Understanding as Conceptualizing

Pei Wang

Department of Computer and Information Sciences
Temple University, Philadelphia, USA
pei.wang@temple.edu

September 27, 2017

Abstract. Understanding, as a cognitive function, can be studied in the context of either communication or general cognition. In both contexts, it is possible to understand the process as the forming and reforming of concepts. The major missing components in the current AI techniques include experience-grounded semantics and resource-bounded spontaneous inference. The relevant aspects of an AGI system, NARS, are discussed as an example on how understanding can be achieved in a general-purpose system.

1 What is Understanding

Like many basic notions in cognitive science and artificial intelligence, "understanding" has no commonly accepted definition, though researchers agree that it is important. In AI study, at least two senses of "understanding" can be discussed. In a narrow sense, a system understands a message by getting its intended meaning; in a broad sense, a system understands an object, event, or situation by getting the necessary information to properly deal with it.

The previous works in AI started at the narrow sense, as in "natural language understanding" [13, 5] and "speech understanding" [1]. In recent years, related topics extend to include "text understanding" [16], "image understanding" [3], as well as to the related topics, such as "machine translation" [15]. The works involving the broad sense of understanding include various approaches to model an object or event [7].

The AI techniques on "understanding" roughly divide into two groups: the old projects typically use symbolic representation and binary logic, while the recent projects mostly use deep neural networks and other statistical methods. These two schools conceptualize the understanding process differently. In the symbolic tradition, understanding is seen as mainly consisting of a mapping from input message to an internal semantic representation, followed by an inference process that reveals the intended meaning of the message. In the neural network tradition, understanding is seen as an input-output mapping to be learned from training samples.

From an AGI perspective, all previous treatments of understanding have major weakness.

Though deep learning has achieved remarkable results in recent years, it still cannot be considered as a proper solution of the "understanding" problem, mainly because it takes "learning" as function approximation or optimization [12], so every speech, text, or image is mapped into a "correct" description. This treatment has the following limitations:

- In communication, the literal meaning of a message is often not the intended meaning of the sender, and that is why *inference* is needed for understanding to combine the message with the receiver's relevant knowledge to find out the intended meaning.
- To understand something usually means to be able to solve various problems about it (such as description, prediction, explanation, manipulation), and that is why its internal *representation* should have a relatively clear meaning. This is beyond what the hidden layers in a neural network can naturally offer.
- The result of understanding is usually *context-sensitive*, and here "context" includes not only the input that immediately precedes and follows the message to be understood, but also the current internal and external circumstances (e.g., the receiver's goals and anticipations).

The above issues suggest that "understanding" cannot be properly handled as an end-to-end mapping as assumed in neural networks, deep or shallow, because there is no unique understanding of something that can be learned as a function from sample data. On these issues, the symbolic approaches may actually be more proper, as semantic representation (the "interlingual" in [4]) and inference have been considered as necessary for understanding from the very beginning [13, 5].

Of course, symbolic approaches towards understanding have their own problems, and some of them are well-known:

- For symbolic representation to have meaning, the symbols must be somehow related to the environment. This is the "symbol grounding" problem [2]. The traditional (model-theoretic) semantics requires an interpretation to be provided for the symbols to get meaning, so such a system does not really understand the symbols, as pointed out by Searle in the "Chinese Room" thought experiment [6].
- Since the previous and current AI projects are usually special-purpose, the inference involved in understanding is oriented to a small number of predetermined tasks in each system. In AGI systems, on the contrary, result of understanding should serve many task-specific inference. For example, if a message is merely used as a command to start a process, it would be far-fetched to claim for full understanding.
- The above "general-purpose inference" cannot be implemented by exhaustively generating all possible implications of the existing knowledge or "facts about the world" [14], as that will usually lead to combinatorial explosion, and surely won't provide real-time responses.

Therefore, for AGI we need a form of understanding that is closer to what happens in the human mind than achieved by the mainstream AI techniques, both from the symbolic school and the neural net school.

Usually, a communication process is considered as successful if the receiver "understands" the message from the sender. However, what is understanding? In previous research, it is often assume that the "message" has a "true meaning", represented in certain way, and if the receiver converts the message into that form, it is considered as understood. Very often, this process is formally described using Shannon's Information Theory, where the information carried by a message is the uncertainty it reduced in the receiver's mind.

According to the previous description about the communication process, this understanding of "understanding" is too rough an approximation for our purpose, though it may be good enough for some other purposes. Since communication is an goal-oriented activity for both the sender and the receiver, whether the process is successful should be judged by whether they achieved their goals, respectively.

For the sender, the direct goal of the communication is to present a task to the receiver, to be processed by the latter in an expected way. In the simplest case, the goal is indeed to add a task into the receiver's memory as a literal translation of the meaning of the message. However, very often it is not the case. Instead, the expected result is a derived task that come from the literal meaning of the message, plus certain beliefs in the mind of the receiver. This explains why an incomplete sentence or a question is often received as a meaningful message. Therefore, for an outgoing message to be properly understood, the sender's expectation about how the message will be processed should match what actually happen in the receiver's mind, otherwise it is misunderstanding, from the sender's point of view.

The receiver's goal in the communication is often not exactly the same as the sender's, though usually related to it. In the simplest case, the receiver just tries to estimate the sender's goal from the message, and forms tasks accordingly, or takes certain immediate actions. As far as the receiver has a strong belief on its estimation about the goal of the sender, the receiver considers the message "understood". Similarly, misunderstanding can happen when the receiver's beliefs about the sender do not match what happened in the mind of the sender.

As soon as the experience of the sender and the receiver become complex enough, "perfect understanding" is nearly impossible, and whether a message is "understood" becomes a matter of degree. During real-time multi-round communication (such as dialog), some misunderstandings can be recognized and corrected by the following communication, but it will be much harder if the communication is one-way (such as broadcasting) or delayed (such as reading). However, a misunderstanding does not means a failed communication, especially from the view point of the receiver. Very often a receiver achieves important goals by process the received message in a way that is completely different from what the sender expected, though still triggered by the message. Therefore, the "intended meaning" (of the sender) is not necessarily the "obtained meaning" (of the receiver), and the former is not always "more valuable" to the receiver.

"Don't understand", on the other hand, is indeed a failure of communication, which is very different from misunderstanding. Here the receiver fails to achieve any goal from the received message.

Outside communication, people often distinguish "no understanding", "shallow understanding", "deep understanding" as descriptions of one's knowledge about an object, according to whether one can successfully deal with it (prediction, manipulation, etc.), though in this situation there is no objective standard, so understanding is a matter of degree.

2 How to Achieve Understanding

The above analysis shows that to achieve human-like "understanding", several cognitive functions are needed. In this paper I only discuss the relevant aspects in semantics and inference in the AGI system NARS [10].

As a reasoning system, NARS is equipped with a symbolic logic, which uses a symbolic language for knowledge representation. However, unlike the traditional "symbolic systems", the words and statements in NARS do not get their meaning via an "interpretation" that maps them into physical objects or Platonic notions outside the system. Instead, their meaning is determined by their mutual relations formed according to the system's experience.

Take the understanding of a message as an example. In the simplest case, an input message is expressed in the "native language" of the system, whose grammar rules are built-in, so syntactically valid sentences can be correctly parsed into trees of terms, and each term identifies a concept. Whether the system can "understand" the message firstly depends on whether the system already has the concepts involved in the message. If some of the terms (words) and concepts never occurred in the system's experience, the message has little meaning to the system, and consequently the system cannot understand it, just like we often cannot understand sentences in a foreign language we have not learned. On the other hand, if the system is already familiar with all the concepts involved, this message can be integrated into the system's knowledge base, which will give the system some understanding about it.

According to this Experience-Grounded Semantics (EGS) [9], the "understanding" of NARS on a concept is basically the "meaning" of the concept to the system at the moment, which is determined by the system's experience of the concept. As the system gets more experience about the concept, its understanding about it becomes deeper and richer.

Restricted by the limited processing time and memory space, NARS does not consider all the relations, or the "full meaning", of a concept each time it is involved in thinking. Instead, only part of the relations will be taken into account, so its "current meaning" is usually a very small part of its "full meaning", selected by many factors, such as the stability, usefulness, and relevance of these relations, while these factors change constantly.

EGS does not assume a concept has a "true meaning", nor a "correct understanding", as there is always subjective and idiosyncratic factors in them.

However, when the environment is relatively stable, the accumulated experience will produce relatively stable concepts, so their meaning, and the syatem's understanding of them, will not be arbitrary at all. For systems existing in the same environment and sharing experience via communication, "objective meaning" of a concept appears as the common part of the various versions of it in the individual systems, against which "correct understanding" of the concept can be talked about.

An implication of EGS is that the meaning of a concept, as well as the system's understanding of the concept, changes with context and experience. For the same thing, different systems may have different understandings. Their differences are often not a matter of "right vs. wrong", but "rich vs. poor", or "common vs. peculiar".

Of course, it does not mean that "anything goes" when understanding is under consideration. On the contrary, according to this theory, the traditional AI techniques, both the symbolic ones and the connectionist ones, fail to properly handle understanding because they have been using improper semantic theories. One common problem is that they usually assuming for everything there is a "true understanding", either as an accurate model or function, rather than take understanding as evolving, adaptive, and rooted in the system's experience.

Beside semantics, inference also plays a major role in understanding. In NARS, each piece of new experience triggers a spontaneous inference process, in which the input and the relevant beliefs are used as premises. In this way, to "understand" something not only means to form and remember an internal representation of the input *per se*, but also to recursively generate its implications, and modify the system's memory accordingly. Eventually some actions may need to be taken.

Take communication as an example, very often the "intended meaning" of the speaker is to trigger an action in the listener, or to modify the listener's internal state in a certain way, rather than to add the literal meaning of the message as a new fact into the listener's knowledge base. In NARS, such inference will include the system's beliefs on the intention of the speaker, the context of the communication, etc., which are not explicitly included in the content of the message.

Beyond communication, our understanding of an entity also demands inference on our direct observation about it. For instance, the knowledge on causal relations needed for explanation and prediction is often constructed from observation by inference [11].

A crucial issue in spontaneous inference is how to avoid combinatorial explosion, which would happen if the system attempts to exhaust the implications of an input in a non-trivial situation. Since NARS is based on the assumption of insufficient knowledge and resources, every inference task has a finite time–space budget, determined partly by the innate features of the task (such as the strength and urgency of the motivation) and partly by the system's resource supply (such as the existence of other tasks), and the budget is dynamically adjusted according to the progress of the processing and the changes in the environment [8]. In general, few task is carried to its "logical end" by taking all relevant beliefs into consideration. In this way, spontaneous inference becomes feasible, and it has been designed into NARS from the very beginning.

3 Conclusion and Discussion

Based on the above discussion, we can say that in the AGI context "understanding X" should be understood as "getting the meaning of X", which requires *forming or reforming the system's concept of X*. Here the X can be either a message to be digested in communication, or an object to be thought about and acted on.

In NARS, the "conceptualizing" function is carried out by the same process that is responsible for many other cognitive functions, such as "reasoning" and "learning". As explained in [10], there are reasons to believe that these notions do not referring to difference processes in the human mind, but different aspects of the same process. At least this is the case in NARS.

When we talk about *understanding*, we focus on what the system (a computer, a human, or oneself) knows about an object (a message, a process, or an entity). Accurately we are talking about our concept about it, that is, "our understanding" about the object, not the object itself. A concept is the result of a constructing process, as the result of self-organization of relevant experience.

One characteristic of this understanding of "understanding" is that there is no absolutely *correct* or *full* understanding, though a relative standard is possible. In communication, the intention of the message sender is usually the criteria against which the receiver's understanding is evaluated, though misunderstanding of certain form or extent is usually inevitable, and to the receiver of the message, the more accurate understanding is sometimes even not the most valuable one — we often learn important lessons not intended by the message sender.

Outside communication, when we say we "understand" an object or process, it usually means our concept (or beliefs, knowledge) about it is good enough for our relevant purposes (manipulation, prediction, explanation, etc.). To get a good understanding, the cognitive functions involved are not different from those needed for intelligence in general.

As understanding is the result of processing of relevant experience, systems with different experience may (though not necessarily) form different understandings of the same object. For this reason, we should not expect a perfect mutual understanding between human beings and A(G)I systems, just like we cannot expect a perfect mutual understanding between people with different experience. Instead, "good-enough understanding" should be aimed, and to achieve that, both common experience and common processing mechanism are needed. The former requires human-like sensors and directly feed of human experience at concept-level, though this can only be achieved to a degree. For the latter, understanding demands no less than intelligence and cognition in general.

Compared with the current AI techniques, the major obstacles toward understanding in AGI include the following:

– Narrow understandings of "understanding" as mapping an input instance to a class label, an internal representation, or a command to be executed. This treatment is still common in current projects, especially those using deep learning. This approach works for certain special-purpose systems, but cannot be extended into general-purpose systems, as such a system will not only understand an object in a single way.
– Taking the meaning of a concept as its denotation that is something outside the system. According to this opinion, to understand a concept means to decide its denotation, or to find an interpretation for it, which is isolated from the system's experience. Beside the issue revealed by the "Chinese Room" argument, this approach also fails to show how understanding is established and changed as part of the adaptation process.
– Though some researchers have seen the need for inference in understanding, such inference is usually oriented for a predetermined goal or task. Few AI system has a spontaneous inference process that reveals the implications of an input without running into combinatorial explosion.

As explained above, NARS has overcome these obstacles, and represents a more promising approach to properly understand objects in its environments, as well as the messages in communication with other systems.

References

1. Erman, L.D., Hayes-Roth, F., Lesser, V.R., Reddy, D.R.: The Hearsay-II speech-understanding system: Integrating knowledge to resolve uncertainty. ACM Computing Surveys 12(2), 213–253 (1980)
2. Harnad, S.: The symbol grounding problem. Physica D 42, 335–346 (1990)
3. Karpathy, A., Fei-Fei, L.: Deep visual-semantic alignments for generating image descriptions. IEEE Transactions on Pattern Analysis and Machine Intelligence 39(4), 664–676 (2017)
4. Schank, R.C.: Language and memory. Cognitive Science 4, 243–284 (1980)
5. Schank, R.C., Rieger, C.J.: Inference and the computer understanding of natural language. Artificial Intelligence 5(4), 373–412 (1974)
6. Searle, J.: Minds, brains, and programs. The Behavioral and Brain Sciences 3, 417–424 (1980)
7. Thórisson, K.R., Kremelberg, D., Steunebrink, B.R., Nivel, E.: About understanding. In: The Proceedings of the Ninth Conference on Artificial General Intelligence. pp. 106–117 (2016)
8. Wang, P.: Problem solving with insufficient resources. International Journal of Uncertainty, Fuzziness and Knowledge-based Systems 12(5), 673–700 (2004)
9. Wang, P.: Experience-grounded semantics: a theory for intelligent systems. Cognitive Systems Research 6(4), 282–302 (2005)
10. Wang, P.: Rigid Flexibility: The Logic of Intelligence. Springer, Dordrecht (2006)
11. Wang, P., Hammer, P.: Issues in temporal and causal inference. In: Proceedings of the Eighth Conference on Artificial General Intelligence. pp. 208–217 (2015)
12. Wang, P., Li, X.: Different conceptions of learning: Function approximation vs. self-organization. In: Proceedings of the Ninth Conference on Artificial General Intelligence. pp. 140–149 (2016)

13. Winograd, T.: Understanding Natural Language. Academic Press, Inc., Orlando, FL, USA (1972)
14. Winograd, T.: What does it mean to understand language? Cognitive Science 4(3), 209–241 (1980)
15. Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., et al.: Google's neural machine translation system: Bridging the gap between human and machine translation. CoRR abs/1609.08144 (2016), `http://arxiv.org/abs/1609.08144`
16. Zhang, X., LeCun, Y.: Text understanding from scratch. CoRR abs/1502.01710 (2015), `http://arxiv.org/abs/1502.01710`