics'.

Around the middle of the 20th century a new urgent push was given to logic by investigation of the human agent in his daily activities and behaviour. The impetus started with artificial intelligence and computer science which wanted to develop devices, programs and theories to assist or even replace human activity.

Logic was heavily used by these disciplines and was forced to evolve even further to be able to respond to the needs of the applications. Soon the application of these new devices and programs spread into natural language, law, psychology, decision theory, communications, robotics, computational philosophy, philosophy of science, and so on. The social impact of the development of computers, robotics and communication stimulated the use of logic in the respective sciences and forced a parallel development in logic itself. In practice, this consisted in scientists in each local discipline developing their own logics and systems as they were required. The result was chaotic landscape of new 'logics'.

Under the impetus of computer science, logic is well on its way towards recovering its historic mission as a study of human reasoning. In making this turn from pure mathematics, logicians have shown a readiness to admit agents as load-bearing notions of their theories. In classical computer science, as well as standard systems of belief dynamics and rational decision theory, agents are conceived of as self-aware thinking creatures, whose rationality consists in reflective deliberation about ends and means, central to which is the maximization of subjective utilities. Possibly the most exaggerated versions of this assumption are Savage's [22] invention of utility functions over possible future possibilities of the world, and De Finetti's [4] conceit that once a probability distribution is at hand, all subsequent belief revisions will be effected by conditioning and nothing else. Of course, everything that is known empirically about human reasoning makes it clear that long-term computations are rarely made by individual reasoners on the ground [15]. This faces the would-be theorist of human reasoning with an obvious choice. He can acknowledge what the empirical record reveals but marginalize it in his theory, either by charging actual reasoners with systematic and widespread irrationality or by reconstituting these descriptively inadequate assumptions as normative rules which ideally rational agents follow by definition and which being like us somehow approximate to. The theorist's second option is to accept what the empirical record reveals and give it a central place in his investigations. In previous work, we have tried to make it clear why we think the first of these options should be resisted [10]. Option two orients a logic of human reasoning to the facts of psychological reality. It embeds the assumption that, once we admit agents to logic, it is best to admit them as they actually are, warts and all. Since beings like us come with psychologies as standard equipment, it would be feckless to attempt to divine the norms of human thinking without taking due notice of how the human reasoner is actually constituted, of what his interests are and what he is capable of.

No one doubts that in re-admitting psychological factors into logic, that the logician and the psychologist will bring to bear different preoccupations and different skills. Roughly speaking, they will reflect what they are respectively interested in and will deploy the theoretical techniques that they are respectively good at. In this there is no occasion for what John Macnamara calls a "border dispute" [17]. One measure of the respective parties' success is the research literature that flows from their efforts. Another is the extent to which these literatures take note of one another. Our view is that, in taking the practical turn, logic has re-committed to psychologism. At present one can only speculate on the likelihood of its success. Logicians have yet to generate the requisite research literature. And they have not yet developed the habit of talking to their colleagues in psychology. Short of an impossibility proof, there is only one way to gauge the value of a psychologically faithful logic of reasoning. Produce the literature and engage in the conversations. Then we'll see.

In previous writings,[1] we have generated a conceptual model of individual thinkers as a paradigm of *practical agency*, and we have characterized the practical agent as one who prosecutes his cognitive agendas with comparatively slight cognitive resources, such as information, time, and storage and computational capacities, and who sets the cognitive bar at heights that enable them to be negotiated with the resources at hand. By these lights, it is characteristic of the practical agent to be a cognitive economizer, to favour in most things satisficing strategies over maximizing strategies. This matters for *TR*-logics in an obvious way. It suggests that the norms of individual human reasoning must reflect the constraints on how practical agency operates in actual conditions. In particular, it must seek to find in such norms indications of their value in an individual's cognitive economy of scant resources. In this paper, we want to try to locate the origins of nonmonotonicity in this

[1] [8; 9; 12; 13].

conception of the resource-and design-bound practical agent. We take it as empirically given that the individual agent's reasoning behaviour is dominantly nonmonotonic. When we say that our intention is to seek the origins of these nonmonotonic proclivities, what we have in mind is that they have been generated by the psychological constitution of the individual cognitive agent, in tandem with the nature and extent of his resource-dependencies. As we hope to show, the individual reasoner operates in a cognitive economy of comparatively scant resources, and is rationally conditioned not to set his cognitive targets at heights that typically out-reach those resources. Seen this way, our nonmonotonic proclivities have an economic rationale. They assist the individual agent to prosecute his cognitive agendas efficiently.

Let us first make some formal observations. There is a wide variety of diverse formal logics grouped together under the name "nonmonotonic" since it suffices for a logic to be nonmonotonic that it have a nonmonotonic consequence relation. If Δ is a database and A a wff, $\mid\!\sim$ is nonmonotonic consequence precisely if it is jointly possible that $\Delta \mid\!\sim A$ but $\Delta + B \not\mid\!\sim A$, where $\Delta + B$ denotes the result of adding B to Δ.

The problem with the notion of nonmonotonicity is that it is not a characterising property of logics but rather an indication of a lack of a property, *monotonicity*, which may fail for many different reasons. We therefore seek some thematic principles which can generate nonmonotonicity and hope to classify the wide variety and diversity of the class of nonmonotonic systems, using these principles. We are looking for humanly oriented principles and not formal principles. Formal principles do exist. In fact, they are (formally) very attractive. They have the form:

- Take a monotonic consequence $\vdash$, change it in a prescribed mathematical way (say π) either syntactically or semantically, and this yields a new nonmonotonic consequence $\mid\!\sim_\pi$.
  The semantic methods work by way of the condition:
- $A \vdash B$ iff in all models $\mathbf{M}_A$ of A, B holds.
  Then we use a selection function $\mathbf{f}$ on sets of models and define
- $A \mid\!\sim_{\mathbf{f}} B$ iff B holds in all models of $\mathbf{f}(\mathbf{M}_A)$.
  The syntactical methods work by way of the condition:
- $A \vdash B$ iff for all $X \in \mathrm{Wff}, A \wedge B \vdash X$ iff $A \vdash X$.
  Choose a suitable set $\Delta \subseteq \mathrm{Wff}$ and let
- $A \mid\!\sim_\Delta B$ iff for all $X \in \Delta$ we have $A \wedge B \vdash X$ iff $\vdash X$.

The above formal principles yield a good many logics. But what is their intuitive meaning in terms of our day-to-day existence? We are seeking, as we said, some human principles, which generate the nonmonotonic systems. Here are some candidates.

1. Nonmonotonicity arising from the need of constant revision and change (investigated initially by Gärdenfors and Makinson [14]).
2. Nonmonotonicity arising from bringing meta-level considerations into the object level (new investigations by the authors).
3. Nonmonotonicity arising from resource limitation of reasoning agents (which is the subject of this paper).

We are looking now at the resource and time limitation problems of human beings. We will identify several reasonable 'simplifying' principles humans use to facilitate reasoning in their everyday existence with which they compensate for their limited cognitive resources. We shall use these principles to derive some well known nonmonotonic systems.

Suppose that it is, as we ourselves believe, true that the human agent is one who in general must prosecute his cognitive agendas under press of scant resources, including rather centrally, resources such as *information*, *time* and *storage and computational capacity*. Agents thus positioned we identify as practical agents.[2] It would seem to follow that practical agents execute their cognitive tasks with the aid of various scant-resource adjustment strategies (SRAS), which all applies to serve their cognitive ends rather well.

We need to say how the principle of scarce resource generate new nonmonotonic systems. We need to compare it with the mathematical way. Suppose $A \vdash B$. A human will use proof theory to work his way form A to B. If he is resource bounded he will take shortcuts, compute approximations, estimates, etc. Call his way of handling things quickly by π.

Then the question of showing that $A \vdash ?B$ through some human chain of reasoning (with infinite resource) which may or may not be successful will turn into the question $A {\mid\!\sim}_\pi ?B$, when we apply π to the process, and thus limiting the resources available to the process.

At first sight we may think that less can be deduced now. It may be the case that $A \vdash B$ can be demonstrated if we have enough resources, but with the limitations imposed by π, we may have $A \not\mid\!\sim_\pi B$. This is true if all the processes involved are monotonic. However, if $A \vdash B$ involves some

[2]Correspondingly, a "theoretical" agent is typically a collective or institutional enterprise where command of the requisite resources is greater and cognitive tasks are those appropriate for such resource-reaches. One might mention Nato in this connection, or present-day particle physics. For a more detailed discussion of this interpretation of the distinction between practical and theoretical agents, see [Gabbay and Woods, 2005].

principle such as failure, then π may increase the chances of failure and thus can deduce more.

For example, in logic programming we may have the database Δ below with $\Delta \not\vdash q$, ($\neg$ is negation as failure)

1. $\neg p \rightarrow q$
2. $p \rightarrow a$
3. p

If we limit the 'use' of p to once only, (as we do in linear logic) then by deducing a we use up p and hence we can later deduce q as well!

2. Resource-dependency and heuristics

There is a substantial consensus that, even when it operates satisfactorily, short-cut reasoning is reasoning faute de mieux or "on the cheap". This opinion lies at the heart of the doctrine that what is gained by a heuristic will be offset by its diminished epistemic payoff, and accordingly, that the faute de mieux character of heuristic reasoning convicts the reasoning of a lesser degree of rationality than its superior cousin, were it only available to beings like us. But on a resource-based approach to logic, and the notion of rationality that it embeds, the faute de mieux assumption cannot be right. In particular, it cannot be right to claim superiority for a mode of reasoning that lies beyond a human's competence, provided that the reasoning that does lie within it operates successfully. There is a fabled cow that jumped over the moon. No one doubts that here is a being who jumped higher than any human has ever dreamt of. But it would be a weak joke to conclude from this that even the best of human jumpers are pretty small beer. No, the best of human jumpers are the best that the physical and psychological constitution of the human animal admits of. It is not much different with human reasoners. An individual thinker's success is not constituted by his simulation of the behaviour of a logical god. He achieves his success by tending to his knitting: by tailoring his cognitive targets to his interests and to his wherewithal for attaining them. There is an external test of the soundness of this conception of an individual agent's rational success. It is behaviour which collectively and over time satisfies the Enough Already Principle:

> *Collectively and historically, human agents are right enough enough of the time about enough of the right things to survive and prosper, and occasionally build great civilizations.*

3. Five resource principles

Among a practical agent's resource-compensation strategies are the following principles:

(a_1) He divides his world into natural kinds.

(a_2) He is a hasty generaliser, or 'thin-slicer'.

(a_3) He does not have much time or resource and therefore tends to approximate/cut short any lengthy process.

(a_4) He is cautiously impulsive. In particular, he has efficient feedback mechanisms. If things don't feel right, then he makes a sudden change of course.

(a_5) Central to our agents existence/survival is the need to take quick action and so all reasoning tasks are designed to serve/enable actions. Our agent has little time for irrelevant considerations.

Discussion of (a_1) and (a_2)

The meaning of (a_1) is that the human divides the world into classes (denotata of unary predicates) $C_1(x), C_2(x), \ldots$. As a first attempt at modelling, we assume this classification is crisp, namely the predicates $C_i(x)$ are two valued. However, we know from experience that any classification attempt with limited resources will encounter items which are difficult to classify exactly. This naturally invites fuzzy predicates. We shall discuss this at a later section. The logical machinery developed for the crisp case will need to be modified for the fuzzy case.

(a_2) means that he is hasty in generalising relationships between these classes. If our agent twice felt sick eating certain leaves, then he generalises that these leaves are not edible. There is no time to investigate. Consequently our low-resource agent is going to have three kinds of rules. (Our language has literals of the form $+P(x) = P(x)$ or $-P(x) = \neg P(x)$.):

1. Absolute rules of the form

$$\pm P_1(x) \wedge \ldots \wedge \pm P_n(x) \rightarrow \pm Q(x)$$

   These rules are absolute because they follow from the meaning of the predicates. They are *not* hasty generalisations based on experience. Example:

   - Penguin(x) → Bird(x)

When a rule has the form $\varnothing \to \pm Q(x)$ (i.e. no $\pm P_i(x)$ exist), we regard this as a *fact*. The agent observed something of x and used the predicate Q to characterise it. We use a single arrow '$\to$' for absolute rules.

2. Hasty generalisation rules of the form

$$\pm P_1(x) \wedge \ldots \wedge \pm P_n(x) \Rightarrow \pm Q(x).$$

Our agents noticed some examples $b_1, \ldots, b_k$ where there is a connection between $\bigwedge_i \pm P_i(b_j)$ and $\pm Q(b_j)$ and generalised it into a rule. He may have tried, e.g. Mill's method, to do some tests, but he bears in mind that this rule may hit some counterexamples. We use $\Rightarrow$ (double arrow) for defeasible rules. $\varnothing \Rightarrow Q(x)$ is called a *presumption* (Nute [19]).
Example:

- Bird$(x) \Rightarrow$ Fly(x)
- Penguin$(x) \Rightarrow \neg$Fly(x).

3. The third kind of rule is (what Nute calls a *defeater*) the realisation that under a certain combination of circumstances ($\bigwedge_i \pm P_i(x)$) it is better not to conclude anything. We write this as

$$\pm P_1(x) \wedge \ldots \wedge \pm P_n(x) \longrightarrow\!\!\!\!\triangleright \pm Q(x).$$

(read $\longrightarrow\!\!\!\!\triangleright$ as $\not\Rightarrow$ streamlined).
This is a safety mechanism since some rules are hasty generalisations, and our agent knows when to be careful!

Discussion of (a_3)

For large ranges of cases, a practical agent does not have much time to spend. So if he is looking for something and cannot find it in the obvious places, he jumps to the conclusion it is not there. This is 'negation as failure'. Another aspect of (a_3) are various approximations of various arguments/proof systems: 'Cannot prove A in k steps then it is not provable'. 'Anyone who objects to this marriage, let him come forth now!', etc.

Of course our agent realises his conclusions depend on the information currently available and on how much time and effort he is able to devote to his reasoning.

To draw a conclusion we must use as much data as possible (given our time and other resources) as bears relevantly on our task. However, our conclusions could have changed had we more time and information).

Example 3.1 (for (a_1), (a_2):). *So for example the database*

$Bird(x) \Rightarrow Fly(x)$
$Penguin(x) \rightarrow Bird(x)$
$Penguin(x) \Rightarrow \neg Fly(x)$
$Bird(a)$
$Penguin(a)$

would conclude $\neg Fly(a)$.

Since we must use all available information that bears relevantly on our task, a rule of the form $\bigwedge_i P_i(x) \Rightarrow \pm Q$ *is weaker than a rule of the form* $P(x) \wedge \bigwedge_i P_i(x) \Rightarrow (x)$ *because it is based on less data.*[3]

Our *third principle* (a_3) assumes that in some cases, the agent tries to find all elements in a class $C()$. For some classes he may feel he has information enough to practically assume that if x cannot be deduced to have property $C(x)$, then it is defeasibly in $\neg C(x)$. For such classes we have what is known in modern terms as the negation by failure rule, which we can present as

$$\mathbf{fail}((C(x)) \Rightarrow \neg C(x)$$

Note that this rule is "on the run" rule. **fail** means "fail to deduce using local resources". There may not be enough time to conduct an exhaustive search.

Discussion of (a_4)

Our agent does not have much time to spend. So if he is looking for something or is doing something and it doesn't feel right, he immediately aborts and runs away. Principle (a_4) is really a Neural Net principle, where we form an opinion/recognition all of a sudden.The previous (a_1)–(a_3) are more rule based.

Imagine an insurance salesman trying to persuade you to buy some insurance. At some point he might realise (all of a sudden) that you are not going to buy. He will suddenly thank you and leave. This does not happen, by the way, with Jehovah Witnesses. They never leave. Nor does it happen with large organisations; they are very slow to shift.

It is well known that single agents respond to feedback and criticism much more quickly than large organisations. Large bureaucracies do not like impulsive reactions. They tend to substitute rules, regulations and red tape.

[3]Imagine one harassed hasty generaliser saying to another: "Don't touch these pointed leaves, they are poisonous', and he gets the reply: "Yes, but I am boiling them before eating them!".

Discussion of (a_5)

Another aspect connected to the (a_3) and (a_5) principles is a local default rule. It's role is to save on time and effort in checking the extension of a predicate $C(x)$. We may decide on the basis of evidence $\alpha(x)$ that if $C(x)$ is locally consistent with our current view of the world then we can add $C(x)$ as the true world. We write this rule as

$$\alpha(x) : \frac{C(x)}{C(x)}$$

This may look like Reiter's default rule, but it is not. Reiter's machinery [20] is not practical, and we doubt whether his particular conception of default is at all realistic. Imagine an agent wanting to execute an action **a**. **a** has precondition A and postcondition B. He gathers evidence A' to check whether A holds and finds that it is consistent to assume A given the A' he has got. He might feel confident to proceed with the action. We note here that large institutions are more cautious and have more resources and tend therefore not to use defaults so much. They would investigate further. What we have here is that our agent who is on the run while trying to make some deductions to guide his next step may choose to add some locally consistent $C(x)$ which makes sense in his current context. Note that the fail rule and the default rule complement one another, one for the positive and one for negative information. The local default rule cannot be perceived as a hasty generalisation of the form $\alpha(x) \wedge \textbf{fail}(\neg C(x)) \Rightarrow C(x)$. The latter means a more systematic check of **fail**. Recall $\neg C(x)$ itself involves failure. The rule can be understood in such a way that in the presence of $\alpha(x)$, failure to show $\neg C(x)$ indicates that $C(x)$ holds.[4]

4. Sample systems

We are now ready to formulate our first system. We choose a labelled formulation.

---

[4]Imagine for example that you are trying very quickly to check whether the library has a certain book on logic. You scan the shelves and find nothing. You deduce it doesn't have the book; i.e. $\neg H(b)$. Then you realise they were shelves in the annex. So the search failed to be reasonably exhaustive. This is **fail** ($\neg H(b)$). Of course, ordinarily you would not want to conclude $H(b)$ from this. But if you are an inspector evaluating the library for adequacy, you might have to.

DEFINITION 4.1 (A limited resource model for a simple agent). *The language contains unary predicates, constants and variables, negation ¬ and various other notation as follows:*

1. *A literal has the form $P(x)$ (positive) or $\neg P(x)$ negative. It is grounded if it has the form $\pm P(a)$, where a is a constant.*

   Let A_i, B be literals

2. *An absolute rule has the form*

$$\bigwedge_i A_i(x) \to B(x)$$

3. *A defeasible rule has the form*

$$R(x) : \bigwedge_i A_i(x) \Rightarrow B(x)$$

4. *A safeguard (defeater) has the form*

$$R(x) : \bigwedge_I A_i(x) \rightarrow\!\triangleright : B(x)$$

5. *A negation by failure rule has the form*

$$R(x) : \mathbf{fail}((C(x)) \Rightarrow \neg C(x)$$

6. *A default rule*

$$R(x) : \bigwedge_i A_i(x) : \frac{B(x)}{B(x)}$$

7. *A theory Δ is a set of grounded literals and various rules. We can assume that all items in Δ are labelled (named). The language of names is a completely separate language used just to enable us to trace proofs. Thus Δ can be assumed to have the form*

$$\begin{array}{l} l_1 : A_1, \ldots, l_k : A_k \\ r_1(x) : R_1(x), \ldots, r_m(x) : R_n(x) \end{array}$$

   where A_i are literals and R_i are rules. The letters l_i name the literals available and may be a more complex label giving more information about the literal.

   We now explain how we reason with such a database.

EXAMPLE 4.2. Let Δ be

l_1:	Penguin(a)
l_2:	Genetically-engineered (a) $(GE(a))$
l_3:	Penguin (b)
$R_1(x)$:	Penguin$(x) \Rightarrow \neg$ Fly(x)
$R_2(x)$:	Bird$(x) \Rightarrow$ Fly(x)
$R_3(x)$:	Penguin$(x) \rightarrow$ Bird(x)
$R_4(x)$:	**fail**(Penguin(x)) $\Rightarrow \neg$ Penguin(x)
$R_5(x)$:	$GE(x) \wedge \neg$ Penguin$(x) \Rightarrow$ Fly(x)
$R_6(x)$:	$\text{Fly}(x) : \dfrac{\text{Bird}(x)}{\text{Bird}(x)}$
l_4:	$GE(c)$.

We can prove Bird(c) as follows.

We first note that Penguin(c) fails. Then we deduce $\neg$Penguin(c), and from this we deduce Fly(c).

Note that in the proof theory for such a system the defaults are not Reiter defaults. In other words, given $A : \frac{B}{B}$ and $A : \frac{\neg B}{\neg B}$, we choose the one we want to enable actions and carry on. There is no question of considering extensions,etc.

Another point to consider is that the default $\Delta : \frac{X}{X}$ can be read context-sensitively. If all we know is Δ and X is consistent then we can add X. Reiter's default is monotonic, i.e. if we know more than Δ and X is consistent we can add X. This point has already been addressed in [7, pp. 240–246].

5. Example

It would be good to illustrate our principles in some reasonable real life example.

> Mrs Smith holds a million pound life insurance policy on her husband, issued many years ago, when they were married. The premium is paid July 1st every year by cheque. If the premium is not paid by July 15th the policy lapses. Mrs Smith posted a cheque on July 1st by ordinary post. On July 7th there was a terrorist attack at King's Cross. Mr. Smith was in the area and (as a result of the shock of the bombing) died of a heart attack. Because of an administrative error, Mrs Smith's cheque was not cashed by the insurance company until August. An internal record exists at the insurers showing the cheque

arrived on July 3rd. Mrs Smith's lawyer is presented with two logical statements[5]

Δ_1: Mr Smith dies $\wedge\neg$ by terror $\rightarrow$ insurance money

Δ_2: Mr Smith dies $\wedge$ by terror $\rightarrow$ government money

The policy has the usual clause that cases of terrorism and war are not covered. The Government has special funds to compensate victims of terrorism. The question is was Mr Smith's heart attack to be considered as caused by the terror attack or just a coincidence with it? Δ_1 and Δ_2 are databases of evidence supporting the statements. Δ_1 needs to establish that the policy was in force in July and Δ_2 needs to establish the form and reason of death.
Mrs Smith's lawyer realises he has two options

1. Establish that Mr Smith died as an (indirect) result of the terrorist attack and claim money from the government's special fund.
2. Claim the heart attack is not a result of terror and claim money from the life insurance company.

To pursue option 2 he needs to establish the validity of the insurance policy. To this end he needs to get hold of the record of receipt of the cheque at the insurers. He is afraid that the insurers will delete the record and claim that the cheque was posted too late. They would reason backwards: the cheque was cashed in August, they have 'failed' to find a record of receipt and so, by common sense reasoning, this implies that the cheque arrived at the end of July.

The point of the above story can be grasped by an ordinary person immediately, but its formal logic however is quite complex. It involves

[5]In claiming injury, two factors are generally involved. That of causation ("but for X; Y would not have happened") and that of remoteness (in tort is "reasonable forseeability"). Causation is much easier to show. Remoteness is different. The government wants to limit the number of people who can claim (e.g. claims like "he saw the terrorist attack live on TV and thought his daughter may have been there and had a heart attack as a result"). In such a case causation will be shown but reasonable forseeability may not. Thus they would want to put the highest hurdle possible to exclude a maximal number of claimants. So they will probably require direct presence at the event. The insurance company also wants to exclude the maximal number of people from claiming and they do this by making the exclusion to their liability of terrorist attack as wide as possible. They want their definition of death by terrorist attack to be as wide as possible. So they would have only the causation test and no requirement of remoteness or direct presence. So the two institutions may have different tests for what the commonsense "lay" cause of death is.

If the tests of "lay" causation are different, Mrs Smith could lose both claims.

1. Labels for time
2. A variety of contexts.
3. The lawyer's simulation and reasoning of what the insurance company might do is a yet higher level (call it the *meta-level*) involving the following principles:
   (a) Deletion of data from a database
   (b) Use of negation as failure
   (c) The temporal advance of events using actions
   (d) Considerations of consistency and conflict of laws
   (e) Consideration of actions to be taken now because of possible simulated future scenarios
4. The overriding problem is resource-oriented. Time and money.

Let us first approach the problem intuitively, as friends of Mrs Smith and see how we can reason on her behalf.

We see two weak points in her claim from the insurance company ('weak' in the sense that the insurance company may make it difficult or impossible to claim).

1. The question of whether the policy was renewed in time. There should be a record of the date of receipt. Would the company destroy the record? If it is a big bureaucracy then it is not likely. Even so, we must get confirmation of the date of receipt.
2. The second question is how to establish that the heart attack was not caused by the explosion? This can be very much context dependent on where exactly Mr Smith was at the moment of the explosion and whether he was hit by the blast. It may also depend on whether he had a weak heart or not.

Let us make the following assumptions:

1. It is unlikely that the insurance company would 'not find' a record of recept of the cheque.
2. It is a legal and medical problem whether the blast should be considered as directly responsible for the death of Mr Smith.
3. The government is not likely to pay as much as the insurance company.

$$A \xrightarrow{\pm x} B$$

Figure 1.

Under the circumstances it may look better to claim that the heart attack was indirectly caused by the blast and claim compensation from the government. If the government denies compensation on the grounds that death was not caused by the blast then one can use this ruling and claim from the insurance company.

This raises the question of resources. If a claim is made from the government, who has the burden of proof that the death was caused by the blast? Is it the claimant? or should the government prove the death was not caused by the blast? The process of proof is costly.

Alternatively, should we claim from the insurance company on the basis of easy-to-prove 'death'? Is it then the insurance company's responsibility to prove that the death was caused by the blast and hence by terrorist activity? This seems more likely but what the insurance company might do in practice is to refuse to pay and let Mrs Smith simply be forced to sue. Even so, the burden of proof still shall probably be on them.

Let us now see how we would represent our options and reasoning formally. Writing all the above statements in some logical language is not good enough. We would lose the transparency of the strucutres involved. We think we can do better using several formalisms and linking them together by a suitable meta-level logical device. We try to keep close to how a human would do it.

Obviously, we need at least three major components:

1. a time/action diagram
2. an argument network presenting the major arguments and counterarguments
3. a resource considerations map.

Let A, B denote arguments. We use figure 1 to denote the situation where argument A attacks $(-x)$ B with strength x or supports $(+x)$ B with strength x, for x a number between 0 and 1. For a study of this model see [1].

Using this notation we can write

$D =$ insurance company deletes the receipt record of Mrs Smith's cheque.

$E =$ Insurance company will save £1m.

$F =$ Insurance company is a bureaucratic organisation with distributed responsibility and procedures.

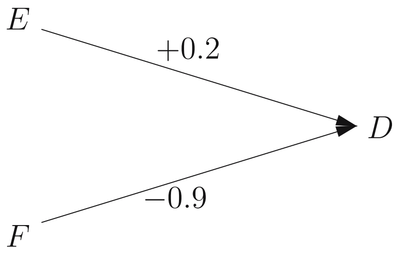

Figure 2.

Figure 2 now represents the relationship between these arguments. We need to combine these two numbers $(+0.2, -0.9)$ to get a new, probably high number, attacking the notion that the receipt record will be deleted.

Similarly let

$P_1 =$ Mr Smith was shocked by the blast

$P_2 =$ Mr Smith died of a heart attack

$P_3 =$ Death was caused by the blast.

$P_4 =$ Legal and medical argument that although the bast is a secondary possible cause for the heart attack it is not to be considered the cause of death.

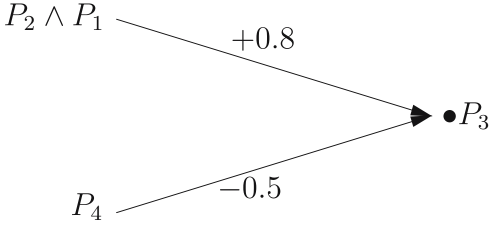

Figure 3.

Available actions

action	precondition	postcondition
a = ask for confirmation of receipt of cheque	none	insurance policy valid
b = file claim with insurance company	valid policy and death certificate	claim considered
c = file claim with government	State Mr Smith was hurt by shock of the blast (cause of death?)	claim considered
d = insurance company claims death caused by act of terror	P_3	claim **b** not valid
e = government claims death not caused by act of terror	$\neg P_3$	claim **e** not valid
f = appeal to public opinion	seeming injustice by public bodies against the victim Mr Smith.	hopefully some resulting solution

Calculation of cost

Obviously a lot hinges on P_3, whether the blast is to be considred cause of death. It is costly to establish P_3 or $\neg P_3$. All the other costs are relatively low. So the course of action to be taken is the one which minimises the cost to Mrs Smith. Here is one option:

Step 1: Take action **a**
Step 2: Take action **b**

In the event that the insurance company takes action **d** then challenge them to establish P_3. Try to get an immediate response from the insurance company on whether they are going to take action **d** or not.

Step 3: In the event of action **d**, claim from goverment (action **c**) providing the proof the insurance company put forward asserting P_3, therefore the government is responsible. If the government take action **e**, then hold a press conference (action **f**).

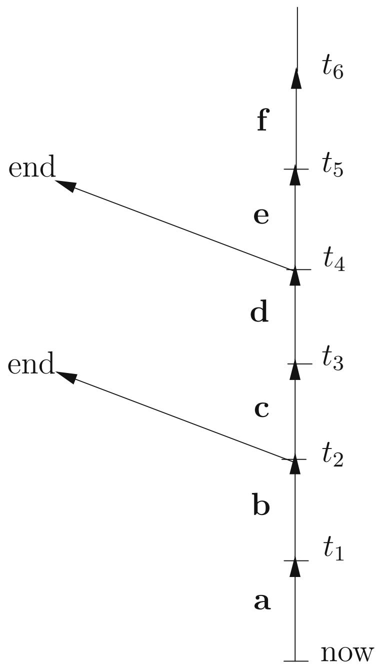

Figure 4.

It is painfully clear that the overriding consideration here is that of cost and resource availability. The common sense normal course of action seems to be to claim from both the insurance company and the government. But Mrs Smith cannot risk the cost of the process of facing government bureaucracy. So she is claiming only from the insurance. The insurance company may also delete documents from the database in order to save money and the only reason they might *not* do that is again the complexity of its bureaucracy.

Time action flow

The above suggest the time action flow of Figure 4, t_i are time points. Time 'ticks' because of actions taken.

6. Another Example: Fallacies

In 1970, C.L. Hamblin fulminated against logicians for having given up on the fallacies programme, abandoning it to the writers of first-year textbooks, who, whether out of incompetence or indifference, confined themselves to rather foolish caricatures [16]. Hamblin was of the view that the revival of the fallacies programme would require the admittance of agents to logical theory, and he saw the logic of dialogue as a natural way to link the behaviour of agents to one of the contexts in which fallacies naturally arise. In so proposing, Hamblin adopted what he took to be (we think rightly) the traditional conception of fallacies as arguments or inferences that are

1. errors
2. attractive[6]
3. universal, that is, widely committed.
4. incorrigible, that is, subject to high levels of post-diagnostic recidivism.

Hamblin was a comparatively rare logician who took pains to learn the history of the subject, and his book is much enriched by this knowledge. But he also tended to accept as loosely canonical in 1970 that list of fallacies which Woods [24] has baptized "The Gang of Eighteen". In [26] they are the following: *ad baculum, ad hominem, ad misericordiam, ad populum, ad verecundiam, ad ignorantiam, post hoc ergo propter hoc,* affirming the consequent, denying the antecedent, begging the question, equivocation, amphiboly, biased statistics, composition and division, faulty analogy, gambler's, *ignoratio elenchi*, and hasty generalization. In the tradition from which the Gang of Eighteen has evolved, the received view is that, whatever else might be said about them, the error they commit is either deductive invalidity or inductive weakness (in the technical sense developed by the main approaches to confirmation theory.) One of the benefits of paying attention to what human individuals are actually like — what they are interested in and capable of doing — is that it helps us see that in a quite general way errors are relative to the cognitive targets agents have set for themselves and to the attainment-standards which those targets embed. If someone seeks a proof of a proposition in category theory, then he will fail if his proferred demonstration is invalid. Producing that demonstration would be an error in the light of that target and that attainment-standard. On the other hand, to take a rather tendentious example, if someone's target is to find out whether his neighbour enjoyed last night's game, it would be seriously misguided to judge his reasoning an error if it involved the invalid inference that the neighbour did enjoy himself from the neighbour's avowal that he did. This teaches us a valuable lesson about the assessment of reasoning.

[6]Michael Scriven calls fallacies the "attractive nuisances of argumentation, the ideal types of improper inference" [23, p. 333].

Before one faults a person's reasoning, it is necessary to identify his target and the standards it imposes. In an obvious sense, then,

- *Attribution precedes assessment.*

While there are contexts, and targets, that call for the deployment of valid reasoning, beings like us have a stake in not setting such targets needlessly. What is known of the empirical record amply demonstrates that by and large, beings like us – individual practical agents – do not set targets that call for validity. The same may be said for targets that call for reasoning that is inductively strong. Nato, MI5 and population biology may more or less routinely set the goal of a scientific confirmation of its lawlike generalizations, and may at times bring this off by the Bayesian machinations of conditional probability. But this, too, is not typically our situation as individuals. As Mill wisely observed, induction is not for the likes of us; it is for *societies.*

From this we have it that as they are traditionally conceived of, the Gang of Eighteen aren't fallacies after all. More carefully,

- *Either the eighteen aren't errors, or if they are, they aren't errors typically committed by us. Either way, they aren't* — for us – *fallacies on the traditional conception.*

This we might call the *Negative Thesis.* It arises out of an examination of what the individual agent is typically *up to* when he reasons. As the present authors see it, the Negative Thesis is not all that earth-shaking. It flows rather directly from the Attribute-Precedes-Assessment Thesis in conjunction with empirical inspection of the reasoning behaviour actually evinced by human individuals on the ground. A more daring thing to venture would be the *Positive Thesis*, to the effect that

- *The Gang of Eighteen are cognitively benign scant-resource adjustment strategies* for beings like us.

This is not the place to try to make good on the Positive Thesis.[7] For the present we shall test it only as it applies to hasty generalization.

What would it take for hasty generalization to be a cognitively virtuous reasoning strategy? What would its virtue consist in? Certainly not its inductive strength, since all agree that this is what it lacks. But the mere

[7] This is attempted in [11] in progress.

fact that the inductive weakness of a benign hasty generalization is not an error does not constitute itself as a strength. So what does?

We are notorious hasty generalizers. Hasty generalization is as natural to us as breathing. How, then, can we say that, although we generalize hastily "all the time", that we don't commit the error of hasty generalization with a frequency required to make it a fallacy? What we want to suggest is that when we generalize hastily, it is typically not the case that the cognitive target we will have set for ourselves is one whose attainment standard is inductive strength, that is, the degree of confirmation necessary for the lawlike generalizations of experimental science. Of course, the reason for this is not epistemic sloth, but tactical realism. Since in the general case individual reasoners lack the wherewithal for meeting these standards, typically they avoid committing themselves to targets that call for them.

If this is right, it must be possible to generalize hastily and correctly in relation to targets that don't demand this high degree of inductive strength. Consider a case. You are tramping in the bush and you see your first ocelot, four legs and all. You say to your guide, "Funny, I always imagined that ocelots were two-legged." In generalizing to four, you relied on a single positive instance, the smallest of non-zero samples. You also got it right. What accounts for this? Why do we generalize on a single instance of four-leggedness but not on a single instance of being-spotted-on-Thursday? Perhaps it has something to do with our aptitude for recognizing natural kinds. Possible it also matters that when we do generalize in this way, the generalization is *presumptive* rather than unqualified, and what is generalized *to* is a *generic* claim, rather than a universally quantified conditional statement – hence a statement with regard to which an error in a particular case need not cost you the generalization's truth.[8] Accordingly, the two cases exhibit some interesting structural differences.

Case 1

1. For all x, if x is an ocelot, x is four legged. (Universally quantified conditional)
2. Therefore, Ozzie the ocelot is four-legged. (An instantiation of (1)).

The "therefore" marks deductive consequence.

[8] Carlson and Pelletier [3] is the definitive single-volume reference work for genericity.

Case 2

1. Ocelots are four-legged. (Generic claim)
2. Ozzie the ocelot is four-legged. (A default from (1))

The "so" marks presumptive consequence.

As we see, with respect to Ozzie the *three*-legged ocelot, it is possible to be wrong in particular without being wrong in general. The reason for this is that, while universally quantified conditionals are *brittle* – i.e., falsified by a single true negative instance, generic claims are *elastic* — i.e., they have the capacity to retain their truth in the face of false instances. In the first case, there are two things to correct, both the faulty instance and the universal quantification it instantiates. In the second case, there is only one thing to correct, the faulty instance which is now seen as a *default* from the generic claim.

A basic appeal of nonmonotonic strategies is that they are *feedback-friendly*. They are natural occasion for the efficient and timely correction of error. They are thus contributions to the agent's cognitive economy. Generalizing from a single instance is likewise nonmonotonic, and so are the subsequent default inferences from what is generalized to. All of this is economically motivated, mandated by the resource-constraints peculiar to individual cognitive agents.

7. Introducing Fuzziness

We have mentioned already that our resource-poor agent divides the world into categories. With lack of resources and lack of time, this division itself cannot be crisp. So, for example, we may have the following two rules about hunting:

- $\text{Vicious}(x) \wedge \text{Big}(x) \Rightarrow\sim \text{Hunt}(x)$
- $\text{Tame}(x) \wedge \text{Small}(x) \Rightarrow \text{Hunt}(x)$
- $\text{Vicious}(x) \wedge \text{Big}(x) \wedge \text{Hungry} \Rightarrow \text{Hunt}(x)$
- $\text{Tame}(x) \wedge \text{Small}(x) \wedge \text{Tired} \Rightarrow\sim \text{Hunt}(x)$

The predicate Vicious may not be crisp and the predicate Big may not be crisp. Hunt can be assumed to be crisp. We either Hunt or not Hunt. Hungry is also not crisp. The question is, when the predicates are fuzzy, what logic is in play here? So if our crisp rule has the form

$$\pm(C_1(x) \wedge \ldots \wedge \pm C_n(x) \Rightarrow \pm C(x)$$

we make it fuzzy by offering a rule of the form

$$(C_1(x) = e_1) \wedge \cdots \wedge (C_n(x) = e_n) \Rightarrow (C(x) = e)$$

where e_i and $e, 0 \leq e_i, e \leq 1$ are the fuzzy values.

Let us simplify and say that we have three values here for the fuzzy predicates $1 =$ yes, $0 =$ no and $\frac{1}{2} =$ not sure. This is reasonable to assume as a first approximation for our resource-poor agent. Ordinary fuzzy logic would use a function $x * y$, combining any two values x and y, which is monotonic commutative, associative and has a unit. For example we have

- $x * y = y * x$
- $x * 1 = x$
- $(x * y) * z = x * (y * z)$

This kind of function already does not work for our nonmonotonic case because of the nonmonotonicity of the rule with the Hungry predicate. We have:

- $\text{Vicious}(x) = 1 \wedge \text{Big}(x) = 1 \Rightarrow \text{Hunt}(x) = 0$

and

- $\text{Vicious}(x) = 1 \wedge \text{Big}(x) = 1 \wedge \text{Hungry} = 1 \Rightarrow \text{Hunt}(x) = 1$

The $*$ function cannot be monotonic if we want it to give this result.

The traditional method of making a system fuzzy is to take each component of the system and make it fuzzy. In the case of nonmonotonic logic we need to look at the two methods of representing a nonmonotonic consequence ($\mid\!\sim_\pi$ and $\mid\!\sim_{\mathbf{f}}$ of Section 1) and make them fuzzy.

See for example the works of Eva Richter about nonmonotonic fuzzy logic [21].

However, this 'global' approach hardly seems relevant to our poor distressed and endangered hunting agent. He really wants to know for any antecedent of his rule of the form

$$(A_1(x) = e_1) \wedge (A_2(x) = e_2) \wedge \cdots \wedge (A_n(x) = e_n)$$

What to do; to hunt or not to hunt?

This means that we cannot use a traditional norm $*$ in our model but we need a new kind of vector function V_* operating on finite sequences of numbers $\mathbf{e} = (e_1, \ldots, e_n), 0 \leq e_i \leq 1$ giving a number $0 \leq e \leq 1$ as value.

In symbols $e = V_*(\mathbf{e})$. Note that we assume that Hunt is a crisp predicate and so in this case $e = 0$ or $e = 1$.

As an example, let us define a workable system. We need to make some assumptions.

1. We divide our predicates into crisp and fuzzy predicates
2. We consider rules of the form

$$\bigwedge_i (C_i(x) = e_i) \Rightarrow C(x)$$

   where $C(x)$ is crisp and $C_i(x)$ are fuzzy.

3. For each crisp predicate C, we have a valuation function V_C giving for any finite set of elements of the form $\{C_i(x) = e_i\}$ a value $0 = \bot$ or $1 = \top$. V_C says whether the assumptions $\bigwedge_i C_i(x) = e_i$ can give us $C(x)$ as a conclusion.
4. V_C satisfies the property that for each fuzzy C_i, it is either monotonic up or monotonic down in $C_i(x)$.

   So consider the nonmonotonicity rule:

- $\text{Big}(x) \wedge \text{Vicious}(x) \Rightarrow\!\sim \text{Hunt}(x)$

Then as far as Hungry is concerned, it encourages one to hunt while Tired encourages one not Hunt.

So as far as Hunt is concerned, each fuzzy predicate either encourages hunting or discouraging hunting.[9]

---

[9]We can take a classical (non-fuzzy) nonmonotonic consequence relation $\mid\!\sim$ and require the same property of it. Namely, given any X and any A, we assume that A either encourages X or the opposite. So we exclude the possibility of items (i)–(iv) all holding: **Encouragement Condition**: For arbitrary X, A, D and E the following four conditions *cannot* hold:

(i) $D \mid\!\sim X$

(ii) $D \wedge A \not\mid\!\sim X$

(A discourages X)

(iii) $E \not\mid\!\sim X$

(iv) $E \wedge A \mid\!\sim X$

(A encourages X)

Such requirement does not turn the nonmonotonic consequence into a monotonic one (i.e. it does not entail monotonicity). We consider an example of Karl Schlechta. Let the language contain only one atomic variable p. There are two models $\mathbf{m}_1 = (p = \bot)$

Our agent looks at the fuzzy data and decides whether to conclude the target or not. This is quite intuitive.

In fuzzy logic there are uninorms which are monotonic up in one variable and down in another. So let $x \oplus y$ be a uninorm (monotonic up in x and down in y) and let $*$ be a norm. Let the rule be

$$\bigwedge_{i=1}^{m}(C_i(x) = e_i) \wedge \bigwedge_{j=1}^{n}(D_j(x) = d_j) \Rightarrow (C(x) = e)$$

where C_i are encouraging and D_j are discouraging.

Then we can calculate

$$(e_1 * \cdots * e_m) \oplus (d_1 * \cdots * d_m) = e.$$

So to turn a crisp system of rules of the form

$$\bigwedge \pm C_i(x) \wedge \bigwedge \pm D_j(x) \Rightarrow \pm C(x)$$

---

and $\mathbf{m}_2 = (p = \top)$. Let the preference relation put $\mathbf{m}_1$ below $\mathbf{m}_2$ (this is actually a circumscription model for p).

Therefore we have $A \wedge B =$ either A or B or $\bot$. We have nonmonotonicity because $\varnothing \mid\!\sim\, \sim p$ but $p \mid\!\not\sim\, \sim p$. So this is a genuinely nonmonotonic consequence. We also have that the encouragement condition holds

First consider p as a target: we have

$$\varnothing \mid\!\not\sim p, \sim p \mid\!\not\sim p, p \mid\!\sim p$$

If we add p we get

$$\sim p \wedge p \mid\!\not\sim p \text{ and } p \mid\!\sim p$$

So p encourages the target p.

Let us add $\sim p$. We get

$$\sim p \mid\!\not\sim p \text{ and } \sim p \wedge p \mid\!\not\sim p.$$

So $\sim p$ discourages the target p.

Second, consider $\sim p$ as a target. We have:

$$\varnothing \mid\!\sim\, \sim p, p \mid\!\not\sim\, \sim p, \sim p \mid\!\sim\, \sim p.$$

If we add p we get

$$p \mid\!\not\sim\, \sim p, p \wedge \sim p \mid\!\sim\, \sim p.$$

So p discourages $\sim p$.

If we add $\sim p$ we get

$$\sim p \mid\!\sim\, \sim p, p \wedge \sim p \mid\!\sim\, \sim p,$$

So $\sim p$ encourages $\sim p$.

into a fuzzy system we provide a $\oplus$ and a $*$ and decide which predicates are encouraging and which are not and turn the rule into a fuzzy rule as above. We need only ensure compatibility with the $\{0, 1\}$ values of the crisp rules. So our choice cannot be arbitrary, and we had better start with a crisp system already satisfying the encouragement condition.

Acknowledgements. For helpful discussion or correspondence, the authors are indebted to Rad Kohanzad, Karl Schlechta, Patrick Suppes and Peter Bruza, as well as our colleagues in the King's College London Logic and Computation Group, Artur d'Avila Garcez, Gabriella Pigozzi and Odinaldo Rodrigues.

References

[1] Barringer, H., D. Gabbay, and J. Woods, 'Temporal dynamics of attack and defence networks', in D. Hutter and W. Stephan, (eds.), *Mechanising Mathematical Reasoning*, in *Springer Lecture Notes in Computer Science*, Springer-Verlag, 2005, pp. 59–98.

[2] Boole, G., *The Laws of Thought*, Cambridge: Macmillan and London, Walton and Baberly, 1854. Reprinted by LaSalle, Ill: Open Court, 1952.

[3] Carlson, Gregory N., and Francis Jeffrey Pelletier, *The Generic Book,* Chicago: Chicago University Press, 1995.

[4] de Finetti, B., *Probability, Induction and Statistics*, New York: John Wiley, 1972.

[5] Frege, G., *Begriffschrift: A Formula Language, Modeled upon that of Arithmetic for Pure Thought*, in Jean van Heijenoort, (ed.), *From Frege to Gödel*, Cambridge, MA: Harvard University Press, 1967. Originally published in German, 1879.

[6] Frege, G., *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concpet of Number*, translation by J. L. Austin, Oxford: Blackwell, 1978. Originally published in German, 1884.

[7] Gabbay, Dov M., *Elementary Logics: A Procedural Perspective*, London: Prentice-Hall, 1998.

[8] Gabbay, Dov M., and John Woods, 'The new logic', *Logic Journal of the IGPL*, 9:157–190, 2001.

[9] Gabbay, Dov M., and John Woods, *Agenda Relevance: A Study in Formal Pragmatics*, volume 1 of *A Practical Logic of Cognitive Systems*, Amsterdam: North-Holland, 2003.

[10] Gabbay, Dov M., and John Woods, 'Normative models of rationality: The disutility of some approaches', *Logic Journal of IGPL*, 9:157–190, 2003.

[11] Gabbay, Dov M., and John Woods, *Seduction and Shortcuts: Fallacies in the Cognitive Economy*, Volume 3 of *A Practical Logic of Cognitive Systems*, 2007, in progress.

[12] Gabbay, Dov M., and John Woods, *The Reach of Abduction: Insight and Trial*, vol. 2 of *A Practical Logic of Cognitive Systems*, Amsterdam: North-Holland, 2005.

[13] Gabbay, Dov M., and John Woods, 'The practical turn in logic', in Dov M. Gabbay and F. Guenthner, (ed.), *Handbook of Philosophical Logic*, 2nd ed., 13, Dordrecht and Boston: Kluwer, 2005, pp. 15–122.

[14] Gärdenfors, P., and D. Makinson, 'Relations between the logic of theory change and nonmonotonic logic', in A. Fuhrmann and M. Morreau, (eds.), *The Logic of Theory Change*, Springer-Verlag, 1990, pp. 185–205.

[15] Gigerenzer, Gerd, 'Fast and frugal heuristics: The tools of bounded rationality', in D. Koehler and N. Harvey, (eds.), *The Blackwell Handbook of Judgement and Decision-Making*, Oxford: Blackwell, 2004, pp. 62–88.

[16] Hamblin, C. L., *Fallacies,* London: Methuen, 1970.

[17] Macnamara, John, *A Border Dispute: The Place of Logic in Psychology*, Cambridge, MA: MIT Press, 1986.

[18] Mill, J. S., 'A system of logic', in J. M. Robson and J. Stillinger, (eds.), *The Collected Works of John Stuart Mill*, Volumes VII an VIII, Torongo University Press, 1974. Originally published in 1843 by Longman and Green, London.

[19] Nute, D., and C. Cross, 'Conditional logic', in D. Gabbay and F. Guenthener, (eds.), *Handbook of Philosophical Logic, volume 4*, 2nd edition, Kluwer, 2002, pp. 1–98.

[20] Reiter, R., 'A logic for default reasoning', *AI Journal*, 13:81–2132, 1980.

[21] Richter, Eva, 'Nonmonotonic inference operators for fuzzy logic', Workshop "Uncertainty and Nonmonotonic Reasoning", KR2002, Toulouse, 2002.

[22] Savage, L. J., *The Foundations of Statistics*, New York: Wiley, 1954.

[23] Scriven, Michael, 'Fallacies of statistical substitution', *Argumentation,* volume 1, No. 3:333–349, 1987.

[24] Woods, John, 'Who cares about the fallacies?', in F. H. van Eemeren, Rob Grootendorst, J. Anthony Blair and Charles A. Willard, (eds.), *Argumentation Illuminated,* , Amsterdam: SicSat Press, 1992, pp. 22–48. Reprinted in *The Death of Argument* as Chapter 1.

[25] Woods, John, *The Death of Argument: Fallacies in Agent-Based Reasoning,* Dordrecht and Boston: Kluwer, 2004.

[26] Woods, John, Andrew Irvine, and Douglas Walton, *Argument: Critical Thinking, Logic and the Fallacies,* 2^{nd} revised edition, Toronto: Prentice Hall, 2004.

[27] Woods, John, R. H. Johnson, D.M. Gabbay and H. J. Ohlbach, Logic and the practical turn, in D. M. Gabbay, R. H. Johnson, H. J. Ohlbach and J. Woods, (eds.), *Handbook of the Logic of Argument and Inference: The Turn Toward the Practical*, pp. 1–39. Volume 1 of *Studies in Logic and Practical Reasoning*, North-Holland, Amsterdam, 2002.

Dov Gabbay
Department of Computer Science
King's College London
Strand
London WC2R 2LS, UK
`dov.gabbay@kcl.ac.uk`

John Woods
Department of Philosophy
University of British Columbia
1866 Main Mall
Vancouver BC V6T 1Z1, Canada
`jhwoods@interchange.ubc.ca`