Chapter 9
The Importance of Grain Size in Communication Within Cyber-Physical Systems

Julia M. Taylor

This chapter will look at various applications of natural language communication to cyber-physical systems. One of the assumptions that it makes is that such communication is not only necessary for the future systems, but also should be done on a level acceptable and natural to humans, rather than training them to accommodate machine capabilities with exact and precise commands. We will address a grain size of commands or descriptions that could be given to a system—at the same time the physical capabilities of a system will be sketched only as needed for purposes of examples. The range of commands that we are talking about is a typical algorithmic description of a task at the low level and a natural one for a human task description on the high level. A low, more detailed, fine-grain-sized level is assumed to exist already. The higher, coarser-grain-sized level is what we are striving for, in the sense of being able to switch to it automatically when convenient, i.e., to pay with some vagueness, as people and language do, for the ease of not having to resolve an ambiguity.

One of the more difficult things that are taught in algorithmic-thinking-101 is things that we do every day but don't think about enough to describe them. Outline, for example, a step-by-step process to boil eggs. We pride ourselves on being able to explain it in fine detail and praise a 7-year-old that can describe such a process on their own. There are two questions here that come to mind. The first one is, do we want to communicate with agents or systems on such a detailed level? We will not pretend to answer this question here but rather propose a solution if the answer happens to be "no". The second question is what should happen when a command of *drop the egg in the water* is given. In other words, should the egg be really *dropped*, or should correction for the smooth and slow execution of the command be allowed.

We will start with the latter, easier-to-answer question and work our way to the former, much more difficult question with an outlined solution. We will build on

J. M. Taylor (✉)
Computer and Information Technology, Purdue University, Indiana, USA
e-mail: jtaylor1@purdue.edu

S. C. Suh et al. (eds.), *Applied Cyber-Physical Systems*,
DOI: 10.1007/978-1-4614-7336-7_9,

the methods of the Ontological Semantics Technology, explained in the OST section. We will also briefly explain how such methods of communication could be used in developing-algorithmic-thinking classes.

Perhaps an easier initial example is giving directions. When one person asks another person how to get from A to B, some negotiation of knowledge takes place and information at the appropriate level is delivered. For example, one can say that B is right next to C, assuming that C is known to both people, and that's all that needs to be said, there is no need for turn-by-turn instructions. While this example and the boiling of an egg look somewhat different, they have one important thing in common: what is stated in a "natural" communication between people is actually "instinctively" limited to what the questioner/recipient of information may not know already because what they are all aware of is not necessary to restate.

This need to communicate only the necessary information can be looked at from various angles. One angle is granularity: we can afford to get to the highest/coarsest possible grain size that would activate the finer grain size of information in people's brain, without explicitly stating fine-grain details. Another angle is that of the processing of the unsaid [1, 2], which we will leave the unsaid and its inferences for later explorations and concentrate here on the grain size.

For the purposes of this paper, we will separate what people say into local granularity and global granularity of information. Global granularity will refer to that of a script-like (see [3, 4]—cf. the seminal [5, 6]) phenomena where some of the components of the scripts are well known and not verbalized. In other words, instead of telling a story of several paragraphs, only a couple of sentences are necessary to outline the picture. We will refer to local granularity where what is explicitly stated can be treated as a hypernym or a hyponym of what is actually meant. This distinction is, of course, not black or white and there is a lot of gray area in the middle, for instance where the needed and known information can be supplied in one or two sentences, within which the known details are omitted.

Ontological Semantic Technology

We rely on the Ontological Semantic Technology (OST) [7–13] for the needed grain size interpretation. Ontological Semantics is not the only theory/methodology/technology that can handle what is described here. Any system that has a solid and representative of the world ontology that has enough reasoning capability and that is linked to a lexicon should do the trick. Some part of the system has also to accommodate common sense knowledge that people use in every day communication and a collection of scripts that can be accessed for a given scenario. What knowledge base (ontology or not) contains this information is not necessarily important, as long as it can be accommodated. We use OST for convenience's sake and because it can be easily modified to reflect the needed changes.

Ontological Semantic Technology is one of the next generation systems that the theory of Ontological Semantics [14] has produced. What we describe here is the significantly modified version of previously developed commercial systems. OST is not a domain specific technology: it attempts to work with any topic that a human would typically hold a conversation about. Within that, certain domains have more emphasis, especially if they are in a particular application of interest.

At the core of OST is the ontology (see Fig. 9.1)—a model of the world that encompasses all of the non-instantiated knowledge that is needed to comprehend information exchange between a human and a machine. The ontology is language independent—any natural language communication is interpreted through the concepts and relationships that the ontology contains. Just like any two speakers that are fluent in the same multiple languages can start a conversation in one language, switch it to the next, and the next, and the information that is delivered in any of them is about the same, the ontology has the power of that representation and provides the underlying power of the conversion or comparison. Ontology contains concepts and properties, it outlines relationships between concepts through properties, as well as more complex situations that can be bundled together and be useful at a needed grain size to compress or expand information.

A lexicon is a language dependent inventory of all senses of the words for a particular language. Each sense is described in terms of how it is pronounced, what kind of morphological forms it can take, what part of speech it is, what syntactic constructions it can participate in and, most importantly, what is the meaning of the sense. An onomasticon is a language dependent inventory of proper names that are required to support the application in question for a particular language. Every

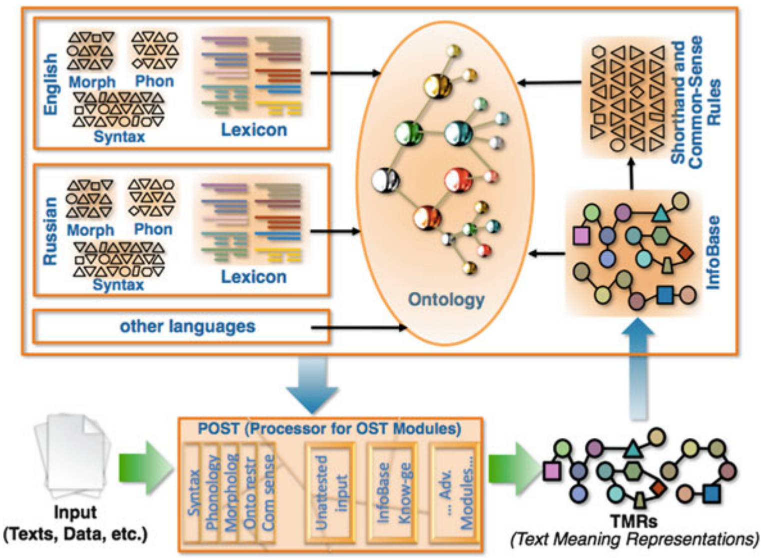

Fig. 9.1 Ontological semantic technology architecture

lexicon and onomasticon sense of a word is anchored in the ontological concept or a "bundle" of concepts that represent the meaning of the sense of the word. What is relevant for cyber-physical systems, is that, through this representation, every command or physical entity loses its ambiguity that is omnipresent in natural language and can be used more precisely. It should be noted that some vagueness will still remain—just like it remains for a human-to-human communication.

Common Sense and Shortcut rules (see [8], cf. other approaches in [14–18]) are rules rather than descriptions and definitions that are in the ontology. Their primary purpose is disambiguation of senses. For example, this repository contains knowledge that you cannot put something larger inside something smaller; that before you end something you have to start it, that a person can be only at one physical location at the same time (grain size is important here too).

Each of these static resources comes with modules that process them. For example, the lexicon comes with syntactic, semantics, phonological and morphological processing. All of these have to return a successful result for the system to return a simple case of a Text Meaning Representation (TMR) of a sentence. It should be noted that for ambiguous sentences there is more than one TMR. A more complicated processing is involved when a word is unknown or when inferences are required. A set of all modules that are responsible for processing information is accessible by the Processor for OST Modules (POST). POST does not only take into account information in text, but may, if needed, look for prior knowledge into InfoBase.

InfoBase is the most "knowledgeable" component of OST—it is where all processed data are stored. InfoBase contains instantiated information of all concepts that were needed to process a particular text. Thus, a generic CAT information is stored in the ontology, but information about particular cats, is stored in InfoBase.

As an example, consider a command "find a kid who knows his name and address." What is of interest here as far as the disambiguation of a lexical sense is concerned is that the word *kid* has at least two meanings, that of a human child and a baby goat (see Fig. 9.2). The restriction of the concept KNOW—that the word *know* is anchored in—for the agent of the event should be an ANIMATE, and both a human child and a baby goat could be applicable here.[1] The restriction of KNOW with a topic of NAME and ADDRESS can only be applicable to a human child [13], thus the sense of a baby goat disappears from the consideration.

Notice that while the command is disambiguated, the rest of the task remains to be performed: a child that knows his name and address still have to be found. We are now looking at the ontological concept KNOW and the common sense knowledge repository of how to check whether somebody knows something and what does it take to find out. From there we will find out that question/answering is the common tool for knowledge solicitation and verification. Therefore, whoever was

[1] Let us assume here that animals do know things and can be valid agents of the concept KNOW just like people are valid agents of it.

Fig. 9.2 Kids chasing a kid (http://www.thisistheplace.org/what_we_do/special_events.shtml)

given the command to find a child, should ask a question of every child and listen for the answer. Notice again, that the task is performed in natural language, and further conversation (other than simple question/answer) may be necessary. Also notice that while some answers are received, unless there is known data about the names and the addresses of these children, what is collected by the executor of our command cannot be verified.

Local Granularity: What is it All about and Should We Be Concerned?

Concepts in natural language are typically easier to disambiguate than to find a correct relationship between the words. For example, even though there were 2 types of kids (person and goat) in the picture, the language itself describing the task dictated that the sense of a goat be rejected.

Of course, it is possible that, instead of giving the command of *finding a kid that knows its name and address*, the command of *catching a kid* be given. All of a sudden no ontological restriction can help the disambiguation mechanism: it is possible to catch both children and goats—so what was meant? One way of handling it is to ask a differentiating question: should a child be caught or a goat? Notice, that if the answer to the question is the former, than more questions arise: any child or a particular child? It might be easier to start the negotiation with asking, which kid, but that, again, depends on whether there is an understanding that there is more than one possibility. And the minute we realize that there is more than one possibility, they have to be represented (see Fig. 9.3 for a hierarchy).

For the purposes of this paper, we refer to local granularity as a phenomenon where ambiguity cannot be resolved and it is masked by coarsening the grain size

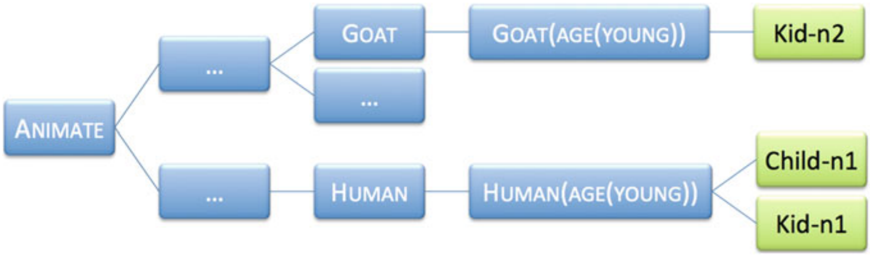

Fig. 9.3 Ontological hierarchy of concepts (in *blue*), with some anchored lexical items (in *green*)

of representation.[2] It is a choice, on the part of the architects, whether to go this route or not. The choice depends not only on the layers of hierarchy between the mutual parents, but also on the number of finer-grain size ontological items in which the ambiguous lexical senses are anchored.

Local granularity phenomenon is even more noticeable in the cases of relationships, especially where words that would be anchored in them are not explicitly stated. A typical example used within OST is that of an *IBM lecture* where it is not clear what the intended relationship between *IBM* and *lecture* is, even if both lexical units used in the phrase are disambiguated. It is tempting to think that information in a sentence would reduce the number of relationships to a minimum, but as demonstrated in [22], it is often not the case. For example, for a sentence *What level of industry expertise exists at local level*, human subjects came up with at least 6 interpretations of *industry experience* within that sentence.

The question of how to represent such vagueness is not just of the semantic nature, especially in the cases of complex nominals (e.g. cat milk bottle). While it is possible to go up the hierarchy of the relationships until the root is reached, it may be just as useful to rely on the syntactic constrains and thus indicate that no semantic restrictions are found at this time [21]. It is possible that this distinction is lost entirely on a native speaker, but it may ease the processing for the machine.

[2] The author is grateful to Victor Raskin for pointing out that this is not that different from considerations underlying Weinreich's [19] objections to what he referred to as Katz and Fodor's [20] "infinite polysemy" in their semantic theory. Why, Weinreich asked, does the theory have to differentiate between two senses of *ingest* (eat solids/drink liquids) but not between two senses of *eat* (with a fork/with a spoon)? Reversing it to fit our discussion, we can say that English masks the latter distinction with the word *eat* but reveals the former distinction with two different words, both, incidentally, in much more common usage than the masking *ingest*.

Global Granularity: A More Interesting Case

We start our discussion with an example of brushing teeth: we all do it every day, and we can probably all easily generate an algorithm of how to do it. Yet, if many of these algorithms were collected, most of them would represent information at a different grain size. Why is it that something that we perform so often and in a similar enough manner, produces such a different description? Let us pretend for the purposes of this discussion that we can describe situations and actions without relying on specific types of memories and that these types do not influence the description.

Now suppose, that you run out of toothpaste before you start brushing your teeth. No matter how different the collection of algorithms were, the result is going to be the same: either the teeth will not be brushed, or more toothpaste will have to be found. In other words, all those algorithms, different in details, will halt at the same place. The same could be said if, all of a sudden, one would run out of water before one can rinse their mouth: the procedure would halt at approximately the same place.

Assuming that one can talk, we would probably not explain why we need toothpaste or water, if we were to ask for more. Moreover, we would expect a certain response: a tube (or some other container) of toothpaste, or an amount of water needed to finish, depending on what was asked. We would not need, nor expect, several gallons of water, for example. Thus, with somewhat different algorithms in mind, we can still assist others if a task is familiar, without describing the whole routine.

We would like the machines to follow the same communication scenario: without explaining step by step what is needed to be done, the knowledge of a particular scenario should be accessed and assistance provided with a request naturally understood by a human. It should be enough to state that one ran out of toothpaste, as a command that another tube should be retrieved (similarly to Searle's indirect speech acts in [22]). Notice also, that just like a human should know that if there is no toothpaste in the house, it should be bought, a machine should understand the same thing.

Where will all this knowledge come from? According to OST, generic knowledge of the world rests either in the ontology or in the common sense and shortcut repository. A generic human script for brushing teeth (overlap between descriptions of how to do it) can be entered also in the appropriate location. An individualized script (instantiated) can be adapted to a particular situation and stored in the InfoBase. This instantiated script can be a result of several conversations. When a human signals that (s)he ran out of toothpaste, an ontological concept triggered by the word's appropriate sense would be retrieved from the ontology and a search of the InfoBase would be initiated. Upon retrieving the needed (instantiated and individualized) script from the appropriate repositories, a halting point will be found and a correction would be provided. What is also of interest here is that since OST's ontology is nothing but a graph, all (weighted) links to where toothpaste can be retrieved from will be found, and processed until the solution in the form of toothpaste is found. And thus, if no toothpaste exists at the location that the person is at, a link that connects toothpaste to an event of BUY will be found as well.

Natural Language Communication with Robots: Examples and Analysis

Most situations where a robot would have to react to natural language are likely to be both local and global. We described how to adapt an OST natural language lexicon to a "Robotese" in [23]. We thus assume here that any robot that we work with comes with a lexicon suitable to perform the tasks that it is capable of.

The terms local and global granularity may not be the best terms for the description of the phenomena at hand. Whenever lexical senses create ambiguity that cannot be resolved, but can be represented by a single higher-grain concept, local ambiguity is at play. In other words, the senses are anchored at a fine-grain concept and we climb up to make the description more compact. Global granularity does the opposite: a sense is anchored in a course-grain concept and in order to understand the text, information from finer-grain concepts has to be brought up. It has a top-down flow of representation, rather the climb that happens in the local granularity scenarios.

Consider a scenario where you are communicating with several robots [24], but you wish them to perform commands that they are capable of. Suppose, you have a robot that is on wheels and a humanoid that can jointly perform a task. Also, suppose, that a humanoid can not only walk, but also run. A command: *move to [name your object that they can both recognize or aware of]* should start them on their "journey." Several things of interest here: while you may anchor your function that is responsible for physical movement of a robot in a fairly generic concept MOVE, it is probably a better choice to anchor it in something that corresponds to MOVE (INSTRUMENT (WHEEL)) in the ontology. If that is the case, then a mechanism of global granularity would have to lower the grain size to the needed concept. At the same time, a command that was given only a named object but not a direction of movement or distance, for that matter, has to be adjusted to what the robot expects and needs to receive in order to function.

On the other hand, the humanoid can both run and walk, and it was given a command to move, which is an ancestor of the ontological concepts representing its real capabilities. It can consider several things in order to decide how to move, including its most stable mode of transportation, or the speed of the wheeled robot. Again, the direction would have to be taken into account here, just as it had to be taken into account for the wheeled robot.

Now, suppose the command is *move inside the building and retrieve a table next to the door.* Let us assume that whoever is giving he command is aware of the physical constraints of the robots and will not ask a humanoid to move the piece of furniture—notice a different sense of the word *move* used here—something that it is not capable of. Let us also assume that whatever humanoid can lift can be placed on the wheeled robot and brought it back that way. The question that remains to be answered is what is a table: is it a small enough piece of furniture that is light enough or some kind of a flat surface with a chart on it? While the furniture should have a much higher weight since it is retrieved from the house, and thus has a stronger association with it, the sense of a chart cannot be dismissed either. This means that one command will have both local and global granularity mechanisms in play.

It could be argued that it is unnecessary to account for both local and global granularity and it only overloads the system: it is just as easy to give exact commands to the robots, according to how they are programmed. One could only understand the word *move* as use your wheels to move forward, and for the humanoids to understand a command of movement either *walk* or *run* have to be mentioned. It may also be possible to argue that even if they can tell a furniture table and some flat surface apart from other objects, they could be disambiguated by a human, or count on the fact human could do the work. But then, again, our goal is seamless communication between humans and machines, in a way that is natural to humans.

Our goal is also for a GPS device to be able to negotiate with a person about what they really want to know and voice just that information. It should not only use streets and intersections for such negotiation, but also buildings and other reference points naturally noticeable to a human. And, it should not use a reference point that may appear frequently—not turning after red barn should be mentioned. Finally, it should adjust to what human wants to know at the grain size that is acceptable to a human in communication with another human.

And then, maybe, just maybe, one day it will be good enough to suggest a better solution to a proposed plan.

References

1. J. M Taylor, V. Raskin, C. F. Hempelmann, & S. Attardo, An Unintentional Inference and Ontological Property Defaults. Proc. IEEE SMC 2010, Istanbul, Turkey, 2010.
2. V. Raskin, J. M. Taylor, & C. F. Hempelmann, "Ontological Semantic Technology for Detecting Insider Threat and Social Engineering." *Proc NSPW*-2010, Concorde, MA, 2010.
3. V. Raskin, S. Nirenburg, I. Nirenburg, C. F. Hempelmann, & K. E. Triezenberg, "The Genesis of a Script for Bankruptcy in Ontological Semantics, in G. Hirst and S. Nirenburg, Eds., Proceedings of the Text Meaning Workshop, HLT/NAACL 2003: Human Language Technology and North American Chapter of the Association of Computational Linguistics Conference. ACL: Edmonton, Alberta, Canada, May 31, 2003.
4. J. A. Crowder, J. M. Taylor, & V. Raskin, "Autonomous Creation and Detection of Procedural Memory Scripts." International Conference on Artificial Intelligence, Las Vegas, NE, July 2012.
5. R. C. Schank, Computer Models of Thought and Language. San Francisco: Freeman, 1973.
6. R. C. Schank, & R. Abelson, Scripts, Plans, Goals, and Understanding. New York: Wiley, 1977.

7. V. Raskin, C. F. Hempelmann, & J. M. Taylor, "Guessing vs. Knowing: The Two Approaches to Semantics in Natural Language Processing," in A. E. Kibrik, Ed., Proceedings of Annual International Conference Dialogue, Moscow, Russia: AABBY/Yandex.
8. J. M. Taylor, V. Raskin, & C. F. Hempelmann, "On an Automatic Acquisition Toolbox for Ontologies and Lexicons in Ontological Semantics." Proc. ICAI-2010: International Conference on Artificial Intelligence, Las Vegas, NE, 2010.
9. C. F. Hempelmann, J. M. Taylor, & V. Raskin, "Application-Guided Ontological Engineering." Proc. ICAI-2010: International Conference on Artificial Intelligence, Las Vegas, NE, 2010.
10. J. M. Taylor, V. Raskin, & C. F. Hempelmann, "From Disambiguation Failures to Common-Sense Knowledge Acquisition: A day in the Life of an Ontological Semantic System." Proc. Web Intelligence Conference, Lyon, France, August, 2011.
11. J. M. Taylor, V. Raskin, & C. F. Hempelmann, "Post-Logical Verification of Ontology and Lexicons: The Ontological Semantic Technology Approach." International Conference on Artificial Intelligence, Las Vegas, NE, July, 2011.
12. J. M. Taylor, & V. Raskin, "Graph Decomposition and Its Use for Ontology Verification and Semantic Representation, Intelligent Linguistic Technologies Workshop at International Conference on Artificial Intelligence, Las Vegas, NE, July, 2011.
13. J. M. Taylor, & V. Raskin, "Understanding the unknown: Unattested input processing in natural language," FUZZ-IEEE Conferenc*e*, Taipei, Taiwan, June, 2011.
14. S. Nirenburg, & V. Raskin, Ontological Semantics. Cambridge, MA: MIT Press, 2004.
15. J. McCarthy, "Programs With Common Sense." Proceedings of the Teddington Conference on the Mechanization of Thought Processes, 75–91. London: Her Majesty's Stationary Office, 1959.
16. D. B. Lenat, "CYC: Toward Programs With Common Sense." *Communications of the ACM* 33:8, 30–49, 1990.
17. J. M. Gordon, & L. K. Schubert, "Quantificational Sharpening of Commonsense Knowledge", in: Havasi 2010, 27–32.
18. C. Havasi, D. B. Lenat, & B. Van Durme, Eds., Commonsense Knowledge: Papers from the AAAI Fall Symposium. Menlo Park, CA: AAAI Press, 2010.
19. H. Liu, & P. Singh, "ConceptNet—a Practical Commonsense Reasoning Tool-Kit, BT Technology Journal, 22:4. 211–226, 2004.
20. U. Weinreich, "Explorations in Semantic Theory." In: T. A. Sebeok, Ed., Current Trends in Linguistics, Vol. 3, 395–477, The Hague: Mouton, 1966.
21. J. J. Katz, & J. A. Fodor, "The Structure of a Semantic Theory," Language 39:1, 170–210, 1963.
22. J. M. Taylor, V. Raskin, & L. M. Stuart, "Matching Human Understanding: Syntax and Semantics Revisited." Proc. IEEE SMC 2012, Seoul, Korea, 2012.
23. E. Matson, J. Taylor, V. Raskin, B.-C. Min, & E. Wilson, "A Natural Language Model for Enabling Human, Agent, Robot and Machine Interaction." The 5th IEEE International Conference on Automation, Robotics and Applications, Wellington, New Zealand, December 2011.
24. D. Erickson, M. DeWees J. Lewis, & E. T. Matson, "Communication for Task Completion with Heterogeneous Robots." 1st International Conference on Robot Intelligence Technology and Applications, Gwangju, South Korea, 2012.