# Reasoning in the presence of apparent contradiction

Paul Freund
University of Cape Town
Cape Town, South Africa
frnpau013@myuct.ac.za

## ABSTRACT

Reasoning is an integral part of the human activity. Classical logic is the default way of modelling idealized human reasoning. Yet it fails to accurately model many inferences we make in our everyday lives, inferences that we are not willing to give up. This paper examines three extensions to propositional logic that purport to better model some aspect of our inferential practices, namely the KLM approach to defeasible reasoning, the AGM approach to belief revision, and the KM approach to belief update.

## CCS CONCEPTS

• **Theory of computation → Semantics and reasoning**; *Logic.*

## KEYWORDS

nonmonotonic inference, defeasible reasoning, belief revision, belief update

## 1 INTRODUCTION

Penguins are birds. Birds fly. So penguins fly. But everyone knows that penguins do not fly. So penguins fly and do not fly. So pigs fly. Something has gone wrong here. Pigs do not fly. But the argument is classically valid.

Perhaps we could give up one of the premises, that birds fly, as penguins are birds that do not fly. But then from being told that one has seen a bird, we could not draw the conclusion that that bird is capable of flight. On the contrary, we would want to draw that conclusion; yet we would retract it if we are further informed that that bird was, in fact, a penguin. To maintain consistency in the face of apparent contradiction, our reasoning is thus nonmonotonic or defeasible, is sensitive to the addition of further information, and hence nonclassical. Typically we justify this nonmonotonicity through holding that the assertion that birds fly is shorthand for something along the lines of birds *normally* fly [16], a rule that only holds *prima facie* and thus may admit exceptions, and hold that penguins are one such exception.

The example is of course, silly, a toy example. But it is emblematic of a larger phenomenon - that human reasoning systematically deviates from the predictions of classical logic, in such a way that this cannot be blamed on lack of attention or insufficient reasoning ability [20]. So classical logic is insufficient for the task of modelling human reasoning. Indeed nonmonotonic, nonclassical, inference plays an important role in understanding inference in areas as diverse and significant as the study of ethics and the nature of scientific laws, and diagnoses of system failure [18].

While thus significant, nonmonotonicity of our inferential practices does not, and can not, fully explain how we deal with apparent contradictions in our lives. Sometimes what we believe of the world may simply be *false*, such that such falsity cannot not be explained away through further clarification as to the meaning of our beliefs. In such situations, upon coming across undeniable evidence of such falsity, the appropriate response *would* be to 'give up a premise'; to revise what we believe [1]. Our practice would then follow the classic abhorrence of contradiction. But where the classical account is silent, is in the particulars of such a process, of *how* the process of belief change takes place in the presence of new information. What is needed then, is an account of such belief change.

Using the framework of propositional logic, this paper shall survey certain technical approaches to the modelling of nonclassical inferential tasks, with a view to testing the *a priori* presuppositions of these models against actual human reasoning. Section 2 shall provide the necessary background in propositional logic. In Section 3 the KLM approach to defeasible reasoning shall be outlined. Sections 4 and 5 describe two distinct, but related, approaches to the modelling of the process of belief change, namely that of belief revision and belief update, with discussion of the appropriate postulates. Finally, the connections between these approaches shall be discussed.

## 2 BACKGROUND

Let $\mathcal{P}$ be a non-empty set, and $p, q, \ldots$ refer to members of $\mathcal{P}$, henceforth atoms. Let $\mathcal{L}$ be the set of all $\alpha$ generated by $\alpha ::= \top \mid \bot \mid p \mid \neg\alpha \mid \alpha \wedge \alpha \mid \alpha \vee \alpha \mid \alpha \rightarrow \alpha \mid \alpha \leftrightarrow \alpha$. $\mathcal{L}$ is the language of propositional logic, and its members are propositional sentences, denoted by $\alpha, \beta, \ldots$ . Let $\mathcal{U}$ refer to the set of all interpretations $u : \mathcal{P} \mapsto \{0, 1\}$, denoted by $u, v, \ldots$ . Intuitively, if an atom is assigned 1 under an interpretation this means that under that interpretation that atom is true, and if it is assigned 0, then it is false.

Given an interpretation, the interpretation is extended to all sentences of $\mathcal{L}$ by the valuation functional $V : \mathcal{U} \mapsto (\mathcal{L} \mapsto \{0, 1\})$ in the manner outlined in Table 1. We say $u \in \mathcal{U}$ satisfies $\alpha \in \mathcal{L}$ if and only if $V_u(\alpha) = 1$, and if such a $u$ exists then we say $\alpha$ is *satisfiable*. The set of interpretations that satisfy a sentence, $\alpha$, is the set of models for that sentence, denoted as $Mod(\alpha)$. The models of a set of sentences, $Mod(\mathcal{K})$, is the set of interpretations such that for each $\alpha \in \mathcal{K}$, $V_u = 1$.

### 2.1 Entailment

Consider now a set of sentences, $\mathcal{K}$, which represents what an agent knows about the world. Such a $\mathcal{K}$ is dubbed a knowledge base. Given a particular $\mathcal{K}$, intuitively we may reason from the sentences of $\mathcal{K}$ to new sentences, and thus come to know more about the world.

For example, take $\mathcal{K} = \{p, p \rightarrow q\}$, where $p$ represents the proposition that Socrates is human, and $p \rightarrow q$ represents the proposition that if Socrates is human then Socrates is mortal (the classical reading of $\alpha \rightarrow \beta$ is "if $\alpha$ then $\beta$"). Then we may infer

**Table 1: Valuation of sentences [2]**

| | |
|---|---|
| $V_u(\alpha) = 1$ | if $\alpha = \top$ |
| $V_u(\alpha) = 0$ | if $\alpha = \bot$ |
| $V_u(\alpha) = u(\alpha)$ | if $\alpha \in \mathcal{P}$ |
| $V_u(\neg\alpha) = 0$ | if $u(\alpha) = 1$ |
| $V_u(\neg\alpha) = 1$ | if $u(\alpha) = 0$ |
| $V_u(\alpha \vee \beta) = 0$ | if $u(\alpha) = 0$ and $u(\beta) = 0$ |
| $V_u(\alpha \vee \beta) = 1$ | otherwise |
| $V_u(\alpha \wedge \beta) = 1$ | if $u(\alpha) = 1$ and $u(\beta) = 1$ |
| $V_u(\alpha \wedge \beta) = 0$ | otherwise |
| $V_u(\alpha \rightarrow \beta) = 0$ | if $u(\alpha) = 1$ and $u(\beta) = 0$ |
| $V_u(\alpha \rightarrow \beta) = 1$ | otherwise |
| $V_u(\alpha \leftrightarrow \beta) = 1$ | if $u(\alpha) = u(\beta)$ |
| $V_u(\alpha \leftrightarrow \beta) = 0$ | otherwise |

**Table 2: KLM Properties**

| | |
|---|---|
| (Ref) | $\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \alpha$ |
| (LLE) | $\dfrac{\alpha \equiv \beta,\ \mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \gamma}{\mathcal{K} \approx\!\!\!\models \beta \mid\!\sim \gamma}$ |
| (RW) | $\dfrac{\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \beta,\ \beta \models \gamma}{\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \gamma}$ |
| (And) | $\dfrac{\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \beta,\ \mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \gamma}{\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \beta \wedge \gamma}$ |
| (Or) | $\dfrac{\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \gamma,\ \mathcal{K} \approx\!\!\!\models \beta \mid\!\sim \gamma}{\mathcal{K} \approx\!\!\!\models \alpha \vee \beta \mid\!\sim \gamma}$ |
| (CM) | $\dfrac{\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \beta,\ \mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \gamma}{\mathcal{K} \approx\!\!\!\models \alpha \wedge \beta \mid\!\sim \gamma}$ |
| (RM) | $\dfrac{\mathcal{K} \approx\!\!\!\models \alpha \mid\!\sim \gamma,\ \mathcal{K} \not\approx\!\!\!\models \alpha \mid\!\sim \neg\beta}{\mathcal{K} \approx\!\!\!\models \alpha \wedge \beta \mid\!\sim \gamma}$ |

$q$, that Socrates is mortal. By way of contrast, few would regard the inference that the moon is made of cheese as warranted given this $\mathcal{K}$. Whether a particular proposition follows from a knowledge base is dependent both on the contents of that knowledge base, and the meaning of that sentence.

So consider a relation $\models \subseteq P(\mathcal{L}) \times \mathcal{L}$, where $\mathcal{K} \models \alpha$ is to be interpreted as $\mathcal{K}$ entails $\alpha$. It were Tarski and Gentzen (in e.g. [23] and [8]) who popularized the idea that by studying the properties of such a relation, we may thereby come to better understand the deductive process.

One such property is truth-preservation; that $\mathcal{K} \models \alpha$ may hold if and only if whenever all the sentences in $\mathcal{K}$ are all true, that $\alpha$ be true as well. Phrased semantically, this is equivalent to $\mathcal{K} \models \alpha$ if and only if $Mod(\mathcal{K}) \subseteq Mod(\alpha)$. Clearly such a property uniquely defines a single $\models$ and this is classically considered the only notion of entailment. Thus, henceforth $\models$ shall refer solely to that classical notion. The consequence set of a knowledge base is denoted $Cn(\mathcal{K})$ and is equal to $\{\gamma : \mathcal{K} \models \gamma\}$.

## 3 DEFEASIBLE REASONING

Truth-preservation is a wonderful property for certain domains, such as mathematics, where certain knowledge is sought. But the classical account of entailment fails to accurately model human reasoning in many domains. In particular, human reasoning often fails to satisfy the following property of $\models$ [22]:

$$\textbf{(M)} \quad \frac{\mathcal{K} \models \phi}{\mathcal{K} \cup \mathcal{K}' \models \phi}$$

, where the schema is to be read as if what is on top of the horizontal line holds, then what is at the bottom of that line must also be true. (M) is the principle of monotonicity. It states the principle that once a particular $\mathcal{K} \models \phi$, adding more information to such a $\mathcal{K}$ will never result in this not being the case.

Monotonicity makes sense as a property in domains in which truth-preservation holds - if $\alpha$ is true*whenever* the members of $\mathcal{K}$ are, then this must trivially include any cases where all the members of $\mathcal{K}$ *and more* are true.

Yet often we draw inferences which while not deductively (read: classically) valid, are nonetheless rationally compelling. In such cases our reasoning is dubbed defeasible [13], and monotonicity must fail, for if our reasoning is not deductively valid we may say that $\mathcal{K}$ entails $\alpha$ yet $\mathcal{K} \cup \{\neg\alpha\}$ will be consistent. It is for this reason that such reasoning is often dubbed nonmonotonic.

The so-called KLM approach (after Kraus et. al [14]) of analysis of defeasible reasoning begins with extending propositional logic by adding the binary connective $\mid\!\sim$. Just as material implication is given by $\alpha \rightarrow \beta$, the intended reading of $\alpha \mid\!\sim \beta$ is that of defeasible implication, and is often paraphrased as $\alpha$ normally or typically implies $\beta$.

Given this extension of propositional logic the natural question arises as to the nature of defeasible *entailment*, denoted by $\approx\!\!\!\models$; in what circumstances is a defeasible implication entailed by a knowledge base consisting of statements in this extended propositional logic. Here it is commonly held that, unlike classical entailment, there are multiple acceptable notions of defeasible entailment (see e.g. [16] and [17] for two such approaches, those of *lexographic* and *rational* closure). Instead the focus is on discerning the necessary properties required of *any* acceptable notion of defeasible entailment. Table 2 summarizes the relevant properties as found in [17], although an equivalent set of properties is first discussed in [14]. What follows is a discussion of each of the properties:

- (**Ref**): the property of Reflexivity, any $\alpha$ defeasibly implies itself. Seemingly *any* notion of logical consequence, defeasible or not, should satisfy this [14].
- (**LLE**): the property of Left Logical Equivalence, which states that the defeasible implications of logically equivalent formulae are equivalent.
- (**RW**): the property of Right Weakening, states defeasible implication is closed under classical entailment. Note that this

implies that logically equivalent formulae may be swapped out for one another on the right side of $\mid\!\sim$. So (**RW**) and (**LLE**) taken together imply that the analysis is proceeding at the *knowledge level* [6], i.e. the level of meaning or semantics as opposed to syntax.

(**And**): states that defeasible implications are closed under conjunction.

(**Or**): states that if $\phi$ normally follows from each of $\alpha$ and $\beta$, then in the case that at least one of $\alpha$ or $\beta$ holds, then normally $\phi$ is implied.

(**CM**): the property of Cautious Monotonicity. A restricted form of monotonicity introduced in [5] and argued for on the following basis: if both $\beta$ and $\phi$ are normally expected given $\alpha$, then given $\alpha \wedge \beta$, $\phi$ should still be expected, as $\beta$ were expected given $\alpha$ anyway.

(**RM**): the property of Rational Monotonicity. Another restricted form of monotonicity. It states the principle that defeasible implication from $\alpha$ to $\phi$ may only be given up in the case that something normally *unexpected* (i.e. its negation is expected) given $\alpha$ were to occur with $\alpha$. In other words defeasible implications hold unless the situation is explicitly nonnormal.

## 4 BELIEF REVISION

Just as truth-preservation of entailment is a wonderful property for certain domains, but ultimately limited in its applications to modelling human reasoning, so is more generally the Tarski-Gentzen approach to the study of the deductive process, that of the study of consequence relations, similarly so limited. For in the study of consequence relations we are ultimately concerned with what can be concluded given a *fixed* knowledge base. The facts are already laid out before us, and all that remains is what may be concluded *a priori* given these facts. So where such an approach must be silent is in the dynamics of reasoning; what may be concluded in the case that a new fact is discovered within the context of an already established body of knowledge.

The central insight of the AGM approach [1] to studying such dynamics, is that they may be modelled by a binary operator $*$ between an already extent theory (a deductively closed set) which an agent believes, henceforth $\mathcal{T}$, and a new fact, where the output of that binary operator is *another* theory. The intended interpretation is that we start with all that may be concluded, model the finding of new information by the binary operation, and by studying the properties of the operator we may thereby better come to understand what may be concluded given this new information, for that is the result of the operation. However the language of the sentences used in the theories remains that of propositional logic (at least in this context: the approach here follows [10]). So in a sense the approach is both more and less general than that of the KLM approach: more general in that it models dynamics that the KLM approach cannot; less general in that the base language used is less expressive.

Similarly to the KLM approach, it is well accepted that such a $*$ need not be unique. Again, the focus is on the properties that any such $*$ need fulfill. Table 3 presents the commonly accepted properties. The guiding notion is that $*$ need respect the so-called

**Table 3: AGM Properties**

| | |
|---|---|
| (G*1) | $\mathcal{T} * \alpha = Cn(\mathcal{T} * \alpha)$ |
| (G*2) | $\mathcal{T} * \alpha \models \alpha$ |
| (G*3) | $\mathcal{T} * \alpha \subseteq Cn(\mathcal{T} \cup \{\alpha\})$ |
| (G*4) | If $\neg\alpha \notin \mathcal{T}$ then $Cn(\mathcal{T} \cup \{\alpha\}) \subseteq \mathcal{T} * \alpha$ |
| (G*5) | $\mathcal{T} * \alpha = Cn(\alpha \wedge \neg\alpha)$ only if $\models \neg\alpha$ |
| (G*6) | If $\alpha \leftrightarrow \beta$ then $\mathcal{T} * \alpha = \mathcal{T} * \beta$ |
| (G*7) | $\mathcal{T} * (\alpha \wedge \gamma) \subseteq Cn(\mathcal{T} * \alpha \cup \{\gamma\})$ |
| (G*8) | If $\neg\gamma \notin \mathcal{T} * \alpha$ then $Cn(\mathcal{T} * \alpha \cup \{\gamma\}) \subseteq \mathcal{T} * (\alpha \wedge \gamma)$ |

*principle of minimal change* or *principal of the informational economy*, as presented in [21]:

- When an agent learns $\alpha$, she should adopt a posterior belief set, $\mathcal{T}'$, such that (i) $\mathcal{T}'$ is deductively cogent, (ii) $\mathcal{T}'$ includes $\alpha$, and (iii) $\mathcal{T}'$ is the closest belief set to her prior belief set $\mathcal{T}'$, which satisfies (i) and (ii).

What follows is a discussion of the properties within the context of this principle:

(**G*1**): states that the output of the binary operation is closed under consequence, hence another theory.

(**G*2**): states that the new fact must be included or proved (these are equivalent for theories) when revising under that fact. Ensures (ii).

(**G*3**): states that $\mathcal{T} * \alpha$ is a subset of the smallest deductively closed set containing all sentences of $\mathcal{T}$ and $\alpha$. Trivially fulfilled in the case that $\mathcal{T} \cup \{\alpha\}$ is inconsistent.

(**G*4**): Taken together with (**G*3**) ensures that in the case that $\mathcal{T} \cup \{\alpha\}$ is consistent, then $\mathcal{T} * \alpha$ is the smallest deductively closed set containing both $\mathcal{T}$ and $\alpha$. These two properties are part of ensuring (iii).

(**G*5**): states that the only way to lapse into inconsistency is if we revise on something already inconsistent. Taken together with (**G*1**) ensures the requirement that $\mathcal{T}'$ is deductively cogent.

(**G*6**): similar to the KLM approach, this requirement ensures that syntax is irrelevant to the results of the operation. Note that theories are deductively closed, so equivalent theories must have identical syntax, hence this requirement only mentions the syntax of the new fact.

(**G*7**): states that the effect of revising on the conjunction of two facts is a subset of the deductive closure of revising on the first fact then adding the second. Again trivially fulfilled in the case that $\gamma$ is inconsistent with $\mathcal{T} * \alpha$.

(**G*8**): taken with (**G*7**) ensures that if $\mathcal{T} * \alpha \cup \{\gamma\}$ is consistent then $\mathcal{T} * (\alpha \wedge \gamma)$ is equal to the deductive closure of this set. Within the context of (**G*3**) and (**G*4**) implies that providing $\gamma$ is consistent with $\mathcal{T} * \alpha$ then $\mathcal{T} * (\alpha \wedge \gamma)$ equals $(\mathcal{T} * \alpha) * \gamma$. In other words, in the right conditions, revising on a conjunction is equivalent to revising on the first conjunct, and then the second. Together (**G*7**) and (**G*8**) help to fulfill (iii) [10].

## 5 BELIEF UPDATE

Within the context of belief revision we are concerned with modelling the dynamics of reasoning upon the learning of new information of an agent. It were in the database community [12] [24] that it were first noticed that depending on the context in which that new information is learned, different dynamics of reasoning may be appropriate, even for formally identical examples [11]. In particular a distinction is drawn between learning new information about an unchanging world, and learning of the world that it has undergone new changes. The former is what is associated with belief revision, and the latter is dubbed that of belief update.

To get an intuitive grasp of the distinction between belief update and revision, take the following example adapted from [11]. Let $b$ be the proposition that the book is on the table, and $m$ be the proposition that the magazine is on the table. Say that our belief set, $\mathcal{T}$, may be represented by the deductive closure of $(b \land \neg m) \lor (\neg b \land m)$, that is the book is on the table or the magazine is on the table, but not both. We send a student in to report on the state of the book. She comes back and tells us that the book is on the table, that is $b$.

Consider what the AGM approach would say here. By (**G*3**) and (**G*4**) if $\mathcal{T}$ and $b$ are consistent then adding $b$ to the set is equivalent to revising on it. The two *are* consistent so by simply adding $b$ to our set we get that $\mathcal{T} * b$ is equivalent to $Cn(b \land \neg m)$, that is the book is on the table and the magazine is not. And this seems right and correct.

But consider now that instead of asking the student to report on the state of the book we had instead asked her to ensure that the book were on the table. After reporting back that she had indeed ensured that the book is on the table, we again are faced with the new knowledge that $b$. But here it seems presumptuous to assume that the magazine is not on the table [11]. Either the book were already on the table and the magazine were not, in which case the student would have done nothing and left, or the magazine were on the table and the book not, in which case the student presumably would have simply put the book on the table and left the magazine similarly so on.

It is on the basis of such examples that Katsuno and Mendelzon [11] argue that within the context of belief update the what they dub *disjunction rule* is appropriate: the result of updating $\gamma \lor \phi$ with $\alpha$ is equivalent to the disjunction of $\gamma$ updated with $\alpha$ and $\phi$ updated with $\alpha$. The disjunction rule is inconsistent with the AGM postulates. So they postulated some new properties, ones more appropriate to belief update.

These properties are summarized in Table 4. Again, they do not single out a unique operator, rather they specify a family of admissible operators. To understand them it first has to be understood that the formal context in which Katsuno and Mendelzon are working is an adaption of the AGM approach [10]. Instead of working directly with deductively closed sets, $\mathcal{T}$, they work only with such sets that have a finite cover, that is a finite set of sentences $\mathcal{K}$ such that $Cn(\mathcal{K}) = \mathcal{T}$, and instead of working directly with that $\mathcal{K}$ they work with a representational sentence $\phi \in \mathcal{L}$ such that $\mathcal{K} = \{\gamma : \{\phi\} \models \gamma\}$. So the binary operator they are working with is a sentential operator, it takes in sentences of propositional logic and outputs a new sentence, with the new sentence standing as representative of the new knowledge base. As per usual a discussion of the postulates ends the section:

**Table 4: KM Properties**

| | |
|---|---|
| (U1) | $\phi \diamond \alpha \models \alpha$ |
| (U2) | If $\phi \models \alpha$ then $\phi \diamond \alpha = \phi$ |
| (U3) | If both $\phi$ and $\alpha$ are satisfiable then $\phi \diamond \alpha$ is satisfiable |
| (U4) | If $\phi_1 \leftrightarrow \phi_2$ and $\alpha_1 \leftrightarrow \alpha_2$ then $\phi_1 \diamond \alpha_1 \leftrightarrow \phi_2 \diamond \alpha_2$ |
| (U5) | $(\phi \diamond \alpha) \land \gamma \models \phi \diamond (\alpha \land \gamma)$ |
| (U6) | If $\phi \diamond \alpha_1 \models \alpha_2$ and $\phi \diamond \alpha_2 \models \alpha_1$ then $\phi \diamond \alpha_1 \leftrightarrow \phi \diamond \alpha_2$ |
| (U7) | If $\phi$ is complete then $(\phi \diamond \alpha_1) \land (\phi \diamond \alpha_2) \models \phi \diamond (\alpha_1 \lor \alpha_2)$ |
| (U8) | $(\phi_1 \lor \phi_2) \diamond \alpha \leftrightarrow (\phi_1 \diamond \alpha) \lor (\phi_2 \diamond \alpha)$ |

- **(U1)**: similarly to belief revision, updating with the new fact must ensure that the new fact is a consequence of the update.
- **(U2)**: states that updating on a fact that could in principle be already known has no effect. Note that a consequence of this is that for any inconsistent $\phi$, $\phi$ will remain inconsistent even with updates. This is in contrast to the case with belief revision, where (**G*5**) ensures this will not happen. This is consistent with the philosophy of belief update such that new information happens within the context of the old world *changing*, as opposed to bringing to light more information about the world - if we already start with an impossibility any changes to that impossibility will themselves be impossible [11].
- **(U3)**: states the reasonable requirement that we cannot lapse into impossibility unless we either start with it, or are directly confronted by it.
- **(U4)**: the usual requirement that syntax is irrelevant.
- **(U5)**: corresponds to (**G*8**) in the AGM approach - first updating on $\alpha$ then simply adding the new information $\gamma$ is at least as strong (i.e. entails) as updating on the conjunction of $\alpha$ and $\gamma$.
- **(U6)**: states that if updating on $\alpha_1$ entails $\alpha_2$ and if updating on $\alpha_2$ entails $\alpha_1$, then the effect of updating on either is equivalent.
- **(U7)**: applies only to complete $\phi$, that is $\phi$ which have only one model. If some situation arises from updating a complete $\phi$ on $\alpha_1$ and it also results from updating that $\phi$ from $\alpha_2$ then it must also arise from updating that $\phi$ on $\alpha_1 \lor \alpha_2$ [11].
- **(U8)**: The aforementioned disjunction rule.

## 6 DISCUSSION

Of the four outlined approaches to the study of reasoning presented here, the linkages between them are now discussed. Both the classical account of entailment and the KLM approach to defeasible reasoning share the same general approach to the study of reasoning in that they both view it as the study of consequence relations, of what follows from a fixed body of knowledge, and the associated properties of such a relation. They differ in that the defeasible approach is more permissive than the classical, both in what counts as entailment (*normally* is an easier requirement to fulfill than *always*), and in that the defeasible approach does not pick out a unique consequence relation.

By way of contrast, both belief revision and belief update are rather concerned with examining the structure of what happens when our body of knowledge is *expanded*, in that they both examine the changes such a body of knowledge undergoes when it comes into contact with new information. They differ in that (at least it is argued) belief revision is more appropriate in the case that the world our beliefs pick out is considered as unchanging, and only the information we get about it is new, while belief update is more appropriate when the information we get is seen as marking a new change in the world.

This has however, been disputed, with it being argued that what is essential to belief revision is not that the *world* is unchanging, but that the language we use to describe it is [4]. And if we consider a time-indexed language (each statement is appended to a specific time) then this seems correct - the problems raised by the book and magazine example by Katsuno and Mendelzon for belief revision would not arise if the reports were indexed to two specific times, one to the initial time which our belief set is based on, and another for a later time indicating that the student's action (or lack thereof) has taken place. It is such considerations that have led some to see belief update as most appropriate for modelling the logic of action progression [15].

There is one more set of linkages between the approaches, so far undiscussed, which are mentioned for completeness here. While concerned with two seemingly quite different topics, defeasible reasoning and belief revision share several important formal linkages. First, postulates for defeasible reasoning and postulates for belief revision may be translated from the one context to the other, in such a way that the process of doing either may be seen as "basically the same process, albeit used for two different purposes [7]." Second, and more remarkably, it can be shown that the KLM postulates correspond for defeasible reasoning correspond to the AGM postulates for belief revision, while both were developed independently. Indeed this may be seen as part of the reason for the widespread acceptance of the KLM postulates [3]. While beyond the scope of this paper, the interested reader should consult [9] for details.

There is one final matter. The central lens through which the three examined approaches have been presented is that of their relation to human reasoning. This is not uncommon in the literature, even in the artificial intelligence community the relevance of defeasible reasoning is often presented within the context of the necessity of an artificial agent having to perform reasoning tasks which more closely model human reasoning than they would if they used classical reasoning to effectively function [18]. Yet all the postulates presented so far have been postulated on an *a priori* basis; what seems reasonable to the academic in the armchair. While the postulates may be considered normative in nature, i.e. what idealized human reasoning should be in the respective domains, this seems slightly disingenuous given that the purported relevance of the topics is often based in *actual* human reasoning. So even without going so far as adopting full-blown psychologism in these domains (that there is no standard apart from that of everyday human practice) such as in [18], there seems to be space for testing empirically whether the postulates accurately model human reasoning. As far as the author is aware research in this area is still nascent, with what attempts that have been made being focused on testing whether human reasoning follows the postulates of various systems of defeasible reasoning [20] [19] [18].

# 7 CONCLUSIONS

Classical logic is the usual way of modelling human reasoning. However it is inadequate to the task of modelling many reasoning tasks in our everyday lives. The three extensions to propositional logic considered here purport to better model reasoning in some way or another. Yet they do so on an *a priori* basis. There is thus space for research to empirically test the extent to which they accurately model human reasoning.

# REFERENCES


[1] C. Alchourrón, P. Gärdenfors, and D. Makinson. 1985. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic* 50 (1985), 510–530.

[2] M. Ben-Ari. 2012. *Mathematical Logic for Computer Science.* Springer Press, New York.

[3] R. Booth, T. Meyer, and I. Varzinczak. 2013. A Propositional Typicality Logic for Extending Rational Consequence. In *Trends in Belief Revision and Argumentation Dynamics*, Eduardo Fermé and Guillermo Simari (Eds.). King's College Publications, London, 1–31.

[4] N. Friedman and J.Y. Halpern. 1999. Modeling Belief in Dynamic Systems Part II: Revision and Update. *Journal of Artificial Intelligence Research* 10, 1 (1999), 117–167.

[5] D.M. Gabbay. 1985. Theoretical Foundations for non-monotonic reasoning in expert systems. In *Proceedings of the NATO Advanced Study Institute on Logics and Models of Concurrent Systems*, R. Krzysztof (Ed.). Springer-Verlag, 439–457.

[6] P. Gärdenfors. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* MIT Press, Cambridge, Massachusetts.

[7] P. Gärdenfors and D. Makinson. 1994. Nonmonotonic inference based on expectations. *Artificial Intelligence* 65, 2 (1994), 197–245.

[8] G. Gentzen. 1932. Über die existenz unabhängiger axiomensysteme zu unendlichen satzsystemen. *Math. Ann.* 107 (1932), 329–350.

[9] S.O. Hansson. 1999. *A Textbook of Belief Dynamics: Theory Change and Database Updating.* Kluwer, Berlin.

[10] H. Katsuno and A. Mendelzon. 1991. Propositional Knowledge Base Revision and Minimal Change. *Artificial Intelligence* 3, 52 (1991), 263–294.

[11] H. Katsuno and A. Mendelzon. 1992. On the difference between updating a knowledge base and revising it. In *Belief revision*, P. Gärdenfors (Ed.). Cambridge University Press, 183–203.

[12] A.M Kellet and M. Winslett. 1985. On the use of an extended relational model to handle changing incomplete information. *IEEE Transactions on Software Engineering, SE-11* 7 (1985), 620–633.

[13] R. Koons. 2017. Defeasible Reasoning. In *Stanford Encyclopedia of Philosophy* (winter 2017 ed.), E.N. Zelta (Ed.). Metaphysics Research Lab, Stanford. https://plato.stanford.edu/archives/win2017/entries/reasoning-defeasible/

[14] S. Kraus, D. Lehmann, and M. Magidor. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44 (1990), 167–207.

[15] J. Lang. 2007. Belief update revisited. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI 07).* IJCAI Press, 1534âĂŞ1540.

[16] D. Lehmann. 1995. Another perspective on default reasoning. *Annals of Math. and Art. Int.* 15, 1 (1995), 61–82.

[17] D. Lehmann and M. Magidor. 1992. What does a conditional knowledge base entail? *Artificial Intelligence* 55 (1992), 1–60.

[18] F.J. Pelletier and R. Elio. 2005. The Case for Psychologism in Default and Inheritance Reasoning. *Synthese* 146, 2 (2005), 7–35.

[19] M. Ragni, C. Eichhorn, T. Bock, G. Kern-Isberner, and A. Ping Ping Tse. 2017. Formal Nonmonotonic Theories and Properties of Human Defeasible Reasoning. *Minds & Machines* 27 (2017), 79–117.

[20] M. Ragni, C. Eichhorn, and G. Kern-Isberner. 2016. Simulating human inferences in light of new information: A formal analysis. In *Proceedings of the Twenty-Fifth*

*International Joint Conference on Artificial Intelligence (IJCAI 16)*, S. Kambhampati (Ed.). IJCAI Press, 2604–2610.

[21] T. Shear and B. Fitelson. 2018. Two Approaches to Belief Revision. *Erkenntnis* (2018).

[22] C. Strasser and G. Antonelli. 2018. Non-monotonic Logic. In *Stanford Encyclopedia of Philosophy* (summer 2018 ed.), E.N. Zelta (Ed.). Metaphysics Research Lab, Stanford. https://plato.stanford.edu/archives/sum2018/entries/logic-nonmonotonic/

[23] A. Tarski. 1956. *On some fundamental concepts of metamathematics. [1930] Logic, Semantics, Metamathematics. Papers from 1923 to 1938, translated by J.H. Woodger. Pages 30–36.* Clarendon Press.

[24] M. Winslett. 1988. Reasoning about action using a possible models approach. In *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI 88).* AAI Press, 89–93.