Cognitive Logic versus Mathematical Logic

Pei Wang
Department of Computer and Information Sciences
Temple University
pei.wang@temple.edu
http://www.cis.temple.edu/∼pwang/

Abstract

First-order predicate logic meets many problems when used to explain or reproduce cognition and intelligence. These problems have a common nature, that is, they all exist outside mathematics, the domain for which mathematical logic was designed. Cognitive logic and mathematical logic are fundamentally different, and the former cannot be obtained by partially revising or extending the latter. A reasoning system using a cognitive logic is briefly introduced, which provides solutions to many problems in a unified manner.

1 Mathematical logic and cognition

An automatic reasoning system usually consists of the following major components:

1. *a formal language* that represents knowledge,
2. *a semantics* that defines meaning and truth value in the language,
3. *a set of inference rules* that derives new knowledge,
4. *a memory* that stores knowledge,
5. *a control mechanism* that chooses premises and rules in each step.

The first three components are usually referred to as a *logic*, or the *logical part* of the reasoning system, and the last two as an *implementation* of the logic, or the *control part* of the system.

At the present time, the most influential theory for the logic part of reasoning systems is mathematical logic, especially, first-order predicate logic. For the control part, it is the theory of computability and computational complexity. Though these theories have been very successful in many domains, their application in cognitive science and artificial intelligence shows fundamental differences from human reasoning in similar situations.

(1) Uncertainty

Traditional theories of reasoning are *certain* in several aspects, whereas actual human reasoning is often *uncertain* in these aspects.

- The meaning of a term in mathematical logic is determined according to an interpretation, therefore it does not change as the system runs. On the contrary, the meaning of a term in human mind often changes according to experience and context. Example: What is a "language"?

- In mathematical logic, the meaning of a compound term is completely determined by its "definition", which reduces its meaning into the meaning of its components and the operator (connector) that joins the components. On the contrary, the meaning of a compound term in human mind often cannot be fully reduced to that of its components, though is still related to them. Example: Is a "blackboard" exactly a black board?

- In mathematical logic, a statement is either true or false, but people often take truth values of certain statements as between true and false. Example: Is "Tomorrow will be cloudy" true or false?

- In mathematical logic, the truth value of a statement does not change over time. However, people often revise their beliefs after getting new information. Example: After learning that Tweety is a penguin, will you change some of your beliefs formed when you only know that Tweety is a bird?

- In mathematical logic, a contradiction leads to the "proof" of any arbitrary conclusion. However, the existence of a contradiction in a human mind will not make the person to do so. Example: When you experience conflicting beliefs, do you believe $1 + 1 = 3$?

- In traditional reasoning systems, inference processes follow algorithms, therefore are predictable. On the other hand, human reasoning processes are often unpredictable, and very often an inference process "jumps" in an unanticipated direction. Example: Have you ever postponed your writing plan, and waited for an "inspiration"?

- In traditional reasoning systems, how a conclusion is obtained can be accurately explained and repeated. On the contrary, the human mind often generates conclusions whose sources and paths cannot be backtracked. Example: Have you ever said "I don't know why I believe that. It's just my intuition"?

- In traditional reasoning systems, every inference process has a prespecified goal, and the process stops whenever its goal is achieved. However, though human reasoning processes are also guided by various goals, they often cannot be completely achieved. Example: Have you ever tried to find the goal of your life? When can you stop thinking about it?

(2) Non-deductive inference

All the inference rules of traditional logic are *deduction* rules, where the truth of the premises guarantees the truth of the conclusion. In a sense, in deduction the information in a conclusion is already in the premises, and the inference rule just reveals what is previously implicit. For example, from "Robins are birds" and "Birds have feather," it is valid to derive "Robins have feather."

In everyday reasoning, however, there are other inference rules, where the conclusions seem to contain information not available in the premises:

Induction produces generalizations from special cases. Example: from "Robins are birds" and "Robins have feather" to derive "Birds have feather."

Abduction produces explanations for given cases. Example: from "Birds have feather" and "Robins have feather" to derive "Robins are birds."

Analogy produces similarity-based results. Example: from "Swallows are similar to robins" and "Robins have feather" to derive "Swallows have feather."

None of the above inference rules guarantees the truth of the conclusion even when the premises are true. Therefore, they are not valid rules in traditional logic. On the other hand, these kinds of inference seem to play important roles in learning and creative thinking. If they are not valid according to traditional theories, then in what sense they are better than arbitrary guesses?

(3) Paradoxes

Traditional logic often generates conclusions that are different from what people usually do.

Sorites paradox: No one grain of wheat can be identified as making the difference between being a heap and not being a heap. Given then that one grain of wheat does not make a heap, it would seem to follow that two do not, thus three do not, and so on. In the end it would appear that no amount of wheat can make a heap.

Implication paradox: Traditional logic uses "$P \rightarrow Q$" to represent "If P, then Q". By definition the implication proposition is true if P is false or if Q is true, but "If 1+1 = 3, then the Moon is made of cheese" and "If life exists on Mars, then robins have feather" don't sound right.

Confirmation paradox: Black ravens are usually taken as positive evidence for "Ravens are black." For the same reason, non-black non-ravens should be taken as positive evidence for "Non-black things are not ravens." Since the two statements are equivalent in traditional logic, white sacks are also positive evidence for "Ravens are black," which is counter-intuitive.

Wason's selection task: Suppose that I show you four cards displaying A, B, 4, and 7, respectively, and give you the following rule to test: "If a card has a vowel on one side, then it has an even number on the other side." Which cards should you turn over in order to decide the truth value of the rule? According to traditional logic, the answer is A and 7, but people often pick A and 4.

2 Different assumptions on reasoning

None of the problems listed in the previous section is new. Actually, each of them have obtained many proposed solutions, in the form of various non-classical logics and reasoning systems. However, few of these solutions try to treat the problems altogether, but see them as separate issues.

There is a common nature of these problems: they all exist outside mathematics. At the time of Aristotle, the goal of logic was to find abstract patterns of valid inference in all domains. It remained to be the case until the time of Frege, Russell, and Whitehead, whose major interest was to set up a solid logic foundation for mathematics. For this reason, they developed a new logic to model valid inference in mathematics, typically the binary deduction processes that derives theorems from axioms and postulations.

What is the deference between a "cognitive logic" as used in everyday life and a "mathematical logic" as used in meta-mathematics? A key difference is their assumptions on whether their knowledge and resources are sufficient to solve the problems they face. On this aspect, we can distinguish three types of reasoning systems:

Pure-axiomatic systems. These systems are designed under the assumption that both knowledge and resources are sufficient. A typical example is the notion of "formal system" suggested by Hilbert (and many others), in which all answers are deduced from a set of axioms by a deterministic algorithm. The axioms and answers get their meaning by being mapped into a concrete domain using model-theoretical semantics. Such a system is built on the idea of sufficient knowledge and resources, because all relevant knowledge is assumed to be fully embedded in the axioms, and because questions have no time constraints, as long as they are answered in finite time. If a question requires information beyond the scope of the axioms, it is not the system's fault but the questioner's, so no attempt is made to allow the system to improve its capacities and to adapt to its environment.

Semi-axiomatic systems. These systems are designed under the assumption that knowledge and resources are insufficient in some, but not all, aspects. Consequently, adaptation is necessary. Most current non-classical logics fall into this category. For example, non-monotonic logics draw tentative conclusions (such as "Tweety can fly") from defaults (such as "Birds normally can fly") and facts (such as "Tweety is a bird"), and revise such conclusions when new facts (such as "Tweety is a penguin") arrive. However, in these systems, defaults and facts are usually unchangeable, and time pressure is not taken into account [Reiter, 1987]. Fuzzy logic treats categorical membership as a matter of degree, but does not accurately explain where the degree come from [Zadeh, 1965]. Many learning systems attempt to improve their behavior, but still work solely with binary logic where everything is black-and-white, and persist in always seeking *optimal* solutions of problems [Michalski, 1993]. Although some heuristic-search systems look for less-than-optimal solutions when working within time limits, they usually do not attempt to learn from experience, and do not consider possible variations of time pressure [Simon and Newell, 1958].

Non-axiomatic systems. In this kind of system, the assumption on the insufficiency of knowledge and resources is built in as the ground floor. Such a system will be described in the following.

Pure-axiomatic systems are very useful in mathematics, where the aim of study is to idealize knowledge and questions to such an extent that the revision of knowledge and the deadlines of questions can be ignored. In such situations, questions can be answered in a manner so accurate and reliable that the procedure can be reproduced by an algorithm.

We need intelligence only when no such pure-axiomatic method can be used, due to the insufficiency of knowledge and resources. Many arguments against logicist AI [Birnbaum, 1991, McDermott, 1987], symbolic AI [Dreyfus, 1992], or AI as a whole [Searle, 1980, Penrose, 1994], are actually arguments against a more restricted target: pure-axiomatic systems. These arguments are valid when they reveal many aspects of intelligence and cognition that cannot be produced by a pure-axiomatic system (though these authors do not use this term), but some of the arguments seriously mislead by taking the limitations of these systems as restricting all possible AI systems.

Outside mathematics, a system often has to work with insufficient knowledge and resources. By that, I mean the system works under the following restrictions:

Finite: The system has a constant information-processing capacity.

Real-time: All tasks have time constraints attached to them.

Open: No constraint is put on the content of the experience that the system may have, as long as they are representable in the interface language.

For a system to work under the above assumption, it should have mechanisms to handle the following situations:

A new processor is required when all processors are occupied;

Extra memory is required when all memory is already full;

A task comes up when the system is busy with something else;

A task comes up with a time constraint, so exhaustive processing is not affordable;

New knowledge conflicts with previous knowledge;

A question is presented for which no sure answer can be deduced from available knowledge;

etc., etc.

For traditional reasoning systems, these situations usually either require human intervention, or simply cause the system to reject the task or knowledge involved. However, for a non-axiomatic system, these are *normal situations*, and should be managed smoothly by the system itself. To deal with these situations, the design of a reasoning system needs to be changed fundamentally.

3 NARS overview

NARS (Non-Axiomatic Reasoning System) is an intelligent reasoning system designed to be adaptive and to work with insufficient knowledge and resources [Wang, 1995]. Here the major components of the system are briefly introduced. For detailed technical discussions, please visit my website for related publications.

Rationality under insufficient knowledge and resources

When a system has to work with insufficient knowledge and resources, what is the criteria of *validity* or *rationality*? This issue needs to be addressed, because the aim of NARS is to provide a normative model for intelligence in general, not a descriptive model of human intelligence. It means that what the system does should be "the right thing to do," that is, can be justified against certain simple and intuitively attractive principles of validity or rationality.

In traditional logic, a "valid" or "sound" inference rule is one that never derives a *false* conclusion (that is, it will by contradicted by the future experience of the system) from *true* premises. However, such a standard cannot be used in NARS, which has no way to guarantee the infallibility of its conclusions. However, this does not mean that every conclusion is equally valid.

Since NARS is an adaptive system whose behavior is determined by the assumption that future situations are similar to past situations, in NARS a "valid inference rule" is one whose conclusions are supported by evidence provided by the premises used to derive them.

Furthermore, restricted by insufficient resources, NARS cannot exhaustively check every possible conclusion to find the best conclusion for every given task. Instead, it has to settle down with the best it can find with available resources.

Experience-grounded semantics

With insufficient knowledge and resources, what relates a formal language **L**, used by a system, to the environment is not a *model*, but the system's *experience*. For a reasoning system like NARS, the experience of the system is a stream of sentences in **L**, provided by a human user or another computer. In such a situation, the basic semantic notions of "meaning" and "truth" still make sense. The system may treat terms and sentences in **L**, not solely according to their syntax (shape), but in addition taking into account their relations to the environment. Therefore, What we need is an *experience-grounded semantics*.

NARS does not (and cannot) use "true" and "false" as the only truth values of sentences. To handle conflicts in experience properly, the system needs to determine what counts as positive evidence in support of a sentence, and what counts as negative evidence against it, and in addition we need some way to measure the *amount* of evidence in terms of some fixed unit. In this way, a truth value is simply a numerical summary of available evidence. Similarly, the meaning of a term (or word) in **L** is defined by the role it plays in the experience of the system, that is, by its experienced relations with other terms.

The "experience" in NARS is represented in **L**, too. Therefore, in **L** the truth value of a sentence, or the meaning of a word, is defined by a set of sentences, also in **L**, with their own truth values and meanings — which seems to have led us into a circular definition or an infinite regress. The way out of this seeming circularity in NARS is “bootstrapping.” A simple subset of **L** is defined first, with its semantics. Then, it is used to define the semantics of the whole **L**.

As a result, the truth value of statements in NAL uniformly represents various types of uncertainty, such as randomness, fuzziness, and ignorance. The semantics specifies how to understand sentences in **L**, and provides justifications for the inference rules.

Categorical language

As said above, NARS needs a formal language in which the meaning of a term is represented by its relationship with other terms, and the truth value of a sentence is determined by available evidence. For these purposes, the notion of (positive or negative) evidence should be naturally introduced into the language. Unfortunately, the formal language used in first-order predicate logic does not satisfy the requirement, as revealed by the “Confirmation Paradox” [Hempel, 1943].

A traditional rival to predicate logic is known as *term logic*. Such logics, exemplified by Aristotle’s Syllogistic, have the following features: [Bocheński, 1970, Englebretsen, 1981]

1. Each sentence is *categorical*, in the sense that it has a *subject term* and a *predicate term*, related by a *copula* intuitively interpreted as “to be.”

2. Each inference rule is *syllogistic*, in the sense that it takes two sentences that share a common term as premises, and from them derives a conclusion in which the other two (unshared) terms are related by a copula.

In NARS, the basic form of knowledge is a statement “$S \rightarrow P$”, in which two terms are related together by an *inheritance relation*. The statement indicates that the two terms can be used as each other in certain situations. The truth value of the statement measures its evidential support obtained from the experience of the system.

Traditional term logic has been criticized for its poor expressive power. In NARS, this problem is solved by introducing various types of *compound terms* into the language, to represent *set*, *intersection* and *difference*, *product* and *image*, *statement*, and so on.

Syllogistic inference rules

The inference rules in term logic correspond to inheritance-based inference. In NARS, each statement indicates how to use one item as another one, according to the experience of the system. A typical inference rule in NARS takes two statements containing a common term as premises, and derives a conclusion between the other two terms. The truth value of the conclusion is calculated from the truth values of the premises, according to the semantics mentioned above.

Different rules correspond to different combinations of premises, and use different truth-value functions to calculate the truth value from those of the premises, justified according to the semantics of the system. The inference rules in NAL uniformly carry out *choice, revision, deduction, abduction, induction,* *exemplification, comparison, analogy, compound term formation and transformation*, and so on.

Control mechanism

NARS cannot guarantee to process every task optimally — with insufficient knowledge, the best way to carry out a task is unknown; with insufficient resources, the system cannot exhaustively try all possibilities. Since NARS still needs to try its best in this situation, the solution used in NARS is to let the items and activities in the system compete for the limited resources.

In the system, different data items (tasks, beliefs, and concepts) have different priority values attached, according to which the system's resources are distributed. These values are determined according to the past experience of the system, and are adjusted according to the change of situation. A special data structure is developed to implement a probabilistic priority queue with a limited storage. Using it, each access to an item takes roughly a constant time, and the accessibility of an item depends on its priority value. When no space is left, items with low priority will be removed. The memory of the system contains a collection of concepts, each of which is identified by a term in the formal language. Within the concept, all the tasks and beliefs that have the term as subject or predicate are collected together.

The running of NARS consists of individual inference steps. In each step, a concept is selected probabilistically (according to its priority), then a task and a belief are selected (also probabilistically), and some inference rules take the task and the belief as premises to derive new tasks and beliefs, which are added into the memory. The system runs continuously, and interacts with its environment all the time, without stopping at the beginning and ending of each task. The processing of a task is interwoven with the processing of other existing tasks, so as to give the system a dynamic and context-sensitive nature.

4 Discussion

In the following, I roughly summarize how NARS solves the problems listed at the beginning of the paper.

- In NARS, the truth value of a statement is determined by available evidence. Since evidence can be either positive or negative, and future evidence is unpredictable, the system cannot be certain about its beliefs. Depending on the source of the evidence, the uncertainty may correspond to randomness or fuzziness [Wang, 1996].

- In traditional binary logic, the truth value of a statement only depends on the existence of negative evidence, but in NARS, as in the everyday reasoning of human beings, both positive evidence and negative evidence contribute to truth value. Therefore, "Ravens are black" and "Non-black things are not ravens" are no longer equivalent, because they have different positive evidence (though the same negative evidence) [Wang, 1999]. In

Wason's selection task, to choose 4 can be explained as looking for positive evidence, and it is not always invalid [Wang, 2001].

- In NARS, the validity of non-deductive inference rules is justified according to the experience-grounded semantics. The system believes on a statement, not because it is guaranteed to be confirmed in the future, but because it has been supported in the past. "To predict the future according to the past" is the defining property of an adaptive system, though its predictions are not alway realized.

- The meaning of a term in NARS is determined by its experienced relations with other terms. Therefore, it changes according to the system's experience and context. For compound terms, their relations to their components consist only part of their meaning, and are therefore not fully compositional.

- With syllogistic rules, in each inference step the premises and conclusions must have shared terms, and so they are semantically related. Whether one statement implies another is not merely determined by their truth values, and unrelated statements will not be linked together in implication statements. A pair of conflicting beliefs will not imply everything. Instead, they only influence the semantically related statements.

- The inference processes in NARS does not follow predetermined procedures. Instead, inference steps are data driven, and are chained together at run time. The selections of concept, task, and belief in each step are determined by many factors, and are usually unpredictable and unrepeatable. When an inference process stops, usually it is not because it has reached its logical end, but because the system has no resources for it anymore.

Different from the various non-classical logics, the above solutions are produced in NARS by a unified system.

The above discussion shows that many traditional problems are caused by the application of mathematical logic outside mathematics, and these problems cannot be properly solved within the mathematical logic paradigm. A "cognitive logic" is different from a mathematical logic firstly in its assumption on the sufficiency of knowledge and resources. Because of this difference, the formal language, semantic theory, and inference rules of these two types of logic are different too, and so are the memory structure and control mechanism when they are implemented in a computer system.

One century ago, Frege changed the subject matter of logic from cognition to mathematics. Today, we feel the need of a reverse change to bring the study of logic, cognition, and intelligence into a new stage.

References

[Birnbaum, 1991] Birnbaum, L. (1991). Rigor mortis: a response to Nilsson's "Logic and artificial intelligence". *Artificial Intelligence*, 47:57–77.

[Bocheński, 1970] Bocheński, I. (1970). *A History of Formal Logic.* Chelsea Publishing Company, New York. Translated and edited by I. Thomas.

[Dreyfus, 1992] Dreyfus, H. (1992). *What Computers Still Can't Do.* MIT Press, Cambridge, Massachusetts.

[Englebretsen, 1981] Englebretsen, G. (1981). *Three Logicians.* Van Gorcum, Assen, The Netherlands.

[Hempel, 1943] Hempel, C. (1943). A purely syntactical definition of confirmation. *Journal of Symbolic Logic*, 8:122–143.

[McDermott, 1987] McDermott, D. (1987). A critique of pure reason. *Computational Intelligence*, 3:151–160.

[Michalski, 1993] Michalski, R. (1993). Inference theory of learning as a conceptual basis for multistrategy learning. *Machine Learning*, 11:111–151.

[Penrose, 1994] Penrose, R. (1994). *Shadows of the Mind.* Oxford University Press.

[Reiter, 1987] Reiter, R. (1987). Nonmonotonic reasoning. *Annual Review of Computer Science*, 2:147–186.

[Searle, 1980] Searle, J. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences*, 3:417–424.

[Simon and Newell, 1958] Simon, H. A. and Newell, A. (1958). Heuristic problem solving: the next advance in operations research. *Operations Research*, 6:1–10.

[Wang, 1995] Wang, P. (1995). *Non-Axiomatic Reasoning System: Exploring the Essence of Intelligence.* PhD thesis, Indiana University.

[Wang, 1996] Wang, P. (1996). The interpretation of fuzziness. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(4):321–326.

[Wang, 1999] Wang, P. (1999). A new approach for induction: From a non-axiomatic logical point of view. In Ju, S., Liang, Q., and Liang, B., editors, *Philosophy, Logic, and Artificial Intelligence*, pages 53–85. Zhongshan University Press.

[Wang, 2001] Wang, P. (2001). Wason's cards: what is wrong? In *Proceedings of the Third International Conference on Cognitive Science*, pages 371–375, Beijing.

[Zadeh, 1965] Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8:338–353.