

Grounding the Meaning of Symbols on the System's Experience

Pei Wang

IntelliGenesis Corporation

and

Center for Research on Concepts and Cognition, Indiana University

Homepage: www.cogsci.indiana.edu/farg/pwang.html

E-mail: pwang@cogsci.indiana.edu

NARS is an intelligent reasoning system, whose interaction with its environment can be described as a stream of input sentences in a formally defined language and a stream of output sentences in the same language. These two streams are called the system's "experience" and "responses", respectively (Wang, 1994; Wang, 1995a; Wang, 1995b). (Detailed description of the system can be found from the author's homepage.)

Each sentence in the language represents an inheritance relation between two terms. By definition, a sentence " $S \subset P$ " indicates that the subject term " S " is in the extension of the predicate term " P ", and " P " is in the intension of " S ". Because the relation " \subset " is defined to be reflexive and transitive, " $S \subset P$ " also indicates that " S " inherits the intension of " P ", and " P " inherits the extension of " S ". Intuitively, the subject is a specialization of the predicate, while the predicate is a generalization of the subject.

Besides complete inheritance relations, NARS also represents partial inheritance relations by attaching a truth value to each " \subset " to numerically measure the extent to which the proposed inheritance relation is confirmed/refuted by the system's experience.

Based on the above definition of truth value, a set of inference rules are established in NARS to derive new inheritance relations from experienced relations. There are rules for deduction, abduction, induction, revision, analogy, compound term formation, and so on, and they correspond to different ways to get inheritance relations.

Intuitively, the knowledge base of NARS can be viewed as a symbolic network, with the terms as nodes, and the inheritance relations as links. The experience of the system consists of nodes and links directly formed during the interaction between the system and its environment, and the system's inference activity adds derived links and nodes into the network.

In NARS, the meaning of a term is its extension and intension. In other works, it is the total (directly

or indirectly experienced) relations between this term and the other terms.

NARS is designed to work in real-world situations where the system's knowledge and (time-space) resources are generally insufficient to solve the problem imposed on it by its environment. In these situations, what we can expect from a system (either a human or a computer) is not to provide perfect solutions, but to give reasonable ones in which the system's available knowledge and resources are used as efficiently as possible.

With insufficient resources, it is impossible for the system to exhaust all implications of its direct experience. Instead, the system prioritizes the tasks to be fulfilled and the knowledge to be used, to spend more resources on the urgent and promising tasks, and to give useful knowledge (proven in the past) more chance to be used again in the future.

This means that the system cannot know everything about a term — it only know what it has experienced about the term and what it can derive from its relevant experience. Furthermore, when a term is involved in a problem-solving process, usually only part of the knowledge the system has about the term is actually used.

According to model-theoretic semantics, the meaning of a term in a language is determined by an "interpretation", which provides an isomorphism between the language and a domain in which the language is used. Consequently, the meaning of a term has nothing to do with the system's activity, and is not "grounded" on anything of the system itself.

In the contrary, in NARS the meaning of a term or the truth value of a statement is grounded on the experience of the system. In this way, they are no longer objective and constant, but subjective and dynamic — they are determined by the individual history of the system, rather than something in the outside world.

This does not mean that meaning and truth become arbitrary. If two systems share the same environment

and communicate to each other, their experience will overlap, and they will have common (though not identical) opinions. Within a system, though new experience and the system's inference activity constantly changing the meaning of terms, certain relations will become relatively stable.

This theory can be extended into systems that have sensory-motor capacity, such as humans, animals, and robots. We can still view the knowledge base of such a system as a network, though here the nodes can also correspond to sensors or operators of the system. Even though, we can still say that the knowledge of the system are links formed either during the interaction between the system and the environment or during the information-processing activity within the system, and it is these links that defines the meaning of the concepts in the system.

References

Pei Wang. From inheritance relation to nonaxiomatic logic. *International Journal of Approximate Reasoning*, 11(4):281–319, November 1994.

Pei Wang. Grounded on Experience: Semantics for Intelligence, CRCC Technical Report No. 96, 1995.

Pei Wang. *Non-Axiomatic Reasoning System: Exploring the Essence of Intelligence*. PhD thesis, Indiana University, 1995.