

# Contemporary theories of reasoning: an analysis



**ASSIGN  
BUSTER**

## Computational and algorithmic challenges to contemporary theories of reasoning

- Kattja Madrell

Reasoning is the process of using given information to draw valid conclusions and produce new information (Goel & Dolan, 2003) based on a combination of beliefs and language of thought (Fodor, 2001).

The language of thought hypothesis proposed by Fodor (2001) states that thought and thinking occurs in a mental language; mental representations of reasoning are like sentences and this is why language of thought is sometimes also known as Mentalese (Murat 2010). Fodor (2001) admitted, however, that language of thought alone could not be used to explain reasoning; instead a combination of language of thought and a person's belief is now accepted as the basis of human reasoning. Evans, Barston, & Pollard (1983) found that a person's beliefs about the conclusion of an argument influenced whether or not they deemed that conclusion to be valid; the truth value of a conclusion was based upon its logical relationship to a belief (Goel & Dolan, 2003).

Marr's Levels of Analysis (1982) is a tri-level hypothesis that provides us with a critical framework to analyse and evaluate models of psychology thoroughly and consistently. There are three different levels; the computational level, the algorithmic level and the implementational level. In the field of cognitive psychology these levels have also been referred to as the semantic, the syntactic, and the physical (Pylyshyn, 1984).

Marr (1982) describes the three levels of analysis as the following:

<https://assignbuster.com/contemporary-theories-of-reasoning-an-analysis/>

*“ 1. Computational Theory: the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?*

*2. Representation and algorithm: How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?*

*3. Hardware implementation : How can the representation and algorithm be realized physically?”*

In other words, the computational level of analysis is concerned with what the model or system in question does and why does it do so. The algorithmic level builds upon this and analyses the way in which the system performs its computation whilst the implementational level is concerned with the way in which the system is physically implemented. Each level is a realisation of the level before it providing a more complete explanation of the system than its predecessor. This allows for the preservation of many of the properties of inter-level relationships in complex systems (McClamrock, 1991).

This essay will discuss some of the critical issues and challenges to various contemporary theories of reasoning using Marr’s levels of analysis.

Monotonic reasoning is based upon a series of logical rules. These rules are strict, rigid and cannot be altered by the addition of new information; instead this leads to the production of new beliefs (Brachman & Levesque, 2004). In the absence of justifications that would make a rule non-monotonic, we use monotonic reasoning as a default (Lakemeyer & Nebel, 1994).

For example:

A bass guitar (A) has four strings (B)

$A = B$ .

James' instrument (C) is a bass guitar (A)

$A = C$

Therefore James' instrument (C) has four strings (B)

$(C = B)$

This is an example of monotonic reasoning; the rules are consistent and based on logic and do not appear to be problematic. But what happens when we learn that James's bass guitar actually has five strings? Reasoning monotonically forces us to learn a new rule ( $A = \neg B$ ) that contradicts a rule that is already known to be true ( $A = B$ ).

The principle of contradiction proposes that statements which contract each other - such as " a bass guitar has four strings" and " a bass guitar does not have four strings" - are mutually exclusive and cannot both be true in the same sense at the same time (Whitehead & Russell, 1912). Monotonic reasoning displays a computational crisis when faced with logically contradicting information; as the rules cannot be manipulated or altered, the goal of the reasoning cannot be achieved.

As we gain new information on various things on a regular basis, it is inappropriate to reason monotonically, like in classical logic (Isaac, Szymanik

& Verbrugge, 2013), as we will not be able to incorporate any new information to our established beliefs. It stands to reason that the only appropriate time to rely on monotonic reasoning is when in a situation in which one has complete knowledge; this, however, is still risky as one may believe that they have complete knowledge of a situation as long as they are not aware of any reason or evidence to suspect otherwise, demonstrating a false belief of what is known as the Closed World Assumption, an example of non-monotonic reasoning (Etherington, 1986)

Non-monotonic reasoning is computationally more complex than monotonic reasoning; with its main forms all sharing the same level of complexity (Eiter & Gottlob, 1992). This is because the system is malleable and based on various different connections being made. Unlike in monotonic reasoning, the addition of new information that may be contradictive of beliefs already held can alter what is already known; this occurs in two main ways belief revision and belief update. Belief revision is the addition of new information into a set of old beliefs without any logical contradictions or inconsistencies; preserving as much information as possible. Belief update is the changing (or 'updating') of old beliefs to take into account any differences (Gärdenfors, 2003).

Non-monotonic reasoning leads to common-sense conclusions being drawn that are based upon the combination of both supporting evidence and the lack of contradictory evidence; Monotonic reasoning encounters problems with this due to the fact that the beliefs being reasoned about do not consider the absence of knowledge (Etherington, 1986). Non-monotonic reasoning shows a level of tautology that is not present in its monotonic

<https://assignbuster.com/contemporary-theories-of-reasoning-an-analysis/>

counterpart; as beliefs are revised or updated to incorporate new information they become harder to negate.

Take the previous example:

A bass guitar (A) has four strings (B)

$A = B$

James' instrument (C) is a bass guitar (A)

$A = C$

Therefore James' instrument (C) has four strings (B)

$(C = B)$

We now know that the bass guitar in question has 5 strings. Using non-monotonic reasoning we can now amend our initial belief that a bass guitar has four strings so that it now shows:

A bass guitar (A) usually has four strings (B) unless it does not have four strings ( $\neg B$ )

$A = B$  unless  $A = \neg B$

This example demonstrates a common display of default reasoning (Reiter, 1980); statistically most A's are B's so it is acceptable to make a general assumption based on the statistical majority. As well as making general assumptions, default reasoning is also based upon conventional and persistent assumptions, along with a lack of contradictive information

(Brachman & Levesque, 2004). Various rules of inference in non-monotonic reasoning have been proposed and explored, including circumscription (McCarthy, 1980) and negation as failure (Clark, 1978).

The closed world assumption is a form of non-monotonic reasoning based on the assumption of complete knowledge. Proposed by Reiter in 1978 the closed world assumption is described as follows:

*“ If we assume all relevant positive information is known, anything which is not known to be true must be false. Negative facts may simply be inferred from absence of positive counter parts ”* (Reiter, 1978).

To put it in other terms, if P is not provable from the knowledge base available then we must assume not P ( $\neg P$ ) (Etherington, 1986). This assumption has one major flaw; should a person not be in possession of all the relevant information, then the assumption can no longer apply. When (and only when) there is a complete and expert knowledge of the matter being reasoned about is it truly appropriate to employ the closed world assumption.

In order to prevent unwanted inferences of non-monotonic logic, such as the false belief of the closed world assumption, it is necessary to retract any assumption of complete knowledge; this leads to the use of implicit general assumptions (Brachman & Levesque, 2004). If the addition of any newly learned information is contradictive to these general assumptions, adjustments are made (Etherington, 1986) and beliefs are updated or revised (Gärdenfors, 2003).

The general assumptions made when reasoning non-monotonically are based upon normalcy obtained from knowledge and experience; we may assume that James' bass guitar has four strings as bass guitars normally do so. But what statistical probability can be assigned to an assumption to label it as 'normal' and what situational factors determine which assumptions can be made? When does a situation deem it appropriate to assume? The complexity of the ever-changing algorithms behind non-monotonic reasoning lead to different results being produced; for example, due to slight changes in situation, individual differences and varying information.

Default reasoning is arguably one of the most popular forms of non-monotonic reasoning (Reiter, 1978). Based on the principles of default logic (see Nebel, 1991; Goldszmidt & Pearl, 1996), default reasoning demonstrates a serious computational crisis known as the specificity principle. The specificity principle states that, when faced with a logical conflict, people make assumptions based more commonly upon more specific defaults than general ones (Brachman & Levesque, 2004); this can lead to stronger conclusions and, although at times, these conclusions are correct, the assumption itself that more specific defaults should be preferred is logically lacking (Brewka, 1994). In order to "make up" for this problem of specificity, one would have to overtly assign the appropriate priority levels to the defaults in regards to the situation in question.

According to the principle of contradiction proposed by Whitehead & Russell in 1912, when faced with a logical contradiction, a logical person should be able to disregard the restrictions of their system of reasoning to arrive at a logical conclusion. This however is not the case. In fact, much literature to <https://assignbuster.com/contemporary-theories-of-reasoning-an-analysis/>



date has shown human beings to behave in an illogical manner, demonstrating various logical fallacies that people reason with when using argumentation to negotiate life in a complex world (Hahn & Oaksford, 2013). A few examples of this are *ad hominem*, *ad Hitlerum* and the *slippery slope argument*.

When the character of an individual is attacked, it is suggested that any proposition they put forward should be disregarded; this is known as Ad Hominem (Hahn & Oaksford, 2013). Ad hominem is a logical fallacy that proposes that once the character or credibility of an individual has been questioned, it is no longer possible for one to have absolute confidence in what that individual says (Harris, 2012).

The term ad Hitlerum was coined by Leo Strauss in 1953; it is the name given to the logical argumentation that an idea or a view can be refuted if it is compared to one that may be held by Adolf Hitler, leader of the Nazi Party. Harris et al., in 2012, conducted a series of experiments to see whether or not participants agreed or disagreed with an opinion that may have been similar to a view shared by Hitler. They found that participants demonstrated sensitivity to probabilistic information when they were evaluating whether or not the ad Hitlerum argument was convincing. This showed that people based some of their conclusions on the origin of an argument rather than current facts.

The slippery slope argument is another logical fallacy based upon belief or assumption rather than evidence, in this case not doing something for fear of what negative consequences that action may lead to. Corner, Hahn, and

Oaksford (2011) outlined four defining components of the slippery slope argument:

- . *“ An initial proposal (A).*
- . *An undesirable outcome (C).*
- . *The belief that allowing (A) will lead to a re-evaluation of (C) in the future.*
- . *The rejection of (A) based on this belief.”*

Within beliefs in the slippery slope argument there appears to be some sort of implied mechanism which leads to the consequent action (C) directly from the antecedent action (A), even though this belief is not based upon prior knowledge nor empirical findings (Hahn & Oaksford, 2013).

These logical argumentations provide a computational challenge as, should human beings operate logically, conclusions should not be drawn based upon these fallacies however empirical evidence has shown that they frequently are (Harris et al., in 2012).

Bayes' Theorem is a formula proposed by Thomas Bayes that can be used to calculate probability in everyday reasoning (Bayes & Price, 1763). Bayesian reasoning is the process of reasoning probabilistically under uncertain circumstances when not all information is known or available (Korb & Nicholson, 2011). Using Bayes theorem, we can calculate the likelihood of different outcomes based on prior knowledge and experience of the world, assign probabilistic values and act accordingly (Oaksford & Chater, 2007).

The use of Bayesian reasoning has provided a new perspective in the analysis of psychological research; results from empirical studies have shown great deficits in human ability to reason logically (Wason, 1972). Where it would be most logical for participants to seek evidence that negated their hypothesis, they instead searched for and selected evidence that could only lead to the confirmation of their hypotheses (Hahn, Harris & Oaksford, 2013). Using Bayes Theorem, however, Oaksford & Chater (1994) demonstrated that this confirmatory response was actually the most probabilistically logical response; it involved the selection of data that provided the most information about the truth or falsity of the hypotheses (Hahn, Harris & Oaksford, 2013).

Persuasion is the process of sending a message to change a belief or incite an action. As well as its personal use, persuasion plays a major role in advertising, politics, law and many more public activities (Kamenica & Gentzkow, 2009). There are a variety of different Bayesian persuasion mechanisms, such as talk games (Crawford & Sobel, 1982), persuasion games (Milgrom & Roberts, 1986), and signalling games (Spence, 1973); Bénabou and Tirole (2004) further adapted the use of Bayesian persuasion to investigate mechanisms of self-signalling and self-regulation. Throughout all aspects of Bayesian reasoning, one thing remains constant; a person (A) can affect the actions of another (B) only by first changing the beliefs of B (Kamenica & Gentzkow, 2009).

Bayesian persuasion has been criticised in terms of its computational properties. Unlike argumentation, persuasion is concerned with what persuasive techniques work and why regardless of whether or not the

reasoning was rational (Madsen et al., 2013). Empirically, the results of study into persuasion have shown that the effects on a person's beliefs rarely persist (Cook & Flay, 1978). There is also a lack of evidence in literature demonstrating that belief change resulting from a persuasive argument produces behaviour that corresponds with the change in belief (Festinger 1964).

Bayesian reasoning shows a great deal of algorithmic complexity. The type of information being reasoned about has an effect upon the conclusions drawn with people showing greater difficulty in reasoning with conditional information than joint information (Lewis & Keren, 1999). The probability estimates for a hypothesis are frequently updated with the addition of new relevant information using Bayesian inference. Gigerenzer & Hoffrage (1995) analysed thousands of Bayesian problems and found that the adaptation of Bayes theorem using a frequency formats can be used to reduce algorithmic complexity.

Bayesian persuasion is also a very complex process, most successful persuasion of belief happens after multiple persuasion attempts over a long period of time (Kamenica & Gentzkow, 2009). Hahn and Oaksford (2013) proposed that the most influential factor of persuasion is the quality of the argument being put forward; because the quality of an argument is subject to personal opinion it provokes the question ' what makes an argument good or bad?' Human beings are not perfect Bayesians (Mullainathan, Schwartzstein & Shleifer, 2008) and while some persuasive activities may reflect a person's failures of rationality, Kamenica and Gentzkow (2009)

concluded that a complete understanding of a Bayesian persuasion is needed in order to fully assess results in literature.

Recently, psychological study has begun addressing the current issues in the computational and algorithmic levels of different types of reasoning. The effects of emotion upon the ability to reason logically have been called in to question (see Blanchette, 2013; Ayesh, 2003) as has the much greater issue of subjectivity in Bayesian reasoning (see Press, 2009; Ben-David & Ben-Eliyahu-Zohary, 2000).