Догадка или знание: два подхода к семантике при обработке естественного языка

Guessing Vs. Knowing: The Two Approaches to Semantics in Natural Language Processing

Raskin Victor (vraskin@purdue.edu),
Hempelmann Christian F. (khempelmann@riverglassinc.com),
Taylor Julia M. (jtaylor1@purdue.edu)

Purdue University & RiverGlass Inc, USA

В статье рассматривается принципиальное различие между двумя подходами к семантике при обработке естественного языка. Первый пытается установить статистически, по совместному употреблению слов как цепочек без значения, о чем говорится в тексте. Авторы предпочитают противоположный подход, построенный на прямом и полном представлении значения текста.

1. Preface: The Brief History of the Dichotomy

This paper discusses two approaches to semantics in natural language processing (NLP), the prevalent statistical/machine learning approach (SML) and the persistent meaning-based minority approach, the "real" computational semantics, based on direct and comprehensive meaning access (DMA). Roughly the same division is captured by contrasting SML to the rule-based approach. Yet another way to expose the dichotomy is to separate the meaning and usage of NLP and computational linguistics—contrary to the common practice of using the two terms synonymously. And, finally for now, the difference is often placed in the attitude towards the "knowledge acquisition bottleneck" (Navigli 2009—see also Section 4 below), with SML recognizing it and dismissing knowledge acquisition as subjective, non-scalable, and non-feasible, and DMA making it the cornerstone of its endeavor.

This dichotomy is not new: there is a great deal of similarity between its current guise and the naïve "95 % vs. 100 %" debate in early machine translation (MT: Akhmanova *et al*. 1961), with the "100 %" side insisting on knowing and understanding everything before computing it, and the "95 %" side interested in trying all kinds of formal methods, from an extremely limited repertoire then available, and finding out what they can tell us. Bar-Hillel (1954) introduced the notion of the semantic bottleneck on the way to a fully automatic high-quality MT very early on, and in the next decades, whether informed by his opinion or not, most people in NLP tried non-semantic methods while a small minority attempted to remove the bottleneck. The former, many of them linguists, but not semanticists, attempted to perfect the syntactic parsers, thinking that this would somehow make semantics unnecessary.

The US government funding for MT disappeared with the publication of the "Black Book" (Languages and Machines 1966) and resumed slowly in the mid 1980s with a new, knowledge-based approach. Other applications, such as information retrieval, information extraction, text mining, summarization, search, and question answering were emerging. By 1990 or so, the syntacticians were replaced, in the same non-, a-, and outright anti-semantic camp, by people, mostly non-linguists, who were interested in applying statistical and machine-learning methods instead of syntax, first admitting openly and even bragging about fully sharing the non-semantic orientation of their predecessors and then, since roughly around 1997, when most government funding started stipulating the necessity of semantics, claiming that theirs was the only way to the meaning of texts. On the other side, a small group of computational semanticists was making a case for the feasibility of acquiring semantic resources, such as the ontology and lexicons.

Another dimension bearing on the same dichotomy is the relation between NLP and artificial intelligence (AI). During the peak of excitement about AI, in the 1980s, NLP proudly referred to itself as the NL AI. Against the background of the Chomskians' (typically, much more stridently than Chomsky's himself) claims that the transformational generative grammar, which was mostly syntax, was not just a model of what was internalized in the mental mechanisms underlying language but actually the content of those mechanisms (see a critique in Raskin 1979), the "just a model" position allied itself with the "weak AI" principle; the "content" position with the strong AI thesis. Very few NLP principals adopted the latter, which claimed that it would develop machines which would think exactly like humans—or, conversely, that the principles, on which those machines will be built, will be identical with the principles on which the human mind operates. But the weak AI thesis was very appealing and more realistic—and compatible with the black box basis of cybernetics at its inception.

NLP systems may be seen in two different lights depending on whether their developers are interested in their results, no matter what methods are applied, or in the methods emulating human thinking. The SML methods can be easily seen as not motivated by the weak AI thesis while the DMA methods appear clearly informed by it. Simplistically and not fully accurately, people process language by knowing the meanings of the words and phrases and how they are combined together in sentences—they do not process large massives of statistical data, certainly not rapidly enough for that to be useful for instantaneous understanding of text.

Finally, an interesting historical parallel invites itself into this discussion. Late in the 19th century, when Frege (1884) introduced the preoccupation with the foundations and justification of well-established disciplines that has led to the emergence of the philosophy of science and the philosophies of specific disciplines, Russell wanted the exact definitions of various terms. Having failed to find them from the linguists because linguistic semantics, barely 40 years old, had not yet graduated to meaning definitions, he wrote it off and developed the parallel discipline of the philosophy of language to deliver what he needed (it has not). In the 1990s, when meaning became a *sine qua non* of NLP, the computer scientists and statisticians could not find any semantic help in the field and proceeded to develop the non-linguistic methods to do something apparently meaning-related.

2. The Plane of Expression vs. the Plane of Content

The distinction between what is perceivable by the senses and what it can symbolize, always known to humans and possibly to animals, had received a conceptual support in philosophy (Peirce 1991), from which the new discipline of semiotics emerged, before it reached linguistics in the writings of Saussure (1916): his *signifiant* was the material side of the sign that people can hear or see, while his *signifié* is the meaning of the sign. His maverick European successor, Louis Hjelmslev (1953), captured the same distinction in his plane of expression for the *signifiant* and plane of content for the *signifié* and bifurcated each plane into form and substance, thus imposing structure on both planes and introducing the notion of commutation between the two, so that the changes on the one plane are reflected on the other. He then proceeded to remove the two substance disciplines, phonetics (substance of expression) and semantics (substance of content) from his linguistics, leaving only phonology (form of expression) and grammar (form of content) in. Deviating from his own decree, however, Hjelmslev later (1958) made one of the very few structuralist contributions to semantics by trying to impose structure on content by comparing the dissection of the same narrow semantic fields (*tree/wood/forest*; *brother/sister/sibling*) in multiple languages (see also Trier 1931, Weisgerber 1950).

It is interesting to throw the two-plane perspective and their commutation on NLP. The syntacticians of the 1960–80s, quite independently of Hjelmslev, who did not leave a school and was—and is—largely forgotten, also attempted to get to the content through its form. These efforts, reinforced first by Chomsky's (1965) venture into a tiny area of semantics, immediately adjacent to syntax (and annexed by him into syntax) and then by a massive infusion of first-order predicate logic, rediscovered by linguists yet again, probably in part thanks to the elegant McCawley (1993), led to the separation of formal semantics from lexical semantics and to the exclusive focus in the thus redefined discipline of semantics on the former (the attempts to save lexical semantics from the charge of being just substance by discovering the grinding rule—*cow/beef, sheep/mutton*—were entertaining but short-lived: see Nirenburg and Raskin 2004: 117 and references there). At its peak, however, formal semantics had to admit that syntactic distinctions and semantic distinctions did not coincide (see Raskin 1994 and references there) but continued to operate, often with admirable virtuosity, on this counterfactual basis and to focus on the most grammatical aspects of meaning, such as quantifiers and other direct reflections in NL of what was clearly defined in logic.

SML, however, operates on the unarticulated assumption that substance is not accessible directly at all, form or content. Instead, it believes that the regularities of co-occurrence, masterfully augmented on multiple parameters, can classify texts without actually understanding its meaning by the computer. Whether they know it or not (and some do), they operate in the Wittgensteinian "language is usage" tradition reflected in semantics by Firth's (1957, cf. Raskin 1971) meaning by collocation: an important part of the meaning of *dark* is its collocation with *night*, and vice versa.

Contrary to that, DMA believes that direct and comprehensive (non-selective) approach is essential for the ultimate success of NLP's growing list of applications. In plain linguistic terms, it means the delivery of the meanings of words and phrasals to the computer. Their opponents may even agree to this position—they simply claim that it is impossible to accomplish (see Section 5 below). In the next two sections, we will sketch out the alternative positions of the two camps. We will then discuss the evaluation attack by the one-planer SML on the two-planer DMA, and the need for the counter-offensive.

3. SML: A Loving View

It is not that the "non-semantic" approach is not interested in the semantics of text: one has to be if one is to compete these days for NLP funding. It is just that they, *a priori*, consider unimplementable the universal program of linguistic semantic theory, which, since and after the much maligned but seminal Katz and Fodor (1963), has included a lexicon and the compositional rules combining lexical into sentential meanings. Instead, they use a combination of statistical methods (Manning and Schuetze 2000) with machine learning (Mitchell 1997).

Because, as mentioned above, ambiguity has been seen as the main cause for Bar Hillel's semantic bottleneck, it is useful to see how this approach handles the problem of word sense disambiguation (WSD: see Kilgariff 1997, 1998, 2006; Ide and Veronis 1998; Ide and Wilks 2006; cf. a very useful and candid survey in Navigli 2009 and references there). We will describe here a common, most typical, generic case rather than any particular implementation, so various improvements may have been already added in later manifestations, without affecting this discussion.

A large corpus of text is divided into a smaller training part and a test part. Selective ambiguous (polysemous or homonymous) words are marked throughout the text, usually no more than one per sentence, and their different sense, commonly no more than two per word, are indicated. Human subjects mark the sense that they prefer contextually. Then the statistical/machine learning system attempts to guess the correct senses of the similarly marked words in the test corpus on the basis of statistical properties it observes in the contexts of the selections, mostly the co-occurrence of certain words significantly more frequent than its random probability. This principle has not changed basically but has been much refined since such early pioneering works as Shaykevich (1963).

A typical application for WSD has been a search for documents from a large control corpus pertaining to a small set of keywords, and the results are evaluated in terms of recall and precision. Recall is the number of hits retrieved divided by all hits in the corpus that should have been retrieved, and precision is the number of hits retrieved divided by the number of all instances retrieved, including false positives. The evaluation metrics are an important part of the approach (see also Section 5 below), and they have been perfected in multiple SemEval/SenseEval (see, for instance, Aguirre *et al*. 2007) competitions that are part US Government-stipulated for all of its grantees and part voluntarily participation by the proponents. The improvement on these scores gives a sense of individual and industry progress; it translates into rankings, awards, prizes, etc. The practitioners of the approach see these evaluations as the NLP standard and attempt to assess other approaches in these terms.

The approach avoids the challenging problems of understanding text beyond judging it pertinent to a keyword set and thus ignores all the problems of understanding how natural language works and how the human processes information. Unattested input, plasticity of meaning, salience, inference, reasoning in NL do not present any problem for it, and the practitioners are proud of it. It is widely believed that even WSD is "an AI-complete problem [Mallery 1988], that is, by analogy to NP-completeness in complexity theory, a problem whose difficulty is equivalent to solving central problems of *artificial intelligence* (AI), for example, the Turing Test[Turing 1950]" (Navigli 2009). He continues to say that "[u]nfortunately, the manual creation of knowledge resources is an expensive and timeconsuming effort [Ng 1997], which must be repeated every time the disambiguation scenario changes (e.g., in the presence of new domains, different languages, and even sense inventories). This is a fundamental problem which pervades the field of WSD and is called the *knowledge acquisition bottleneck* [Gale et al. 1992b]."

In the next section, presenting DMA, we will respectfully question these statements and attribute them to the lack of interest, ideological and disciplinary preferences, and in many cases, insufficient understanding of linguistic semantics and/or experience with descriptive semantics. We will have to agree, regretfully, with Navigli's rather damning (to us) self-assessment that "[t]he hardness of WSD [in SML practice—added by us] is also attested by the lack of applications to real-world tasks."

4. DMA: A Critical Self-Scrutiny

As described in Sections 1–2, the opposite approach, DMA, to which we refer as semantic and sometimes, for emphasis, "semantic semantic" or "real semantic" (see Hempelmann and Raskin 2008), differs in that it: (a) sets itself and the field the much more ambitious goal of direct and comprehensive meaning access to text and (b) proceeds to that goal, in full compliance with the program of linguistic semantic theory, by acquiring the same knowledge resources as it sees the humans as having and using in understanding NL. It also fully subscribes to the weak AI thesis, thus pursuing an interest in modeling/emulating the way humans are hypothesized to process meaning. In other words, besides the interest in getting optimal results in NLP applications, this approach maintains that the only way to achieve this is by programming the computer to emulate human understanding. Capitalizing on the training and experience in descriptive linguistic semantics, it is free of the "fear of semantics" (Raskin 1988) and directly attacks what it does not actually perceive as the "knowledge acquisition bottleneck." It disagrees with the AI-complexity assessment of its task and believes that, while actually fraught with more complexities than SML recognizes, the goal is implementable, and it will suffice to do it only once, with the methodology of domain extension in place and robustness with regard to unattested input built in.

We will illustrate this on the example of a specific approach we have been developing, improving, and implementing for a couple of decades. Ontological Semantics has been extensively reported in a number of publications since its inception in the late 1980s, most comprehensively in Nirenburg and Raskin (2004) and recently at this forum as well (Raskin 2006, Petrenko 2009). The current incarnation of Ontological Semantics, which we call the Ontological Semantics Technology (OST) is characterized by the implementation of many but not all features outlined in Nirenburg and Raskin (2004), resulting in a considerable revision of these features and sometimes radical departures from the earlier views. A proprietary implementation of the system has reached a functional implementation stage, well beyond the earlier academic proof-of-concept demonstrations, such as the MikroKosmos MT system (see the references in Nirenburg and Raskin 2004: 29).

The main resource of OST is the engineered language-independent ontology, consisting of a lattice of concepts, each of which is a set of properties (slots, facets, and fillers), including the subsumption and mereological (part-whole) properties that are common to ontologies but adding several hundred other properties, so that the concept HOUSE will look, in its simplified non-proprietary and unfaceted partial property-filler(s) format as follows:

house
    is-a residential building
    has-object-as-part room staircase balcony entrance
    *location street lot square
    *material wood stone metal glass
    *has-object-as-part wall window door foundation roof
    *theme-of build reside

The asterisked properties are inherited from the parent or ancestor concepts. The English lexical items, such as *house*, of course, but also *cottage, villa, mansion, bungalow, cabin* but not *hut, tent, yurt* will be anchored in this concept, as will indeed the Russian *дом, вилла, замок, дворец, дача* be. Processing the sentence *She lived in a big house*, the OST analyzer will read the words, find them in the lexicon, identify the concepts they are anchored in, if any, and try and identify the fillers for the event RESIDE properties (HUMAN and HOUSE will fit into its AGENT and THEME requirements, respectively).

Most of the OST effort is, of course, devoted to the interesting cases of no easy and ideal fit (see the initial sketches for some of those cases in Nirenburg and Raskin 2004, Ch. 8; see Hempelmann *et al.* 2010 for an easy-to-medium case of ambiguity resolved by the technology). The current implementation, even with many necessary modules in the analyzer not yet coded, already analyzes a large chunk of sentences correctly (if you want exact figures read the next section).

Let us now address the standard charges of non-feasibility, subjectivity, and non-scalability raised against us by SML. Our experience and rapid progress towards the product-level implementation, the effort started around 2004, has allowed us to reduce the cost of an individual concept to sligtly over $3 and of the lexical sense to $2.50. We estimate that we need under 15,000 concepts and under 150,000 senses to provide adequate coverage, complemented by the robust unattested input module (see Nirenburg and Raskin 2004: 279–282), already implemented for proper nouns. Doing this from scratch would cost thus $420,000, but much of it is already done and is licensable for less. We are also considering putting up a legacy 6.500-concept ontology and 21,000-sense lexicon (KBAE 2002), after improvement they badly need, as an open source resource. The extension to a new sublanguage/domain, which we have executed 8 times so far, involves, on the average, 6 person months at the post-doctorate level, or under $30,000, and it enriches the ontology by around 50 concepts and the lexicon by around 400 senses (the domains which have hundreds and thousands of terms, such as genomics, bring in many more phrasals but they are highly structured and easy to handle). Finally, translating the senses (not the words!) from one language to another involves a student-level bilingual who does not have to be a linguist. Executed several times fully or partially, it costs under $20,000 per pair of languages, with about $5,000 more for naturalizing the SYN-STRUC zones of the lexical entries in the target language. How does it all compare with multiples of millions spent so far on the SML efforts?

Feasibility involves not just the cost but whether it is at all possible to do it objectively. Well, there is an extreme and useless, even if somewhat plausible view that we all speak idiolects, our own individual subjective languages (see Raskin 1971). So, if you, the reader, understand what we are saying here, it should be reported as a miracle to an appropriate Facebook group! Because, you see, we are using our own idiolect, mysteriously negotiated among the three of us, while you are using your own. The reality, crude as it may be, is that we understand each other most of the time, and, accordingly, the OST acquistion process is set up as a hybrid computer-human effort limiting the human performance, basically, to a multiple choice and thus ensuring homogenuity and continuity in the effort. Our elaborate and ever developing acquisition toolbox, increasingly automated, the likes of which would have saved CyC (Lenat 1995) from abandoning its noble initial task of structuring all the «lemmas» of common sense, is briefly sketched out in Taylor *et al.* 2010.

Finally, for the non-scalability charge that should really not be raised from a glass house. Our 15,000-concept and 150,000-sense resources will provide pretty adequate coverage for all possible meaning, this combining the maximum descrptive power with a built-in explanatory power, and our unattested input will do an even better-than-already-implemented job of guessing the senses of unattested input. Our extension to new domains has been tested. What other scalability is there? Oh, language changes, we hear. Indeed, there is that. How much have Russian or English changed in the course of our presentation? Our unattested input, again, can handle the few hundred new words a year, including—most prominently—the English *friend* and *unfriend* as verbs, and the US government prolixity to add several hundred acronyms a month to their special Gobbledegook dialect that does defeat an occasional visitor to Washington, DC, but fortunately, these are forgotten at the same or higher speed. We are developing and perfecting an increasingly automated module for new acquisition , so that unattested input, partially with human approval and correction, be learned, which means that OST includes lexicon learning, and possibly ontology learning (where, incidentally, machine learning techniques could be used, but on TMRs rather than on words and sentences as meaningless character strings.). We are not sure, of course, that full automation will ever be possible, and we refuse to be fazed by it. And how many different training corpora need to be tagged to train the statistical/machine-learning systems for different domains, different corpora, more than two senses per ambiguous word? A friendly question to a co-author of a 2004 workshop presentation at ACL/IJCAI revealed that tagging for annotation, a shallow semantic effort in SML, cost about $75 per sentence per person. Isn't there a scalability/cost issue there?

5. The Evaluation Game That DMA Should Learn From SML (Not Really!)

SML has an enormous advantage over DMA: they have brought the evaluation game to perfection. First, they adjusted it to the very limited, if not only functionality that their technology is capable of, and that they declare the industry/field standard, namely, to identify pertinent documents. The Semeval/Senseval competitions and their informal extensions to other fora (= *forums) give the researchers the quantifiable bragging rights. As Hempelmann and Raskin (2008) polemically claimed, there is definite pride in showing that "our" results outrank "theirs" by .28 %. Early voices objected to these often self-serving metrics as falling far short of real efficacy in product-level applications and user acceptance, and that paucity of real-life applications using SML methods successfully confirms that. The semantic camp is reluctantly developing similar self-encouraging metrics, and we will report them later if we really have to.

The position of the semantic camp has always been that the proof of the pudding is in the eating, not in measuring it according to a number of quantifiable parameters (weight, size, density, color?). The weakness of this position is that the resources have to be developed to a certain minimally functional phase and an application implemented before any informative evaluation. Before it happens, the recall and precision metric is not really applicable. We have been asked whether OST can improve tagging, and our honest response that OST makes tagging unnecessary stuns the younger non-semanticists in NLP who feel that a pillar is being removed from their world. Apparently, the famous Chapaev joke about saddles for the ICBMs is not part of every NLP student's education.

The appropriate evaluation set for the semantic camp should include a number of advanced functionalities that are much harder to develop without understanding the text, such as the paraphrases and similar texts using totally different words that are not from the WordNet synsets; identifying the actual answers to the queries instead of letting users look for them and, more often than not, find them in the documents deemed pertinent; understanding, as humans routinely do, what is left unsaid, namely, inferences, ellipses, implicatures, etc.

With the very low goals set for itself by SML, NLP, like linguistics before it, is getting a very bad name in the field and industries that need our services. The latest disappointed customers are the law firms buying e-discovery products. The yield of the e-mails pertaining to a lawyer's deposition queries is high on recall and very low on precision but the US legislation disallows "fishing expeditions," that is, seeking information on a much broader scope than the case justifies and getting unrelated information. In a growing number of cases, the judge examines the e-discovery yields, sees tons of irrelevant information brought out by the well-pointed but not understood queries, and throws out the entire yield as a "fishing expedition," illegal in US legal system, even if it contains a tiny percentage that is crucial evidence. This leads to lost cases and millions of dollars of damage, and the lawyers are desperate for e-discovery with understanding.

There is an enormous amount of energy and talent on both sides of the semantic divide, and if we set our goals right and channel the energy in that direction we will see an enormous jump in the quality of NLP. We have every reason to believe, as per our skill sets and experience, that the future of NLP depends on the availability of real text understanding. And, incidentally, there is absolutely nothing wrong with statistics or machine learning—as long as we stop applying them to meaningless character strings instead of the elements of meaning, such as ontological property fillers. Until then the competitions will only be marginally meaningful and participation of DMA-type systems not possible, for the simple reason that "AI, linguistics and IR were respectively seeking propositions, sentences and byte-strings and there is no clear commensurability between the criteria for determining the three kinds of entities" (Wilks and Brewster 2009: 47).

6. OST in Action

We have illustrated the OST conceptual apparatus and mode of operation on a simple example of an ambiguous sentence *A dog ate a mouse* in Hempelmann et al. (2010). Now we will demonstrate how OST technology handles a deliberately non-compositional example on, first, *Bill kicked the bucket* and later on *Bill kicked the bucket and dented it.* It should be noted that the Google translation of the latter into Russian arrives at *Билл загнулся и помят его,* thus completely missing the appropriate sense in the first clause.

When the sentence *Bill kicked the bucket* is interpreted by the Semantic Text Analyzer (STAn) with the help of the OST English lexicon and language-independent ontology, the following entries are selected by STAn from the lexicon for consideration:

```
(kick
	[(kick-v1 is domain-dependent and not considered here)]
	(kick-v2
		(cat(v))
		(anno(def "to remove from a place as a result of a violation")
			(ex "the security kicked the offender out"))
		…
		(syn-struc((subject((root($var1))(cat(np))))
			(root($var0))(cat(v))
			(phr((root(out))(cat(phr))))
			(directobject((root($var2))(cat(np))))))
		(syn-struc1((subject((root($var1))(cat(np))))
			(root($var0))(cat(v))
			(directobject((root($var2))(cat(np))))
			(phr((root(out))(cat(phr))))))
		(sem-struc(remove(precondition(value(^$var99
				(should-be-a(sem(minor-crime))))))
			(agent(value(^$var1(should-be-a(sem(human))))))
			(beneficiary(value(^$var2(should-be-a(sem(human))))))))
	)
	(kick-v3
		(cat(v))
		(anno(def "to punch, usually with the foot")
			(ex "he kicked the ball. the foot kicked the ball. the wind kicked the ball")
			(senseprim(1)))
		…
		(syn-struc((subject((root($var1))(cat(np))))
			(root($var0))(cat(v))
			(directobject((root($var2))(cat(np))))))
		(sem-struc(kick(agent(value(^$var1)))
			(instrument(value(^$var1)))
			(precondition(value(^$var1(should-be-a(sem(physical-event))))))
			(theme(value(^$var2)))))
	)
	(kick-v4
		(cat(v))
		(anno(def "to die")(ex "the old man kicked the bucket"))
		…
		(syn-struc((subject((root($var1))(cat(np))))
			(root($var0))(cat(v))
			(directobject((root(bucket))(cat(np))))))
```

```
                    (sem-struc(die(agent(value(^$var1)))))
          )
)
(bucket
          (bucket-n1
                    (cat(n))
                    (synonyms "")
                    (anno(def "a cylindrical vessel or metal, plastic or wood")
                              (ex "he poured water into the bucket"))
                    (syn-struc((root($var0))(cat(n))))
                    (sem-struc(bucket))
          )
)
(bill
          (bill-n1
                    …
                    (sem-struc(beak))
          )
          (bill-n2
                    …
                    (sem-struc(text(has-topic(sem(pay)))))
          )
)
(Bill
          (Bill-pnd1
                    …
                    (sem-struc(human(gender(value(male)))(has-name(value("Bill")))))
          )
)
```

Next, STAn checks all of the above entries for their mutual compatibility on the basis of the infomration in their SYN-STRUCS, SEM-STRUCS. The SEM-STRUCS are checked against the ontological concept that the entries are anchored in or restricted to (see Raskin et al. 2010 for a formal description of the lexicon; Taylor et al. 2010 and Taylor and Raskin 2010 for a formal description of the ontology and the OST reasoning process). The results of STAn's interpretation of the sentence are:

```
TMR 1:Weight(TMR): 4.24 Event:
                              kick-v3,
                                      kick1 agent(value (Bill-pnd1, human1(gender(value(male)))
                                                (has-name(value("Bill"))) ))
                                                theme(value (bucket-n1, bucket1 ))

TMR 2:Weight(TMR): 3.0900002 Event:
                                    kick-v4,
                                            die1 agent(value (Bill-pnd1, human1(gender(value(male)))(
                                            has-name(value("Bill"))) ))
```

Notice that STAn recognizes both interpretations of the sentence: Bill died or Bill hit a physical object. Now let us consider the sentence *Bill kicked the bucket and dented it*.

```
          (dent-v1
                    …
                    (cat(v))
                    (syn-struc((subject((root($var1))(cat(np))))
                              (root($var0))(cat(v))
                              (directobject((root($var2))(cat(np))))))
                    (sem-struc(damage(relative-force(less-than(0.3)))
                              (agent(value(^$var1)))
                              (theme(value(^$var2)))))
          )
```

There are several modules that will be activated to process this sentence, in addition to the previously mentioned ones. The sense of *dent* requires a subject and an object, the subject will be selected to be *Bill*, and the object, the pronoun *it*. The next step is to resolve the pronoun (Co-Reference Module). There is only one satisfactory candidate, and that is the concept that corresponds to the word *bucket*. Notice that while such a concept exists in TMR1, the interpretation of the idiomatic expression in TMR2 removes such possibility. Thus, the combination of the two clauses in the new sentence is only possible with TMR1. This successfully disambiguates *kick the bucket* and removes the sense of dying or загнулся from the table. When STAn's Event Embedment Module checks the possible relationship between KICK and DAMAGE in the ontology, we will also find that DAMAGE is an effect of KICK, resulting in the following interpretation of the sentence:

```
kick1   (agent(value (human1
                (gender(value(male)))
                (has-name(value("Bill")))
        )))
        (theme(value (bucket1)))
        (effect(value(damage
                (relative-force(less-than(0.3)))
                (theme(value(bucket1)))
        )))
```

While this is a constructed example, it demonstrates the capability of the system on the actual, real-life ontology and lexicon of an implemented system. It is this capability that keeps all interpretations of the ambiguous sentences when needed, and removes them when there is enough knowledge to provide accurate results in machine understanding of natural language.

7. Conclusion: Crawling, Flying, and Other Self-Propelling Modes

At a major NLP gathering a few years into this century, a fellow workshop participant, a master grant getter, declared, in fake admiration, that the semantic approach was about flying while he and his non-semantic confrères were crawling, which he made sound as something in hand (his ambitious project has yielded no known result). Maxim Gorky also had something to say about crawling and flying but he never got a single Federal grant… We think it is actually neither about crawling nor flying but rather about walking, preferably running, with one's feet firmly on the ground (most of the time) and the pace rapidly accelerating, to the goal (not Grail) that the industrial, societal and academic needs are making it increasingly urgent for us to reach. One cannot help recalling here this once famous (Hubert) Dreyfus (1992: 100) quote: "…the first man to climb a tree could claim tangible progress toward reaching the moon. Rather than climbing blindly, it's better to look where one is going." Are we, the computational semanticists, just a leg or so ahead because we are already trying? Should we really be spending enormous amounts of money and effort on just divining what is hidden behind the door of meaning that we presume closed, or should we continue the difficult but, ultimately, less costly work on the doorstop that keeps moving the door towards the wall, wider and wider, even if possibly we will never succeed to take it off the hinges? Will the curiosity kill the cat? Meow!

References

1. *Agirre E., Marquez L. and Wicentowski R. (eds.). Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval)//Association for Computational Linguistics, Prague, Czech Republic, 2007.*
2. *Ахманова О. С., Мельчук И. А., Падучева Е. В. и Фрумкина Р. М.* О точных методах исследования языка.//М.: Издательство Московского государственного университета, 1961.
3. *Bar-Hillel Y.* A demonstration of the non-feasibility of fully automatic high quality translation, 1954. Reprinted in his: Language and Information //Jerusalem: Magness, 1960.
4. *Chomsky N.* Aspects of the Theory of Syntax // Cambridge, MA: MIT Press, 1965.
5. *Dreyfus H. L.* What Computers Still Can't Do: A Critique of Artificial Reason.//Cambridge, MA: MIT Press, 1992
6. *Firth J. R.* Modes of meaning. In his: Papers in Linguistics.//London: Oxford University Press, 1957, P. 19–215.
7. *Frege G.* Die Grundlagen der Arithmetik. Eine logisch-mathematische Untersuchung ueber den Begriff der Zahl.// Breslau: Koebner., 1884. English translation by J. L. Austin, Foundations of Arithmetic.//Oxford: Blackwell, 1950.
8. *Gale W. A., Church K. and Yarowski D.* A method for disambiguating word senses in a corpus.// Comput. Human, 26: P. 415–439, 1992.
9. *Hempelmann C. F., Raskin V. and Taylor J. M.* In: Proceedings of ICAI '10.//Las Vegas, NE, July 2010.
10. *Hempelmann C. F. and Raskin V.* Semantic search: Content Vs. formalism. //Proceedings of LangTech 2008, http://www.langtech.it/en/technical_program/technical_program.htm, 2008.
11. *Hjelmslev L.* Prolegomena to a Theory of Language.// International Journal of American Linguistics, Suppl. to vol. XIX, 1953.
12. *Hjelmslev L.* Dans quelle mésure les significations des mots peuvent-elles etre considérées comme formant une structure?//Proceedings of the Eighth International Congress of Linguists, Oslo, 636–654, 1958.

13. *Ide N. and Veronis J.* Word sense disambiguation: The state of the art.//Computat. Ling. *24,* 1: 1–40, 1998.
14. *Ide N. and Wilks Y.* Making sense about sense. // In: E. Agirre and P. Edmonds (eds.), Word Sense Disambiguation: Algorithms and Applications.// Springer, New York, NY, 47–73, 2006.
15. *Katz J. J. and Fodor J. A.* The structure of a semantic theory.//Language 39(1): 170–210, 1963.
16. *KBAE:* Knowledge-Based Acquisition Editor. Purdue Version 2.1. http: //kbae.cerias.purdue.edu:443/; guest login (browsing only), nlpgroup; password, ch@ng3me.//NLP Lab and CERIAS, Purdue University W. Lafayette, IN, 2002.
17. *Kilgariff A.* I don't believe in word senses // Comput. Human. *31,* 2, 91–113, 1997.
18. *Kilgariff A.* Senseval: An exercise in evaluating word sense disambiguation programs?? LREC, Granada, Spain, 1255–1258, 1998.
19. *Kilgariff A.* Word senses. // In: E. Agirre and P. Edmonds (eds.), Word Sense Disambiguation: Algorithms and Applications // Springer, New York, 29–46, 2006.
20. *Languages and Machines.* Computers in Translation and Linguistics. A Report by Automatic Language Processing Advisory Committee, Division of Behavioral Studies, National Academy of Sciences, National Research Council, Publication 1416 // Washington, DC, 1966.
21. *Lenat D. B.* Cyc: A large-scale investment in knowledge infrastructure // Communications of the ACM 38(11), 1995.
22. *Mallery J. C.* Thinking about foreign policy: Finding an appropriate role for artificial intelligence computers.//Unpublished Ph. D. dissertation. Political Science Department, MIT, Cambridge, MA, 1988.
23. *Manning C. D. and Schuetze H.* Foundations of Statistical Natural Language processing. Second Printing with Corrections.//Cambridge, MA: MIT Press, 2000.
24. *McCawley J. D.* Everything That Linguists Have Always Wanted to Know About Logic but Were Ashamed to Ask, 2nd ed. // Chicago: University of Chicago Press, 1993.
25. *Mitchell T. M.* Machine Learning // New York: McGraw Hill, 1997.
26. *Navigli R.* Word sense disambiguation: A survey.// ACM Comput. Surv. 41, 2, Article 10 (February 2009), http://doi.acm.org/10.1145/1459352.1459355, 2009.
27. *Ng T. H.* Getting serious about word sense disambiguation.//In: Proceedings of the ACL SIGLEX Workshop on Tagging Text with Lexical Semantics: Why, What and How? (Washington D.C.). 1–7, 1997.
28. *Nirenburg S. and Raskin V.* Ontological Semantics.//Cambridge, MA: MIT Press, 2004.
29. *Peirce C. S.* On the Nature of Signs. In: J. Hoopes (ed.), Peirce on Signs : Writings on Semiotic by Charles Sanders Pierce // Chapel Hill, NC: University of North Carolina Press, 116–141, 1991.
30. *Petrenko M.* Ontological semantics and abduction: parsing ellipsis. // Компьютерная лингвистика и интеллектуальные технологии, по материалам ежегодной Международной конференции "Диалог 2009." Выпуск 8 (15). /Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue 2009."Issue 8 (15). Moscow: RGGU Press, 2009.
31. *Раскин В.* К теории языковых подсистем///М.: Издательство Московского государственного университета, 1971. Второе дополненное издание.//Москва: Издательство ЛКИ, 2007.
32. *Raskin V.* Theory and practice of justification in linguistics.//In: P. L. Clyne et al. (eds.), Papers from the Parasession on the Elements, Chicago: Chicago Linguistics Society, 1979.
33. *Raskin V.* Fear of Semantics. Invited lecture.//Center for Machine Translation, Carnegie Mellon University, Pittsburgh, 1988.
34. *Raskin V.* Frawley: Linguistic Semantics.//Language 70(3): 552–556, 1994.
35. *Raskin V.* The whys and hows of ontological semantics. // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции "Диалог 2006" (Бекасово, 31 мая — 4 июня 2006 г.). / Под ред. Н. И. Лауфер, А. С. Нариньяни, В. П. Селегея / N. I. Laufer, A. S. Narinyani and V. P. Selegei (eds.), Computational Linguistics and Intellectual Technologies. International Conference "Dialogue 2006" Proceedings (Bekasovo, 31 May — 4 June, 2006)/. Moscow: RGGU Press, 2006.
36. *Raskin V., Taylor J. M. and Hempelmann C. F.* Ontological Semantic Technology for Detecting Insider Threat and Social Engineering.//Submitted to NSPW 2010
37. *Saussure F. de.* Cours de linguistique générale, ed. C. Bally and A. Sechehaye, with the collaboration of A. Riedlinger.//Lausanne and Paris: Payot 1916. English translation by W. Baskin, Course in General Linguistics.//Glasgow: Fontana/Collins, 1977.
38. *Шайкевич А. Я.* Распределение слов в тексте и выделение семантических полей. //Иностранные языки в школе. Вып. 2. М., 1963.
39. *Taylor J. M. and Raskin V. Fuzzy Ontology for Natural Language.//Submitted to NAFIPS 2010,*
40. *Taylor J. M., Hempelmann C. F. and Raskin. V.* On an automatic acquisition toolbox for ontologies and lexicons.// Proceedings of ICAI '10.//Las Vegas, NE, July 2010.
41. *Trier J.* Der Deutsche Wortschatz im Sinnbezirk des Verstandes.//Heidelberg: Winter, 1931.
42. *Turing A. M.* Computing machinery and intelligence.//Mind *54*, 443–460, 1950.
43. *Weisgerber J. L.* Das Gesetz der Sprache als Grundlage des Sprachstudiums.//Heidelberg: Quelle und Meyer, 1951.
44. *Wilks Y. and Brewster C.* Natural Language Processing as a Foundation of the Semantic Web. *Foundations in Web Science 1/3–4,* P. 199–327, 2006.