Grounded on Experience: Semantics for intelligence

Pei Wang
Center for Research on Concepts and Cognition
Indiana University
510 North Fess, Bloomington, IN 47408, USA
pwang@cogsci.indiana.edu

Abstract

Model-theoretic semantics is inappropriate for adaptive systems working with insufficient knowledge and resources. An experience-grounded semantics is introduced in this paper, using NARS, an intelligent reasoning system, as a concrete example. In NARS, the truth value of a sentence indicates the amount of available evidence, and the meaning of a term indicates its experienced relationship with other terms. Accordingly, both truth value and meaning are dynamic and subjective. This approach provides new ideas to the solution of some important problems in artificial intelligence.

1 Introduction

Semantics studies how the items in a language are related to the environment in which the language is used. Concretely, semantics is the theory of *meaning* and *truth*. To ask questions like "What is the meaning of a term?" and "What is the truth value of a sentence?", what we are looking for are the *principles* that determining meaning and truth in general, rather than the meaning of a specific word or the truth of a specific sentence.

A *computerized reasoning system* often uses an artificial language to communicate with its environment, which may be a human user or another computer system. The syntax of the language is usually accurately regulated by a formal grammar. The system carries out inferences in the language according to some formal inference rules. For such a system, we need a semantics for two major reasons. When designing the system, we need to choose inference rules to get desired conclusions; when communicating with the system, we need to understand the system's language.

From the point of view of artificial intelligence, beside the general problems in semantics, we are also interested in how *intelligence* is related to semantics, that is, is there a fundamental difference in the semantics of intelligent systems and that of unintelligent systems?

In this paper, we discuss the semantics of *intelligent reasoning systems*, and use *Non-Axiomatic Reasoning System (NARS)* as a concrete example. We argue that traditional semantical theories, typically model-theoretic theory, cannot be applied to a system like NARS, because of the assumptions under which the system is designed. A new approach, *experience-grounded semantics*, is introduced, which can satisfy the requirements of NARS, and also has many interesting properties.

2 Model-Theoretic Semantics

Model-theoretic semantics, with its variations, is the dominant paradigm for semantics of formal language.

Formal languages was developed in the study of the foundation of mathematics by Leibniz, Frege, Russell, Hilbert, and so on. A basic motive behind formal language is to get ride of the uncertainties in natural language to get an objective and accurate artificial language. Model-theoretic semantics was founded by Tarski's work. Although Tarski's primary target formal language, he also hoped that the ideas could be applied to reform everyday language [17]. This approach is accepted by the "logical approach" to AI [11].

For a language **L**, defined by a finite formal grammar, a model M consists of a description of the relevant part of a domain, in another language ML, and an interpretation I, which maps the items in **L** into the items in ML.

Given the above components, the *meaning* of a term in **L** is defined as its image in M under I, and whether a sentence in **L** is *true* is determined by whether it is mapped by I into a "state of affairs" in M. For a reasoning system, valid inference rules are those that only derive true conclusions from true premises.

According to this opinion, as Tarski said [17], "semantics is a discipline which deals with certain relations between expressions of a language and the objects 'referred to' by those expressions."

Let us see what is implied by the above definitions. According to model-theoretic semantics, for any formal language, the sufficient and necessary condition for its terms to have meaning and for its sentences to have truth value is the existence of a model. In different models, meaning and truth value in the language may change, however, these changes cannot be caused by using the language. A reasoning system R that represents all its knowledge in **L** has nothing to do with the semantics of **L**. That means, on one hand, R has no *access* to the meaning of terms and truth value of sentences — it can only distinguish terms by their forms, and derive sentences from other sentences according to its inference rules, but put no constraint on how the language can be interpreted. On the other hand, what R does to the terms and sentences in **L** has no *influence* to their meaning and truth value. When working *within* such a system, as Russell said, "we never know what we are talking about, nor whether what we are saying is true," [15] unless R can set up models by itself. In that case, however, it no longer works in **L** only, and its models still cannot exclude other possible models.

These properties are good for mathematics and meta-mathematics, where abstract patterns of ideal inference are studied, and the patterns can be applied to different domains by constructing different models. The study of semantics contributes remarkably to the development of mathematics. As Tarski said, "As regards the applicability of semantics to mathematical science and their methodology, i.e., to meta-mathematics, we are in a much more favorable position than in the case of empirical sciences." [17]

However, the attempt to apply this idea to the semantic study of natural language is not successful [12]. It seems that natural language is too subtle and fluid to be put into the frame of model-theoretic semantics. Also, it hardly works for non-deductive inferences [2, 10], though there are various attempts to extend the theory into more flexible variations by introducing ideas like possible world and multi-valued logic [3, 4, 8, 23].

The problems in model-theoretic semantics are often used as arguments against "strong AI". Actually, Searle's assertion that "computers are syntactic, but the human mind is semantic" in his "Chinese room" argument [16] is directly based on the assumption that all computerized symbol manipulations are bounded to model-theoretic semantics, so uninterpreted symbols are meaningless.

Model-theoretic semantics has been criticized by many authors for its rigidness [2, 10]. However, without a powerful competitor, the solution is far from clear. As McDermott said: "The notation we use must be understandable to those using it and reading it; so it must have a semantics; so it must have a Tarskian semantics, because there is no other candidate." [10] Some people believe that it is the idea of "formalizing language and inference rules" that should be abandoned. They try some other ideas, such as neural network and robots, with the hope that they can generate meaning and truth from perception and action [2, 5].

What is the fundamental difference between natural and artificial (formal) languages? Why the latter must be rigid, constant, determined, and clear-cut? It is especially important for artificial intelligence, because here we want a computer system to use a natural language directly, or to use an artificial language in a more fluid and flexible way.

Some people may suggest that the reasoning system itself (human or computer), rather than the world it deals with, should be used as the model of the language the system uses. We can simply say that the meaning of a term is a "concept" the system has, and the truth value of a sentence is the system's "degree of belief". This idea sounds reasonable, but it does not answer the original question: how is the "concept" and "degree of belief" related to the outside world? Without an answer to that question, such a solution "simply pushes the problem of external significance from expressions to ideas." [1]

In this paper, we show another possibility: to abandon model-theoretic semantics, and to find another semantics for an intelligent reasoning system, which still use a formal language and formal inference rules.

3 Model vs. Experience

By claiming the existence of a model M, it is assumed that there is a consistent, complete, and static description of (the relevant part of) the environment in a language ML, at least in principle, and that such a description, as "state of affairs", is at least partially known, so that the truth value of some sentences in $\mathbf{L}$ can be determined accordingly. These sentences then can be used as premises for all the inferences. To insist the soundness of the inference rules, it is implied that only true conclusions (no matter how expensive they are) are desired.

When can we accept these assumptions? It is only when a system has *sufficient knowledge and resources* with respect to the problems to be solved. Sufficient knowledge means that the desired results can be got by inference from available knowledge alone, so no new knowledge is necessary; sufficient resources means that the system can afford the time-space expense of the inference, so no approximation is necessary. These are exactly the assumptions we accept when working *within an axiomatic system*. Therefore it is not a surprise that model-theoretic semantics works fine there.

Now let us see the opposite situation: a system works with *insufficient knowledge and resources*. It means that the system must have the following properties:

Finite. The system has a constant information processing capacity. As a result, it cannot be assumed that all requirements for processor time and storage space can be satisfied.

Real-time. All tasks have time requirements attached. As a result, it cannot be assumed that the system can spend as much time as it wants on a problem. It cannot be assumed, neither, that new problems only show up when the system is idle.

Open. No constraints are put on the knowledge and tasks that the system need to process, as long as they are representable in the formal language. As a result, it cannot be assumed that new knowledge will always be consistent with old knowledge. It cannot be assumed, neither, that all desired results are deductively implied by current knowledge.

It is easy to see that the human mind usually works in such an environment, but few current computer system can. In [22], it is argued that "working with insufficient knowledge and resources" is a definitive property of *intelligence*.

Model-theoretic semantics cannot be applied in such a situation. If we still define truth as "agreement with reality", so cannot be challenged by new knowledge, then no sentence can get a truth value under the above assumptions. On the other hand, the system cannot reorganize its knowledge by generating new concepts — it cannot be confirmed that these new concept really correspond to objects that "exist in the domain."

However, it does not follow that in such an environment semantic notions like "truth" and "meaning" are meaningless — if that is the case, then we cannot find truth and meaning beyond mathematics. It only means that we need a different semantics.

Semantics studies how the items in a language are related to the environment in which the language is used. With insufficient knowledge and resources, the language **L** used by a system R is related to the environment of the system, not by a *model*, but by the *experience* the system, which is the knowledge and tasks provided by the environment to the system during their interaction. To simplify our discussion, in the following we only study a system whose experience can be completely recorded as a stream of sentences in **L**.

In such a situation, the basic semantic notions like "meaning" and "truth" still make sense. The system may treat terms and sentences in **L** differently, not according to their syntax (shape), but according to their relations to the environment.

To a human designer or user, the semantic study is necessary because we want to make the system *adaptive*, that is, to behave according to its experience. For this purpose, the system need to judge the truth value of sentences according to whether, or how much, they are supported by its experience, and to distinguish the meaning of terms according to their positions in its experience.

Therefore, for an adaptive system working with insufficient knowledge and resources, a model-theoretic semantics is no longer applicable. What we need is an *experience-ground semantics*.

As descriptions of an environment, what is the difference between a "model" and a record of "experience"? At least there are the following:

- A model is a complete description of an environment, but a record of experience is only a partial description of it.
- A model must be consistent, but pieces of knowledge in experience may conflict with one another.
- A model is static, but experience extends in time.
- A model of **L** is represented in another language *ML*, and it is not necessarily accessible to the system that use **L**, but experience is represented in **L** itself, and it is accessible to the system.

4 The Semantics of NARS

In the following, we take *Non-Axiomatic Reasoning System (NARS)* [19, 21, 22] as an example, to show how to apply experience-ground semantics to a formal language used by a computerized reasoning system.

As a general-purpose intelligent reasoning system, NARS is designed to be adaptive with insufficient knowledge and resources [22]. As discussed above, in such a system the truth value of a sentence is determined by its relationship with the experience of the system, rather than with "state of affairs" in a model.

Obviously, NARS should not (and cannot) use "true" and "false" as the only truth values of sentences. To be adaptive, it is necessary to measure the amount of experience that support or against a sentence. To do it, we need to determine what is positive evidence that supports the sentence and what is negative evidence that refutes the sentence, then to measure their amount with a certain unit. In this way, a truth value is simply a numerical summary of relevant evidence.

However, as mentioned previously, in NARS "evidence" is represented in **L**, too. Therefore, the truth value of a sentence in **L** is defined by a set of sentences, also in **L**, with their own truth values — it seems to be a circular definition or infinite regression.

The solution used in NARS is "bootstrapping" — taking a subset of **L** to define the truth value of sentences and meaning of terms in **L**. In the following, we only discuss the core language of NARS, defined in [19], and ignore its extensions.

NARS uses a *term-oriented language*, in which each sentence consists of a *subject term* and a *predicate term*, related by a *copula*.

Let us first define a copula for an *ideal inheritance relation*. The copula is written as "$\sqsubset$", and the relation is defined to be *a reflexive and transitive binary relation between two terms*. Intuitively, "$\sqsubset$" correspond to the subset relation in set theory, and "to be" in English, in a highly idealized form. If "$S \sqsubset P$" is true, it means that there is no, and will no, negative evidence for the sentence in the system's experience.

According to the assumption of insufficient knowledge, this type of sentence cannot appear as knowledge in NARS. Actually, the sentences appear in the knowledge base of NARS have the form "$S \subset P$", where "$\subset$" is a copula that is reflexive and transitive *to a certain degree*. Therefore, "$\sqsubset$" is the limit case of "$\subset$", when the uncertainty in the latter completely disappears.

Now we can define evidence for a "$\subset$" relation by two "$\sqsubset$" relations.

A piece of *positive* evidence (with a unit weight) of "$S \subset P$" is a term M such that "$M \sqsubset S$" and "$M \sqsubset P$", or "$P \sqsubset M$" and "$S \sqsubset M$".

A piece of *negative* evidence (with a unit weight) of "$S \subset P$" is a term M such that "$M \sqsubset S$" but "$M \not\sqsubset P$", or "$P \sqsubset M$" but "$S \not\sqsubset M$".

The intuition behind above definition is clear: for positive evidence, the proposed transitivity holds when M is checked as an instance of the relation; for negative evidence, the proposed transitivity fails when M is checked as an instance of the relation.

Given the number of positive instances w^+ (call it the weight of positive evidence) and the number of all checked instances w (call it the weight of available evidence), the truth value of "$S \subset P$" can be naturally represented by a pair of real numbers $<f, c>$, where $f = w^+/w$, and $c = w/(w+1)$ [19]. Here f is the *frequency*, or *proportion*, of positive evidence among all evidence, and c is called *confidence*, indicating the amount of relevant evidence that the system has collected. (See [19] for a detailed discussion about confidence.)

Especially, "$S \sqsubset P$" is identical to "$S \subset P <1, 1>$", that is, "$\sqsubset$" is the limit of "$\subset$" when both w and w^+ go to infinite, while $w - w^+$ (the amount of negative evidence) has an upper bound.

Let us define L_0 as a language in which each sentence has the form "$x \sqsubset y$", where x and y are different terms. Assuming that the experience of the system (until a certain instant) is represented by K, a finite set of sentences in L_0, then it is easy to generate the reflexive and transitive closure K^* (see [19]). Based on K^*, we define an *extension* and an *intension* for each term in K^*:

The *extension* of a term T is a set of terms $E_T = \{x \mid x \sqsubset T \in K^*\}$.

The *intension* of a term T is a set of terms $I_T = \{x \mid T \sqsubset x \in K^*\}$.

The meaning a term is determined by its extension and intension, that is, its relations with other terms, according to the system's experience. Such a definition matches our intuition: the meaning of a word is revealed by its *instances* and *properties*, both represented by other terms.

As discussed previously, for any sentence "$S \subset P$" in **L**, its truth value can be derived from the weights of evidence w^+ and w^-, which can be got from the extensions and intensions of the two terms:

$$w^+ = |E_S \cap E_P| + |I_P \cap I_S|,$$

$$w^- = |E_S - E_P| + |I_P - I_S|,$$

$$w = w^+ + w^- = |E_S| + |I_P|.$$

Now we have finished our basic task. We define the truth value of sentences and meaning of terms in **L** by the system's experience, which is represented in L_0, a subset of **L**. Because the meaning of "$\sqsubset$" is completely determined by its two properties, reflexivity and transitivity, the relation is used as a semantic primitive to define other notions.

According to the assumption of insufficient knowledge, in NARS the confidence of a sentence cannot reach 1 (which means the system has infinite evidence about the sentence), but can approach it as a limit. Therefore, sentences like "$S \sqsubset P$" cannot really appears in the system's experience. However, it does not prevent us from using them to construct an "ideal experience" for semantic purposes. For example, if there is a sentence within the system's knowledge base with the form "$S \subset P$ $< 0.75, 0.80 >$", then from the relationship between truth value and weight of evidence, we get $w = 4$, $w^+ = 3$. Therefore, the system believes the relation "$S \subset P$" to such an extent, as it has tested the relation four times (by checking common extension or intension of the two terms), and the relation is confirmed three times, but fails once. By saying so, it does not mean that the system actually got the truth value by carrying out the testings — this "ideal evidence" cannot be got practically. The system may have checked the relation for more than four times, or the conclusion was derived from other knowledge or even directly provided by the environment. No matter how the truth value $<0.75, 0.80>$ is practically *generated* (there are infinite possibilities to get it), it can always be *understood* in a unique way, as stated above. The "ideal experience" is used here as an "ideal meterstick" to measure degree of truth [7].

Another factor that makes actual experience different from ideal experience is the insufficiency of resources. Due to the lack of space, some knowledge in the experience is forgot by the system; due to the lack of time, some knowledge in the experience is ignored by the system. Consequently, the truth value of a sentence or the explicit meaning of a term (i.e., its revealed relations with other terms) is usually based on *partial experience* of the system. As discussed above, this factor makes the real situation much more complex than the ideal situation, but it does not prevent us from saying that the truth value of a sentence indicates its summarized evidence, and the meaning of a term indicates its experienced relations with other terms.

5 Truth value in NARS

Now let us discuss the properties and implications of the new semantics.

One important character of the theory is its dynamic and subjective nature. Obviously, the truth values of a sentence changes dynamically in NARS, due to the coming of new experience. The system's inference activity also change truth value of sentences by combining evidence from different sections of the experience. Since truth values are based on the system's experience, they are intrinsically *subjective*. Accurately speaking, the knowledge in the system is not a description of the world, but a summary of the experience, so it is the *system's point of view* of its environment. It is highly possible that systems in the same environment have different knowledge, due to their different individual experience.

To say that truth values are dynamic and subjective, it does not mean that they are arbitrary. As Quine said about the human mind, "Observations are the boundary conditions of a system of beliefs." [14] The systems in the same environment can achieve certain "objectivity" by communicating to one another to share experience. However, here "objective" means "common" or "unbiased", rather than "observer-independent."

Such a semantics provides a justification for non-deductive inferences. As revealed by Hume's "induction problem", our predication about future experience cannot be infallible [6]. From limited past experience, we cannot get general descriptions of "state of affairs", neither can we know how far our current knowledge is from such an "objective" descriptions. Based on this, Popper made the well-known conclusion that an inductive logic is impossible [13]. However, from the previous discussion, we can see that what really pointed out by Hume and Popper is the impossibility of an inductive logic *with a model-theoretic semantics.*

Let us see how induction is justified in NARS by considering the following ideal experience: the system gets the knowledge that "$swan \subset bird <1, 1>$" ("Swans are birds.") and "$swan \subset white\text{-}thing < 1, 1 >$" ("Swans are white."). According to previous discussions, here we find a common instance of "bird" and "white-thing", which can be generalized as "$bird \subset white\text{-}thing < 1, 1/2 >$" ("Birds are white."), that is, the inductive conclusion is supported by positive evidence with unit weight. The frequency of the conclusion, 1, means that all evidence summarized by this sentence are positive; the confidence, 1/2, indicates the amount of evidence collected. They do not measure how many birds "in the world" are white. If in another section of experience the system knows that crows are birds but they are not white, it similarly gets an inductive conclusion, "$bird \subset white\text{-}thing$ $<0,\ 1/2>$," from that piece of evidence alone. When these two conflicting conclusions meet, the evidence are combined by the system's revision rule, and the system get a summarized conclusion "$bird \subset white\text{-}thing$ $<1/2,\ 2/3>$", where the frequency is a weighted sum of the competing two, and the confidence is higher than the premises', due to the accumulation of evidence.

Though the above examples are insufficient to uniquely determine the induction rule and the revision rule in NARS (the general situation is discussed in [19]), they do provide boundary conditions that the rules should satisfy. Similar analysis can be done to other non-deductive inferences, such as abduction and analogy. In this way, the validity of the inference rules in NARS are justified. These rules are not truth-preserving in the traditional sense: the conclusions may conflict with new evidence; however, they are truth-preserving according to the new definition of truth value, because the truth value of the conclusion is determined by the experience summarized in the premises.

Another result of such an experience-based definition of truth value is that it provides a unified representation for the various types of uncertainty. As discussed in [20], in a sentence "$S \subset P$," *randomness* usually happens when the *extension* of the *subject* term S is partially included in the extension of the predicate term P (some, but not all, instances of S are P), and *fuzziness* usually happens when the *intension* of the *predicate* term P is partially included in the intension of the subset term S (S has some, but not all, properties of P). On the other hand, *ignorance*, revealed by the phenomenon that judgments have different sensibility to new evidence, can be measured by "lack of confidence" in NARS, so become a function of available evidence [19]. Although these types of uncertainty have different origin, in NARS they are all represented by the truth value of sentences, and processed in a unified manner. Moreover, defining a truth value by a set of binary relations in a section of "ideal experience", we can explain a multi-valued statement about randomness, fuzziness, ignorance, or their mixture, by translating it into a set of two-valued statements. This is exactly what measure theory asks us to do [7].

6 Meaning in NARS

Like truth values, the meaning of terms in NARS is also dynamic and subjective. The meaning of a term is determined by its experienced relations with other terms, and it determines how the term will be used by the system in future. An human observer can still interpret the terms appearing in NARS freely by identifying them with words in a natural language or human concepts, but that is their meaning *to the interpreter*, and has nothing to do with the system itself. For example, if the term "bird" never appears in the system's experience, it is meaningless to the system (though meaningful to English speakers). However, when a sentence "$bird \subset animal < 1, 0.8 >$" appears in the system's input stream, the term "bird" begins to have meaning to the system, revealed by its inheritance relation with "animal". As the system know more about "bird", its meaning becomes richer and more complex. The term "bird" may never means the same to NARS as it means to a human (because we cannot expect a computer system to have human experience [22]), but we cannot say "bird" is meaningless to the system for this (human chauvinistic) reason.

This leads us to Searle's "Chinese room" argument [16] and Harnad's "symbol grounding" problem [5]. As mentioned previously, Searle's argument is based on the assumption that a symbol can only get meaning from a model. If we accept an experience-grounded semantics, it is no longer the case. As long as a term has experienced relations with other terms, it become meaningful to the system, no matter how poor its meaning is. An adaptive system never process a term only according to its shape without considering its position in the system's experience. The shape of a term may be more or less arbitrary, but not is its experienced relations with other terms.

By saying so, we do not mean that a word in a natural language get its meaning *only* by its relation with other words in the language, because *human experience* does not consist of words only. However, the general principle is still applicable here, that is, a word gets its meaning by its experienced relations with the system's other *components*, which may be words, perceptive images, motor sequences, and so on. In a system like this, the meaning of a word is much more complex than in a system whose experience consists of symbols only, but it does not rule out the latter case as a possible way for symbols to be meaningful.

The feeling of meaningless in Searle's "Chinese room" comes from his deliberate isolation of his experience in Chinese from his sensory-motor experience and his experience summarized in his native language. If we put an intelligent computer system into the same situation, there are two possible cases. If the computer system already has profound sensory-motor experience and/or a "native language", it may also judge the Chinese characters as meaningless, because it cannot relate them to its previous experience. However, if the system enters the room with no previous experience, Chinese will become its native language. The system gets the meaning of the characters by how they are related to one another, and it will not attempt to ground them on some more fundamental stuff, or complain about the "meaningless squiggles and squoggles" when it fails in doing so. If the system also has sensory-motor ability, and communicate with other similar computer systems in Chinese, we may find that the meaning of Chinese words, to them, become as rich and complex as to human Chinese speakers, though it is fairly possible that they may have different opinions about what is the "correct" meaning of a word.

If experience-grounded semantics is applied to a symbolic system, all the symbols that the system have are already *grounded* — in the system's experience, of course. The crucial point here is that for a symbol to be meaningful (or grounded), it must be related somehow to the environment. However, such a relation is not necessarily by sensory-motor mechanism. The experience of a system can be *symbolic*, as in the case of NARS. This type of experience is much simpler and "coarse-grained" than sensory-motor experience, but it *is* real experience, so can ground the symbols appear in it, just as words in natural language are grounded in human experience.

This idea sounds like what Harnad calls "dictionary-go-round" — he hopes that meaning of symbols can "be grounded in something other than just more meaningless symbols." [5] Here we should notice a subtle difference: in experience-grounded semantics, the meaning of a term is not *reduced into the meaning of other terms* (that will lead to circular definition in a finite language), but *defined by its relations with other terms*. These relations are formed during the interaction between a system and its environment, and is not arbitrary at all. On the other hand, extending the system's experience to include sensory-motor activities does not fundamentally change the situation. Sensory-motor primitives are still components of a system's experience, rather than components of the "outside world". The meaning they have (to the system) comes from their internal relationship, too.

Human beings judge the truth value of a sentence according to personal experience and determine the meaning of a word according to its relations with other words. This is not a new idea to psychologists and linguistics [9, 12, 18]. However, few people tried to apply it to an artificial language defined by a formal grammar. This is caused by several assumptions, which, though seldom mentioned, are accepted by many people.

It is implicitly assumed that the semantics of a "formal language" has to be model-theoretic. Such an inductive conclusion is valid according to our experience — almost all formal language in history get their meaning and truth values in that way. As a result, people who does not like the semantics usually abandon the language at the same time. Actually, a language can be "formal" in two different senses. In a week sense, it means that the language is artificial, and formed according to a formal grammar; in a strong sense, it means that the language is also used with a model-theoretic semantics. The language used in NARS is "formal" in the week sense only.

Logicians, in distinguishing themselves from other people, such as psychologists, tend to stress the *normative* nature of logical theory. As a result, in the study of semantics, the goal is often set as to look for *the* real, objective meaning of terms or truth value of sentences. Even if we can still justify such an opinion when the object is to study the "logic in mathematics," we cannot do it when we study the "logic in empirical science and common sense."

7 Conclusion

As discussed above, there are two different ways to relate the items of a language to the world: by the interpreting them in a model or by locating them in the experience of a system.

Model-theoretic semantics is good for axiomatic systems. However, it is not appropriate for adaptive systems working with insufficient knowledge and resources. What makes model-theoretic semantics inapplicable is not the language, but the condition under which the language is used.

Experience-grounded semantics takes the truth value of a sentence as a summary of relevant evidence, and takes the meaning of a term as its experienced relations with other terms. Many important and interesting results follow from this theory. With such a semantics, a "formal" language can also be vague and ambiguous, as well as creative and flexible, as a natural language.

A predictable suggestion to this new theory is that it is better to leave "truth value" and "meaning" with their model-theoretic usages, and to replace them with names like "degree of belief" and "association" in the new theory. Such a name substitution does not change the contents of the new theory, and it can prevent certain misunderstanding. However, unless we all agree that "truth value" and "meaning" can only be used in mathematics, the proposed substitution is not very attractive, because many problems about these semantic notions can be reconsidered in the light of the new theory.

The semantic theory introduced in this paper can be extended into more complex systems. The term-oriented language defined previously can be extended to including other inheritance relations and compound terms [21]. In a system with sensory-motor capacity, truth value and meaning are no longer only determined within the language, but also determined by the "non-verbal" components in the experience. In a system that can generate new terms by itself, these terms will correspond to the stable patterns appearing in the system's experience, rather than to the objects existing in the outside world.

Undoubtly, this paper only addresses the basic principles of experience-grounded semantics, and leave many issues untouched. Even so, we can say, from our previous work, that this approach is promising.

Acknowledgment

This work is supported by a research assistantship from the Center for Research on Concepts and Cognition, Indiana University.

References

[1] J. Barwise and J. Perry. *Situations and Attitudes*. The MIT Press, Cambridge, Massachusetts, 1983.

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

[3] R. Carnap. *Logical Foundations of Probability*. The University of Chicago Press, Chicago, 1950.

[4] J. Halpern. An analysis of first-order logics of probability. *Artificial Intelligence*, 46:311–350, 1990.

[5] S. Hardad. The symbol grounding problem. *Physica D*, 42:335–346, 1990.

[6] Hume. *An enquiry concerning human understanding*. London, 1748.

[7] D. Krantz. From indices to mappings: The representational approach to measurement. In D. Brown and J. Smith, editors, *Frontiers of Mathematical Psychology: Essays in Honor of Clyde Coombs*, Recent Research in Psychology, chapter 1. Springer-Verlag, Berlin, Germany, 1991.

[8] H. Kyburg. Semantics for probabilistic inference. In *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*, pages 142–148, 1992.

[9] G. Lakoff. Cognitive semantics. In U. Eco, M. Santambrogio, and Violi P., editors, *Meaning and Mental Representation*, pages 119–154. Indiana University Press, Bloomington, Indiana, 1988.

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

[11] N. Nilsson. Logic and artificial intelligence. *Artificial Intelligence*, 47:31–56, 1991.

[12] F. Palmer. *Semantics.* Cambridge University Press, New York, 2 edition, 1981.

[13] K. Popper. *The logic of Scientific Discovery.* Hutchinson, London, 1968.

[14] W. V. Quine and J. Ullian. *The Web of Belief.* Random House, New York, 1970.

[15] B. Russell. Recent work on the principles of mathematics. *International Monthly*, 4:84, 1901.

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

[17] A. Tarski. The semantic conception of truth. *Philosophy and Phenomenological Research*, 4:341–375, 1944.

[18] A. Tversky and D. Kahneman. Judgment under uncertainty: heuristics and biases. *Science*, 185:1124–1131, 1974.

[19] P. Wang. From inheritance relation to non-axiomatic logic. *International Journal of Approximate Reasoning.* Accepted in June 1994.

[20] P. Wang. The interpretation of fuzziness. *IEEE Transactions on Systems, Man, and Cybernetics.* Accepted in October 1994.

[21] P. Wang. Non-axiomatic reasoning system (version 2.2). Technical Report 75, Center for Research on Concepts and Cognition, Indiana University, Bloomington, Indiana, 1993. Available via WWW at *http://www.cogsci.indiana.edu/farg/pwang_papers.html.*

[22] P. Wang. On the working definition of intelligence. Technical Report 94, Center for Research on Concepts and Cognition, Indiana University, Bloomington, Indiana, 1994. Available via WWW at *http://www.cogsci.indiana.edu/farg/pwang_papers.html.*

[23] L. Zadeh. Test-score semantics as a basis for a computational approach to the representation of meaning. *Literary and Linguistic Computing*, 1:24–35, 1986.