From Wittgenstein to Ontology Engineering

A research framework for ontology-driven AI

A framework for future research into Ontology-driven AI, based on analytical language philosophy and cognitive semantics

Erik David Johnson

www.erk.dk

Abstract

This thesis establishes how the field of natural language processing (NLP), which is related to artificial intelligence and linguistics, concerned with human-computer interaction, seems to have stagnated, in its strong preference towards statistical methods based on the lower levels of language analysis. As a possible solution, it proposes an alternative approach, which instead has the highest level of language analysis – world knowledge – as its outset. A strategy towards endowing NLP systems with world knowledge through ontology engineering, based on an analysis of the works of language philosopher Ludwig Wittgenstein as related to Lakoff & Johnson's cognitive semantic theory on the metaphorical systematicity of language, is suggested and tested in a proof-of-concept. The result is a theoretical framework for further research into the development of a top-level ontology for use in high-level NLP systems, based on the cognitive semantic and language philosophical premise established in this thesis, which also includes suggestions for a computational model in the form of artificial neural networks, contextual segmentation using Lakoff's theory on Frames and metaphor deconstruction and representation using Fauconnier's Mental Space and Blending Theory.

Table of Contents

1 Introduction

Natural language processing (NLP) is a field of artificial intelligence and linguistics concerned with human-computer interaction (HCI), using natural languages.[1] Though much progress has been made in areas such as machine translation and other statistical approaches to NLP, less can be said about advancements within those areas involving high level NLP systems, such as dialogue systems and HCI-interfaces based on natural language use.[2] In many ways, these advances are halted by the need for NLP systems to relate content to human perceptions and representations of the world – in other words, the need for systems to incorporate actual 'world knowledge'.

As a relatively new field in computer science and artificial intelligence (AI), 'ontology engineering' sets out to structure a representation of the world or a limited domain within the world. In ontology engineering, an ontology is simply a structural framework that organises information by defining the type of objects or concepts that exist, as well as their properties and relations.[3] Distinctive to ontology engineering, this structural framework is usually aimed at high level AI applications and is often inspired by abstract, human cognition. In this way, ontology engineering just might have the potential to take important steps towards solving the problem of world knowledge in NLP.

1.1 Problem Statement

This thesis aims at researching the question of how ontology engineering can be used to endow NLP systems with world knowledge, based on analyses of human cognition and language. To be specific, it attempts to provide a possible answer to the following question:

How can analytical language philosophy, AI and cognitive semantics meet in an attempt to answer basic questions on how we best go about structuring and representing the world in relation to human cognition and natural language, through ontology engineering aimed at high level NLP systems?

1.2 Methodology

• ***Presentation***: To provide adequate context for the central term 'world knowledge' and illustrate a propensity towards low level, statistical language analysis within the field of NLP, the different levels of language analysis will be presented, following an example involving machine translation.

[1] For an explanation of natural and artificial languages, please see chap. 2.
[2] This view of the author of this thesis is also shared by NLP authority Prof. Bolette S. Pedersen from the Copenhagen Center for Language Technology (CST); see chap. 12.2 for further information.
[3] Gómez-Pérez, Fernández and Chorcho, p. 1-9.

• ***Analysis***: Then the two major works of the language philosopher Ludwig Wittgenstein are analysed and held up against the field of NLP to illustrate the necessity of incorporating world knowledge in high level NLP systems, and to suggest certain criteria that need to be met.

• ***Analysis***: With this outset, using cognitive semantics theory and other similar, theoretical perspectives on human thought and language, a new approach to developing an ontology for use with high level NLP systems requiring some form of world knowledge, is introduced.

• ***Experiment***: Hereafter, the suggested theoretical approach is applied on a more practical level, in the form of a limited proof-of-concept, which is implemented and documented using a modern programming platform.

• ***Discussion***: Finally, the proof-of-concept is evaluated, followed by an overall discussion of the derived approach and its possible implementations, which also includes an evaluation of the results and general applicability of the given approach within the fields of NLP and ontology engineering.

The report will be written in English, and use English for all its language analysis purposes.

2 A Short Introduction to Natural Language Processing

In this chapter, a brief explanation of the difference between natural and artificial languages is followed by a quick look at the basic functionality of statistical machine translation systems, which is used to exemplify a general propensity towards low level, statistical language analysis within the general field of NLP.

Usually, when we think of the way in which we use language, we are heavily inclined towards a view of it as an entirely human, interpersonal means of communication, associated with a variety of everyday uses and social contexts. However, a basic distinction, which is made in this thesis, is the one between ‘natural language’ and ‘artificial language’.

A natural language is a human language, such as English or German, whose rules and form have evolved into current usage, as opposed to an artificial language, the rules of which are prescribed prior to its construction and use, as is the case of any computer- or programming language, as well as that of the notation used in logic and mathematics. NLP systems are thus computer programs, or systems, that try to handle actual, natural language in a variety of ways involving different approaches and levels of complexity.[4]

[4] Jurafsky & Martin, p. 1-18.

One of the most popular, successful and widely used examples of NLP is statistical machine translation (SMT), which performs automated translation from a source text in one natural language into a target text in another natural language. In its simplest form, SMT provides simple substitution of words in one language into another, but since this usually yields rather poor results, a large corpus of statistical data is most often utilized.

This data constitutes an extensive, parallel corpus of translations from the source-language into the target-language. The basic technique of translation is then based on evaluating statistical probabilities when matching the input data against the content of the parallel corpus. We could say that given a source-sentence f, we want the output-sentence e that maximizes $P(e|f)$ (the probability of *e* given *f*); this being the most likely translation.[5]

This can be written as: $\underset{e}{\mathrm{argmax}}\ P(e|f)$

What this means is that to find out the probability of a sentence in the target-language, given a sentence in the source-language, we simply count occurrences of translation pairs in the parallel corpus. Of course, accumulating and maintaining enough data to match all possible sentence constructions is impossible, given the variety and scope of potential sentences occurring in natural language. Therefore, it is necessary to break sentences into subparts (words, phrases, etc.), and get counts on those subparts, to estimate the probabilities of their individual translation candidates.[6]

The interesting thing about the field of machine translation is that it in no significant way utilizes our knowledge of the different levels of language in relation to analysis of syntax, grammar, semantics, pragmatics, and so on. There are machine translation systems that to a limited extent try to do so, with rule-based approaches, that focus on lexicon, grammar, or semantics, but they have not been nearly as successful and effective as the statistically based, much more low-level approaches, even though the lack of grammatical and semantic coherence in the output of SMT systems is evident.

[5] Clark, Fox & Lappin, p. 531-542
[6] Mitkov, p. 516-518.

As Knight & Koehn from the Information Sciences Institute of the University of Southern California note in their paper on the recent findings within the field of SMT: "*the currently best performing statistical machine translation systems are still crawling at the bottom.*"[7], referring to a conception of what levels of language analysis are used in practice in SMT, as seen in fig. 2.1.[8]

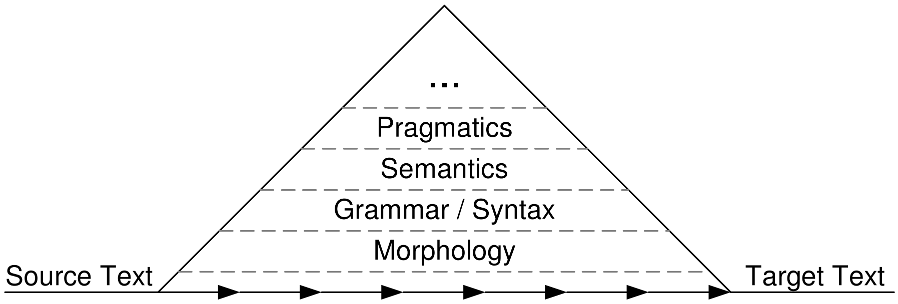


FIGURE 2.1

Though unpopular and marginalized in the field of machine translation, higher levels of language analysis are, however, necessary when it comes to some of the other types of NLP systems, such as HCI using natural language. Though different approaches to NLP system design may rely on different levels of language analysis to a varying degree, each level represents issues relevant to NLP, and as we will later see, it is also the contention of this thesis, that any optimal NLP design solution may very well need to incorporate even the highest levels of language analysis. Such a claim would then stand in stark contrast to the tendency towards low level, statistical language processing exemplified here. But let us start with sketching out an overview of what the different levels of language analysis are.

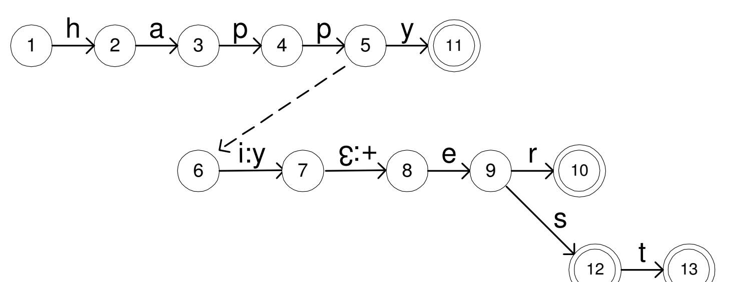
3 Levels of Natural Language Analysis

In this chapter, the different levels of language analysis are described[9], starting from the lowest and moving upwards in complexity and degree of abstraction. The purpose of providing this overview is to properly contextualize, within the functional scope of NLP systems, the highest level, which is also the focus of this thesis – namely that of 'world knowledge'. Finally, this chapter features a short rendition of one of the greatest overall challenges to NLP known as 'the problem of natural language understanding' (sometimes called 'the problem of world knowledge')[10], which is also specifically related to the issue of world knowledge, and to which it is also the ambition of this thesis to provide an additional measure of perspective.

[7] http://people.csail.mit.edu/koehn/publications
[8] This illustrates the way the level of analysis stays at the lowest level instead of going along the *sides* of the pyramid.
[9] Allen, James, p 10.
[10] Zadeh, Lofti, p. 164-166.

3.1 Morphological Analysis

Natural language systems with larger vocabularies often face a serious problem in representing the lexicon[11] of the natural languages which they handle. One way to address this problem is to pre-process words into sequences of morphemes, taking advantage of the underlying systematicity of language at the morphological level. Usually, the root form of a word is combined with different possible suffixes to reduce necessary entries into the lexicon, as in: {[eat]+[-ing, -s, -en], [ate]}.[12] These patterns of word formation can then later help model the morphological rules of language users.



FST for the forms of *'happy'*

FIGURE 3.1

One popular model for doing this is based on 'finite state transducers' (FST), which are used to transform the surface form of words into a sequence of morphemes. Fig. 3.1 illustrates a simple FST for the different forms of the word "happy":[13] The same structure could be used for the words 'crazy', 'lovely' and so on. An entire lexicon encoded as an FST then only has to define the different prefixes and suffixes once, and then let the root forms of different words point to them.

3.2 Grammatical and Syntactical Analysis

With a lexicon of well-formed words, the grammatical label of these words in the language instance being analyzed can now be assessed by different means. One popular, automatized approach is POS-*tagging*[14]. Though there are several strategies and design variations, the common basic functionality most often consists of trying to determine grammatical word labels and resolve grammatical ambiguities, by using statistics or rules together with a large corpus of pre-processed texts[15], where each word in each text has already been assigned the correct, grammatical label.

---

[11] A lexicon, in the context of NLP, is a structure used to store the possible, grammatical categories for each word (such as [man: NP] and [baby: NP, VP]…), while, as we shall see, also ordering these in accordance with morphological rules and sets which are realized through the generative capacities of the lexicon.
[12] Allen, p. 70.
[13] Allen, p. 71.
[14] POS-tagging is short for 'part-of-speech tagging', where 'part-of-speech' is the grammatical label or category of the word in question, such as verb, noun, adjective, etc.
[15] The Brown Corpus, which contains 500 samples of English language text, totalling roughly about two million words, is a good example of such a corpus.

Though POS-tagging precision levels very rarely exceed 90 percent[16] and are easily much lower due to difficulties such as disambiguation, we are at least able to approximate the grammatical category and forms of the individual words which are analysed. This is an important step towards yet another, higher level of language analysis; namely that of grammatical *parsing*, where we identify the grammatical structure of a sentence[17] using a 'grammar' – a formal specification of the grammatical structures allowable in the natural language handled by the NLP system in question.[18] Let us look at a simplified example, where our NLP system needs to be able to make a grammatical analysis of the sentence "Billy read the book". A tree representation would look something like fig. 3.2-A, while fig. 3.2-B shows a simplified grammar with the necessary specifications.

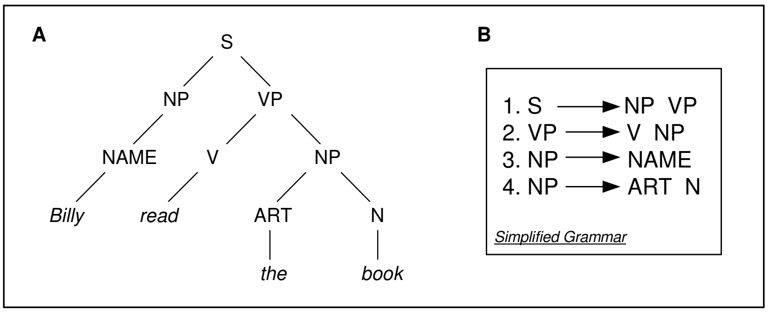


FIGURE 3.2

The sentence 'S' is broken down into its grammatical constituents in accordance with the specified rule-set of the simplified grammar.[19] The connection between 'NAME' and 'Billy', 'V' and 'read', 'ART' and 'the', and 'N' and 'book', we have from the lexicon. As long as the words used are defined in the lexicon, this extremely simple grammar is already able to handle a variety of sentences like "Susan hates carrots", "Joe is driving the car" and so on. From both an analytical and generative outset, 'Syntax' then has to do with the rules that apply to how grammatical constituents form valid phrases/sentences through different patterns, without reference to their actual meaning.[20]

Unsurprisingly, when an entire natural language is to be supported by a single grammar, the level of complexity and the sheer number of sources of ambiguity and rule-exceptions is overwhelming. Often sentence-constituents are themselves sentences, calling for a degree of recursion and complicated transition networks[21], and issues such as number agreement and the like have to be handled by additional features that also extend into the lexicon, and all that is really just scratching the surface of the problems and solutions involved when trying to build a usable grammar for an NLP system which needs to be able to perform grammatical and syntactical analysis.[22]

---

[16] Allen, p. 195.
[17] In addition, a dialogue system would also use the derivations of the grammar to *generate* valid sentences as output.
[18] Allen, p. 41-44.
[19] Allen, p. 42.
[20] Such rules are essential, if the NLP system in question is to have any language *generative* abilities.
[21] The term 'network', in practice implies the use of an advanced and complex data-structure.
[22] Allen, p. 44-75.

3.3 Semantic Analysis

This is the level where most people think that the actual *meaning* of a sentence is determined. However, when observing how computational semantics works in NLP, it quickly becomes clear that the grammatical/syntactical analysis of an input sentence already has provided a form of meaning which integrates into the semantic analysis procedure.

An example of this is a semantic representation in the form of higher-order logic[23]. Here each sentence constituent in a syntactic tree, such as the one seen in fig. 3.2, is mapped into a corresponding expression in a system of higher order logic. For example, the sentence "*John sings*" would look something like:

```
λx[sings(x)](John) (=sings(John))
```

When dealing with more complex sentences which include sub-phrases, the translations into higher-order logic are combined through function-argument application at successively higher phrase levels to generate representations for each intermediate phrase structure until an interpretation for the highest sentence is obtained.[24]

It is worth noting that not only does this and other current theories on semantic representation rely on the syntactic disambiguation[25] of sentences to have already taken place on its lower level, but these two levels also integrate in a way that clearly shows that their functional design is based on the notion that lower levels of language analysis define the upper ones.

As Shalom Lappin, Professor of Computational Linguistics at King's College London, notes in his contribution on semantic representation in the 'Oxford Handbook of Computational Linguistics', each syntactic unit can be seen as two-dimensional, with its syntactic elements also holding semantic information – a view he terms the "*Parallel Correspondence Model (PCM) of the relation between syntax and semantics…*"[26] and which he finds is incorporated in most, current theories and applications of formal grammar.

---

[23] Mitkov, p. 93-95.
[24] Mitkov, p. 99-102.
[25] The *syntactic* disambiguation of the word 'file' would then be to determine whether it is a verb or a noun. The semantic level of analysis will, however, still need to perform *semantic* disambiguation to determine whether it refers to a folder for storing papers, or a tool to shape one's fingernails, or a line of individuals in a queue, and so on.
[26] Mitkov, p. 99.

3.4 Local Discourse Context and Pragmatic Analysis

While the semantic level is occupied with *meaning* as such, this is still meaning which is out of context, or context-*independent*. Pragmatics and local discourse context analysis study the context-*dependent* meaning of language instances.

For local discourse context analysis in particular, this could be the way in which the immediately preceding sentence affects the next in terms of anaphoric reference with respect to pronouns[27], or discourse/text structure recognition, which determines the functions of sentences in the text, which, in turn, adds to its meaningful representation.[28] For example, newspaper articles can be deconstructed into discourse components such as: lead, main story, previous events, evaluation, attributed quotes, expectation, etc.[29]

Further contextualizing language analysis is pragmatics, which has as its goal to be able to read extra meaning into the language instance, based not on its own content, but on its overall context. A good example of this is speech acts, such as order, question, assertion, request, etc. This level of analysis clearly requires at least situational/domain-specific knowledge about the world. For example, as a guest at a friend's party one might ask the question "Is it me, or is it kind of hot in here?" Syntactically it is a question, but in terms of being a speech act, it is a polite request for the host to take measures to lower the indoor temperature, and a dialogue-oriented NLP system would do well to recognize this.

Current progress in computational pragmatics limits speech acts to what is known as 'dialogue acts'. The difference between speech acts and dialogue acts is that the scope of dialogue acts is strictly limited by the domain in which they occur, as well as in terms of their specific task orientation. The rationale behind this is that analysis at this level, has to be closely aligned with analysis performed at lower levels, so once again we see that low-level solutions have a tendency to shape upper-level design. As Leech and Weisser put it in their recent article on 'Pragmatics and Dialogue':

"Dialogue act interpretation, in such terms, has to depend substantially on the information derived from lower decoding levels, including lexical, and syntactic decoding, although contextual information, in terms of what dialogue acts have preceded, also plays a major role."[30]

---

[27] For example, what does 'he' or 'it' refer to?
[28] Allen, p. 429-430.
[29] Mitkov, p. 112-133.
[30] Mitkov, p.150.

The question is, when speech acts and other pragmatic concerns, which are crucial to establishing the intended meaning of language instances, so obviously require contextual world knowledge – how can this only semi-contextual, reductionist, low level-analysis approach to their interpretation and pragmatic analysis in NLP today as a whole, ever hope to reflect natural language freely? The answer is simply that it can not. As Leech and Weisser put it, as they elaborate on the above notion: "*We are here dealing with an area of computational linguistics which is still under development.*"[31]

3.5 World Knowledge and The Problem of Natural Language Understanding

With the language analysis level concerned with discourse and pragmatics showing so little progress, even under such restricted conditions, it might seem untenable to venture even further and turn one's attention to the next and final level of natural language analysis – namely that of world knowledge. However, as we shall soon see, it is entirely possible that the highest level of language analysis might serve as a better outset when attempting more ambitious NLP system design with the potential to exceed current NLP limitations.

But first, let us try to establish a definition of the term itself. In his account of the different levels of language analysis, in his popular and recognized work on natural language understanding (NLU), Prof. James Allen defines world knowledge as *"…the general knowledge about the structure of the world that language users must have in order to, for example, maintain a conversation. It includes what each language user must know about the other user's beliefs and goals."*[32]

As such, the boundary between pragmatic analysis and world knowledge may seem rather unclear, in that pragmatic analysis relies so heavily on these beliefs and goals. Nonetheless, world knowledge as a level of language analysis, needs to be given separate attention, since there are two classes of knowledge with different modes of application. One is the situation or immediate context of the language instance, i.e. specific knowledge about the current situation in which the linguistic instance is occurring. The other is general knowledge about the world and its structure.[33]

Although researchers agree that high-level NLP systems such as dialogue systems require world knowledge to work properly, little progress has been made in this area. Most work concentrates on techniques for generating simple expectations regarding actions, goals and causation; matching interpretations of input sentences. As Allen also notes in his closing comment in his chapter on how to apply world knowledge in NLP: "*Significant work remains to be done before these techniques can be applied successfully in realistic domains.*"[34]

---

[31] Mitkov, p.150.
[32] Allen, p. 10.
[33] Allen, p. 391.
[34] Allen, p. 495.

Perhaps the difficulties in treating language as *meaning* instead of letters and numbers in NLP, stems from the inherent difficulties in finding a precise and set form of the underlying systematicity of natural language, often referred to as 'the problem of natural language understanding' – as mentioned, also sometimes called 'the problem of world knowledge'.[35] Basically, the trouble is that because world knowledge is expressed in natural language, which in turn is a system for describing human perceptions, *and* because human perceptions are intrinsically imprecise, natural language is itself inherently imprecise.[36]

Overcoming the scepticism expressed by this problem thus requires us to uncover the actual way in which natural language organizes our perceptions.[37] It is the contention of this thesis that the highest level of language analysis can be seen as a core of language analysis that affects all outer (lower) levels, and that it in this way can be seen as the issue that must be dealt with *first*, in order to know how to approach lower layers, when attempting an optimal NLP design strategy. Fig. 3.3 illustrates this idea:

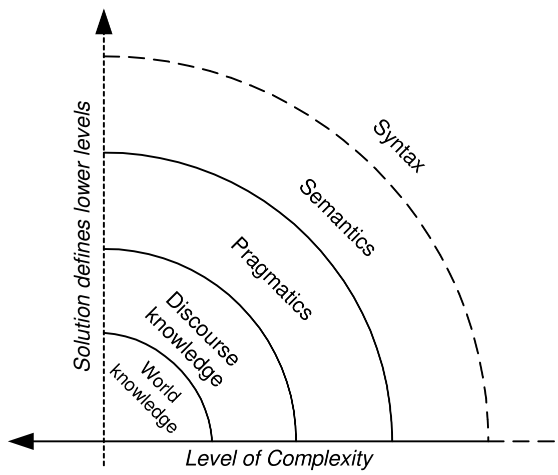


FIGURE 3.3

In stark contrast to this view, the majority of the levels of language analysis reviewed in this chapter showed a pervasive propensity towards basing their design on the lower levels of language analysis. It is not difficult to understand why this is so. As was also seen in the example with machine translation, the statistical or low-level approaches to NLP have been the ones with the earliest and greatest success – yielding concrete and somewhat usable results. As such, they have come to form the basis for further development into higher levels of language analysis.

[35] Zadeh, p. 164-166.
[36] Reusch, p. 2.
[37] A notion which seems overlooked in many areas within the field of NLP, and which will be central to the ideas expressed in this thesis.

The problem with this general development within the field of NLP is that such an approach could be potentially flawed in ways which could result in real progress being unable to move past a certain point. Instead, one could speculate that starting with the highest level of language analysis, which also represents the highest level of abstraction, and then working downwards successively in order to reach a working solution, would improve the chances of the developed NLP architecture to encompass the complexity of attaining full natural language understanding.

Following this line of thought to its conclusion, this thesis will contend that it is entirely possible that an investigation into the issue of world knowledge should preclude or at least characterize and delineate any work done on lower levels of language analysis, in order to entertain hopes of reaching an optimal NLP solution.[38] Another important implication hereof, would then be, that to arrive at an optimal solution, *any* type of NLP system would need to include and utilize *all* levels of language analysis. Much in accordance with this notion, it is already often argued that the true success of even machine translation, which we recall uses only the lowest levels of language analysis, requires the problem of natural language understanding to be solved first.[39]

Demonstrating the correctness of either approach to natural language analysis and NLP is of course impossible, in that the problem of natural language understanding remains unsolved.[40] Currently, it has not been possible for an NLP system to bridge the gap between language instances and actual meaning, as Th. R. Hofmann also points out in his book 'Realms of Meaning – An Introduction to Semantics', in commenting on the apparent 'gap' extending between the syntax level of language analysis and *meaning*; a gap which "*…reflects the fact that we still do not know very well what should really be there.*" As he also notes, it is a complex area: "*…what we are inevitably discussing are the elements of the human mind, the elements of thinking. It is here that linguistics, psychology, logic, philosophy and even anthropology and sociology meet and overlap.*"[41]

Fully transcending the bridge between language and meaning is of course far beyond the limitations and scope of this thesis. Instead, it is the author's ambition to suggest one possible approach, which will have the level of world knowledge as its *outset*, and which thereby may have the potential to lead to new results in the field of NLP, should it be pursued further through comprehensive, additional work and research. The basis for this approach will be an investigation into the intricacies of world knowledge through language philosophical analysis.

---

[38] An example could be a dialogue system able to pass the renowned 'Turing Test', which holds that if a judge using a chat interface cannot reliably tell the dialogue system from a human, the dialogue system is said to have passed the test.
[39] Hutchins & Somers, p. 8.
[40] In fact, it is often hypothesized that NLU is an 'AI-complete' problem, i.e. that it represents difficulties equivalent to solving the central artificial intelligence problem – making computers as intelligent as people, involving computer vision, NLU, and dealing with unexpected circumstances while solving any real world problem.
[41] Hofmann, p. 11-15.

4 Analytical Language Philosophy and World Knowledge

This chapter establishes the notion of regarding the issue of world knowledge as a language philosophical concern, by introducing the four main problems of the field of analytical language philosophy. An approach to a language philosophical investigation of world knowledge, based on an analysis of the works of the influential language philosopher Ludwig Wittgenstein, is then suggested.

When viewing world knowledge, not simply as a general term, but as a specific language analysis problem domain, it is apparent that there are issues pertinent to knowledge organization and structuring, which an NLP system that includes real knowledge in its design should address, and which are partly philosophical in nature. This is well illustrated by the compelling relevance of the four central problems of analytical language philosophy[42] to the issue of world knowledge:

- *The relationship between language and reality.*
- *The nature of meaning.*
- *Language use.*
- *Language cognition.*

The idea is thus to treat these problems as held up against the issue of world knowledge in NLP. Though rarely seen in literature on world knowledge in NLP[43], a language philosophical approach such as this, should not be completely without merits. However, taking on the four central problems of analytical language philosophy in their entirety is far too great a task. Instead, this thesis will use the findings and works of a single language philosopher to shed new light on the matter. In this way, the author needs only provide an analysis of that philosopher's treatment of those problems, and then relate the results to the issue of world knowledge in NLP.

One of the most influential, modern language philosophers to have lived is Ludwig Wittgenstein (April 26, 1889 – April 29, 1951).[44] Uncovering the nature of meaning, language use and the relationship between language and the world was of prime importance to him – an obsession perhaps, and his philosophical work was groundbreaking within these areas.

---

[42] Martinich, p. 1-29.

[43] Though there are of course exceptions to this general tendency – a good example being John F. Sowa's comprehensive work: "Knowledge Representation – Logical, Philosophical, and Computational Foundations".

[44] Alternatively, one could look into Heideggerean ontology or perhaps the work of Bertrand Russell, who was also an important source of inspiration for Wittgenstein.

A key issue regarding Wittgenstein, is that his life and work can be divided into two distinct, *opposing* periods – each represented by one of his two major works. The first is 'Tractatus Logico-Philosophicus' (1921), from hereon referred to as 'Tractatus', and the other his posthumously published 'Philosophical Investigations' (1953), which in 1999 was ranked as the most important philosophical work of the 20'th century, with Tractatus rated as number four.[45] Though some ideas and theories are shared between these two distinct philosophical works, they present completely opposing views on those problems central to analytical language philosophy.

The following two chapters each present an analysis of one of these two works and their language philosophical implications. Their respective conclusions will be held up against the field of NLP, in order to attempt to provide new insights into how one best can approach the issue of world knowledge in NLP. In this way, the lessons learned by Wittgenstein, which prompted him to shift his language philosophical stance so dramatically, can hopefully be applied to current strategies for NLP design that includes world knowledge.

5 Tractatus Logico-Philosophicus and NLP

In this chapter, an overview of the ontological structure of Tractatus is used to form the basis for an account of Wittgenstein's picture theory, and its logical conclusion is then held up against issues relevant to the field of NLP. However, before proceeding with the textual analysis, a short comment on the unusual literary style of Tractatus, would be in order. Notoriously inaccessible to casual readers, Tractatus employs a rigorous and concise literary style which contains almost no arguments. Instead, hierarchically numbered, declarative statements, which are meant to be self-evident, are organized around 7 basic propositions. Each basic proposition then has sub-levels which elaborate and comment on it, and these may also have their own sub-levels, and so on.[46]

5.1 The Basic Ontology of Tractatus

In the beginning of Tractatus, Wittgenstein starts out by stating that "The world is everything that is the case.", and in proposition 1.1 and 1.2 that the world is and can be divided into *facts* – not things. These propositions are to be viewed in the light of proposition 2, which explicates that all which is the case is the existence of *atomic facts* or 'sachverhalten'. The central issue then becomes defining an atomic fact, which is what Wittgenstein elaborates on in proposition 2.01, as he states that "An atomic fact is a combination of objects (entities, things)." In this way the four terms 'the case', 'facts', 'atomic facts', and 'objects' are linked together, to show that the world is an object-world where it is the relations between the objects, i.e. the atomic facts, which in turn are the facts that are the case, that make up the world. [47]

---

[45] Lackey, p. 329-346.
[46] For an example of this and a view of Tractatus in full, see also the folder [Full Texts] on the enclosed CD.
[47] Wittgenstein - Tractatus, p. 30-31.

Wittgenstein then turns his full attention towards ‘objects’ and the way they form atomic facts, since an analysis of any single object would uncover the possible relations it could enter into; not saying that an object can be known only by the external qualities that its relations to other objects constitute, for as Wittgenstein says in proposition 2.01231: “In order to know an object, I must know not its external but all its internal qualities.” Rather, it is its internal qualities – its form – that define which atomic facts it can partake in constituting.

This is what Wittgenstein means, when he in proposition 2.0122 explains that an object’s independent connection with an atomic fact comes from the dependence implicated by the way in which it *must* enter into some connection with an atomic fact. Even if an object were to be defined by its internal qualities, independently of its possible connection to an atomic fact, it is not possible to think of any object apart from the possibility of its connection with other things, as Wittgenstein also says in proposition 2.0121. It is thus the objects that make up the substance of what exits, while it is the atomic facts, i.e. the *relations* of those objects, which are variable and unfixed. [48]

Further detailing his account of objects in proposition 2.02, Wittgenstein claims that “The object is simple.”[49] This statement I see as connected with proposition 1.13, which states that “The facts in logical space are the world.”[50] The ‘logical space’ that Wittgenstein refers to is language, reduced to its logical form. Already, we see Wittgenstein’s early view of the relation between language and the world as being equivalence between the smallest, logical, irreducible constituents[51], and the smallest constituents of the world – the simple objects.[52] As Wittgenstein is also saying in proposition 2.021-2.0212, the object world is a precondition for us having a meaningful language.

5.2 Wittgenstein’s Picture Theory of Language

With the basic ontology in place, we are ready to take a look at Wittgenstein’s picture theory in Tractatus. In proposition 2.1 and on, Wittgenstein claims that we construct *pictures* of facts. Central to Wittgenstein’s picture theory, is that a picture has a number of *elements*, and that these elements represent objects, as is also stated in proposition 2.13/2.131. These elements can be seen as simple names.

[48] Wittgenstein - Tractatus, p. 32-33.
[49] Wittgenstein - Tractatus, p. 35.
[50] Wittgenstein - Tractatus, p. 31.
[51] The reduction implied here is not grammatical, but one that brings language to a logical form.
[52] An important point of criticism has been that Wittgenstein himself never bothers to give examples of these smallest constituents of language and the world.

Reference in Tractatus thus amounts to the relation between a simple name and a simple object. The name *is* the meaning of the object, and it gets its meaning from the object. The way in which a picture then depicts a fact is brought about in that the elements of the picture are ordered in the same way as the objects of the atomic fact which it depicts. As we recall, facts are simply those atomic facts which are the case, and atomic facts are configurations of simple objects.[53]

In proposition 2.141, Wittgenstein goes on to say that "The picture is a fact." I believe that this proposition, along with propositions 2.1511, 2.1512, 2.15121, and 2.1515, describe the essence of Wittgenstein's view of the relation between language and reality. The way in which *language* represents the world/reality is itself a fact. The cross point between reality and language is that they are both *facts*. As proposition 2.15121 and 2.1515 also emphasize, language 'touches' reality by way of this structural likeness.

The meaning of a sentence is in this way its truth conditions, which means that one understands what the world were to look like in order for the sentence to be true. However, it is also important to notice that it is the elements of the picture that represent the objects and not the other way around, as is also clarified in proposition 2.1514: "The representing relation consists of the co-ordinations of the elements of the picture and the things."[54] When this notion is further subjected to the constraints of logic, as we will see in the following, these truth conditions are the very nature of language meaning as portrayed in Tractatus.

Wittgenstein elaborates on his definition of a 'picture' in propositions 2.182 and 2.19, which state that any picture is a logical picture, which can depict reality; not saying that any picture is in accordance with reality / is true. It is a logical picture which constitutes a meaning, in that it depicts a *possible* reality.[55] Whether it is *true* or not, however, depends on whether it agrees with reality by depicting a *positive* fact. The way in which we examine this issue, as Wittgenstein says in proposition 2.223, is by simply comparing the picture with reality.[56]

[53] Wittgenstein - Tractatus, p. 38-39.
[54] Wittgenstein - Tractatus, p. 38-41.
[55] For ex., there is nothing illogical about the sentence: "Finland is the capital of Spain." It is merely a *negative* fact.
[56] Wittgenstein - Tractatus, p. 40-43.

In proposition 3, Wittgenstein's idea of a 'picture' is explicated in that we are told that the logical picture of the facts is the *thought*. Wittgenstein's belief in logic as encompassing both language and the world is seen in proposition 2.182, where he states that we can think nothing illogical in our depiction of reality: "Every picture is *also* a logical picture."[57] Here it is important not to confuse the before mentioned distinctions between true/false and logical/illogical. We can think something untrue, but *not* anything illogical like a round square or an oblong coordinate.

In proposition 3.1, the idea of thought as a logical picture is linked together with language by Wittgenstein's saying that the thought is expressed perceptibly through the senses, in the form of sentences/propositions (language). The type of 'sentence' or 'proposition' and thereby the kind of *language* that Wittgenstein is working with in Tractatus does, however, require a certain liberation from our common and flawed everyday language, which can not always be analysed into a logical form. Language cognition (the way language relates to the mind of the language user) should therefore be viewed from within the special context that is Wittgenstein's notion of an ideal language.

To say it plainly, in Tractatus, Wittgenstein is treating language as it, in his eyes, *ought* to be – not the way it is.[58] Wittgenstein's belief in logic is evident here. He wants a perfect logical language; probably much like the one formulated in the famous work "Principia Mathematica", written by Dr. Whitehead and Wittgenstein's former tutor Bertrand Russell[59], and throughout proposition 4, 5, and their subsidiaries, Wittgenstein explores formal logic and the formal mechanisms required for a logically 'ideal' language, which is then detailed further in proposition 6 and its subsidiaries.

The conclusion of Tractatus is then that all expressions that can not be analysed into a logical form, should be *removed* from language altogether. This is precisely what Wittgenstein means when he in the seventh and final proposition of Tractatus says: "Whereof one cannot speak, thereof one must be silent."[60] A great deal of philosophical problems and most metaphysical issues are in this way 'annulled' and disregarded. For example, with a sentence like "Is there a God?", it is the impossibility of the *question* and not the answer that, according to Wittgenstein, is giving us trouble, since it can not be brought to a proper logical form, and is as thus void of meaning.

[57] Wittgenstein - Tractatus, p. 41.
[58] Wittgenstein - Tractatus, p. 45-51.
[59] Lübcke, p. 37-39 + 54.
[60] Wittgenstein - Tractatus, p. 189.

5.3 The Application of Tractatus in NLP

Now, if Wittgenstein is correct in his assumptions on language and reality in Tractatus, natural language use must be reduced and changed in a way that at first glance seems to fall out to the advantage of the field of NLP. If all an NLP system needs to be able to handle is inherently logical sentences that can be broken down to a logical form, perhaps the problem of getting from a human language to a formal, artificial language, which can be handled by computer systems, is far less overwhelming.

Also, an actual solution to the unsolved problem of natural language understanding would seem within reach, in that if we were only able to account for the way in which human thought and language depicts reality by the ordering of elements parallel to the configurations of objects that make up atomic facts, the way human perception orders information about the world into language could be mapped out as exemplified in fig. 5.1.

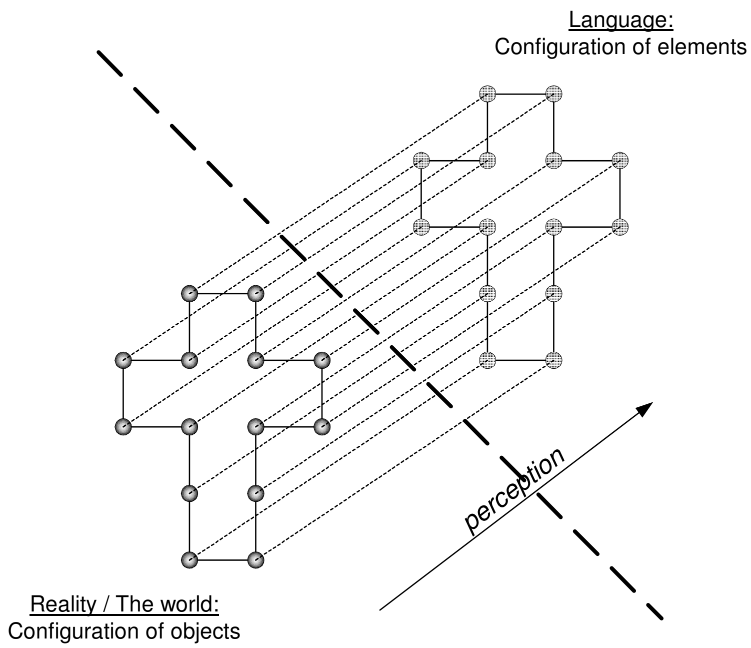


FIGURE 5.1

In fact, we need only know the simple objects whose configurations make up atomic facts, and the simple elements whose configurations make up our depictions of these atomic facts, in order to give an exhaustive account of the systematicity of not only language, but also the reality it describes; this would indeed be a giant leap towards solving the so-far insolvable issue of employing world knowledge in NLP, as well as the before mentioned problem of natural language understanding. Even a field like machine translation would benefit immensely from there being a single highest logical syntax which could form a bridge between the different kinds of natural languages, finally fully utilizing the rule-based machine translation systems in practice.

There are, however, a number of critical problems. The most pressing one, and an ever-occurring criticism of Tractatus, would be the way in which Wittgenstein never gives us any concrete examples of exactly *what* his simple objects and their corresponding picture-elements *are*. The inherent logical atomism of Tractatus certainly dictates the use of such elements, but all we really know of them is that they are simple, irreducible, logical entities. Even more mysterious is Wittgenstein's idea of simple objects which make up the actual *substance* of the world. What are these simple objects? They seem almost metaphysical in their nature, and we are never given any suggestions as to what they really are.[61] Suddenly, the usefulness of the ontology and picture theory of Tractatus to the field of NLP seems somewhat diminished.

Further complicating the applicability of the logical reductionism implied by Tractatus, is the issue of the actual scope of the resulting NLP system, because exactly how much of natural language can really be brought to a logical form? If natural language is forced into the straitjacket of a logically perfected system of semi-artificial language, humans will have surrendered completely to the terms of computers as regard to our mode of communication.

This would entail a significant departure from the modern goals of NLP, HCI and computational linguistics as a whole. Today, instead of searching for an ideal language, most areas within the field of computational linguistics, including NLP, generally strive towards treating language as a natural phenomenon which should be studied and dealt with as it is, leaving natural language as natural language, by way of ambitions to completely avoid artificial reductionism in the scope of NLP.[62] It is the NLP system that must use language on human terms and not the other way around. As we shall see, this is exactly the approach to language that the later Wittgenstein turns towards in his later work 'Philosophical Investigations', which in many ways rejects the views found in Tractatus.

6 Philosophical Investigations and NLP

This chapter provides an overview of Wittgenstein's alternative view of language, which will then be held up against the field of NLP. As we shall see, Wittgenstein modifies his view of language drastically in his later philosophical work. Not only is the ontology and picture theory of Tractatus completely rejected; Wittgenstein also determines an altogether different scope of and approach to *language*. Furthermore, the literary style of this second work is far less constrictive with a simple sequential numbering of paragraphs and chapters of variable length.

[61] Sluga & Stern, p. 88-94.
[62] Jurafsky & Martin, p. 1-18.

6.1 A different View of Language and Meaning

In ‘Philosophical Investigations’ (PI), Wittgenstein starts out by referring to the simple view of language portrayed in Augustine’s “Confessions”; namely that individual words in language simply name objects. A word’s meaning is “the object for which it stands.”[63] On the basis of this view of language, one could imagine a primitive language, exemplified in the second paragraph of PI, where person A calls out the name of an object, and person B hands it to him.

It would seem that Wittgenstein uses this simplified view of language to sketch out his criticism of the view of language found in Tractatus and in philosophy, where focus is on objects – objects that make up the world, and later on in PI, Wittgenstein also completely rejects the ontology of Tractatus, pointing to the obscurity of its mysterious simple objects and logical atomism.

Central to PI is Wittgenstein’s desire to show that there is more to language than objects and their corresponding representations in language. As he says in paragraph 3: “*Augustine, we might say, does describe a system of communication; only not everything that we call language is this system.*”[64] If meaning was simply naming, language would be fairly homogonous as a whole[65], but this does not seem to be the case. We can, however, be induced to take this simplified view of language – especially when we engage in philosophical thinking. As Wittgenstein says in paragraph 11 in PI: *“…what confuses us is the uniform appearance of words… For their application is not presented to us so clearly. Especially when we are doing philosophy!*” [66] The illusory ‘uniformity’ Wittgenstein mentions, is undoubtedly a reference to the field of logic.

Wittgenstein is in this way criticising the uniqueness of reference between an object and its name, saying that even though it is necessary to know the names of the objects of the world for them to be part of language, there is more to language than naming things. Naming is merely *one* aspect of our multifaceted language. It is part of a greater whole. As Wittgenstein says:

“-It is interesting to compare the multiplicity of the tools in language and the ways they are used, the multiplicity of kinds of word and sentence, with what logicians have said about the structure of language. (Including the author of the Tractatus Logico-Philosophicus.)” [67]

[63] Wittgenstein - Investigations, p. 2.
[64] Wittgenstein - Investigations, p. 3.
[65] As we also saw when looking at the *initial* implications of the ontology and picture theory of Tractatus for NLP.
[66] Wittgenstein - Investigations, p. 6.
[67] Wittgenstein - Investigations, p. 10.

Though reference in Tractatus is far more complex than the Augustinian definition of language that Wittgenstein uses as an outset for his criticism of his own former work, Augustine's definition is an excellent caricature of the faults of Tractatus, and shares with it its misguided focus on reference between names and objects, as a juxtaposed example.[68]

6.2 Language-games: Meaning as Use

What is then the totality of language? The answer to this question, and a central aspect of the view of language presented in PI, is 'language-games': "*I shall also call the whole, consisting of language and the actions into which it is woven, a 'language-game'.*"[69] Contrary to the view expressed in Tractatus, Wittgenstein now sees language as a tool for social interaction, the primary function of which is to function as an interpersonal means of communication.

Language-games and language as such is to be seen as something dynamic, and the rules for each language-game depend on the situation and context within which it unfolds. Language is incalculable and constantly changing. An illustrative metaphor of this, which Wittgenstein utilizes in paragraph 18 and 107 of PI, is that language can be seen as a city. Sometimes houses are torn down, sometimes new houses are built, and the city grows and renews itself constantly, and as we are *in* the city, we cannot oversee it completely – the same being true of language. [70]

The critical change since Tractatus is foremost that Wittgenstein's focus has shifted from logical, reduced language to *everyday*-language. Logical statements are no longer his central point of attention; there are many other relevant language-games. This also entails that the meaning of a sentence no longer is its truth-conditions as were the case in Tractatus. Instead, the meaning of a sentence is the way it is *used* in practice: "*…the meaning of a word is its use in the language.*"[71]

6.3 Rule-following, life-form and the Acquisition of Language

A relevant question is now: How are we able to partake correctly in particular language-games? The answer is 'rule-following', which is also a central term in PI. Language-games are rule-bound activities where differing interpretations of rules are possible from different perspectives and in different situations. Rules are in this way relative to our understanding and experience.

[68] Sluga & Stern, p. 15-17.
[69] Wittgenstein - Investigations, p. 4.
[70] Wittgenstein - Investigations, p. 7+40.
[71] Wittgenstein - Investigations, p. 18.

Wittgenstein exemplifies this issue with the way in which one could imagine that a pupil was learning the ‘add 2-game’, where 2 is added again and again, as in 2, 4, 6, 8, and so on. The pupil then exceeds 1000 for the first time and continues on: 1000, 1004, 1008, and so on.[72] We thought he understood the rule as we did until he went beyond 1000, but apparently he did not.[73] We might think he is wrong, but really, Wittgenstein argues, these two different patterns of playing the game could both be correct. In this way, meaning is not fixed in language, but is relative to our unique understanding.

The way we acquire language also illustrates how language-games really are the most basic mechanism of language. As we first learn to speak, the names of objects are taught to us through ostensive definition[74], but notice here that ostensive definition itself relies on our pre-existing ability to play the language-game which dictates that pointing to an object and making a sound means describing the sound of the name that describes that object. Similarly, one can not teach a child to count, but one can teach the child the ‘counting-game’ by way of example, and hope that the child constructs the correct form of rule-following.

In this way, there is not a fixed meaning hidden in the core of any individual word. To find a word’s meaning, we should, instead, travel along with the word's actual uses through “*a complicated network of similarities, overlapping and criss-crossing*”[75] as Wittgenstein puts it. Particular uses of the same word share what Wittgenstein calls ‘family-resemblances’, but they are not identical in a way that permits universal definitions. The same principle can be applied to language as a whole, saying that one can not make inferences from what is particular to what is general in language.

With the rules of language-games relying on contextual issues, with unfixed rules that are relative to our understanding and unique to each language-game, how is it then that we, nevertheless, manage to communicate and play language-games with such a high degree of success? The answer to this question, and the final central term of PI is that it is by virtue of our *common life-form,* which has certain traits that are shared across an interpersonal axis. It is our common ground and shared point of departure that permits us to construct correct forms of rule-following, such as the social context of the language-game, or our likeness in human form and culture. Wittgenstein does not go into any great detail concerning the particularities of the concept of our common form of life, but as he says in paragraph 19: “*…to imagine a language means to imagine a life-form.*”[76]

---

[72] Wittgenstein - Investigations, p. 63-64.
[73] Perhaps he thought that one should add 2 up to 1000, 4 up to 2000, 6 up to 3000, and so on.
[74] The act of teaching someone the name of an object by pointing to it while pronouncing its name.
[75] Wittgenstein - Investigations, p. 27.
[76] Wittgenstein - Investigations, p. 7.

The conclusion of PI is that the search for the ideal language is not so much an actual achievement as it is an impossible demand put forth by logicians, as he himself attempted in Tractatus. Philosophy as a whole has, through the focus on logic and curious attempts to say something concrete on the basis of an *a priori*[77] ordering of the world, strayed from the true nature of language, which is instead found in its everyday use. This misguided approach to language frustrates and confuses us; we "*have got on slippery ice*" and must get "*Back to the rough ground!*"[78] as Wittgenstein puts it.

The solution is to analyse different language-games within their given context, and thus reveal how philosophical problems can be exfoliated as mere surface-problems. As Wittgenstein says in paragraph 116, we have to *"…bring words back from their metaphysical use to their everyday use."*[79] The goal of philosophy thus becomes to act as a kind of therapeutic activity that helps us avoid philosophizing about impossible abstracts and the unattainable completion of perfect rule-systems.

6.4 The Application of Philosophical Investigations in NLP

The implications of Wittgenstein's findings in PI, and the language philosophical shift he makes from the outset of Tractatus, are both uplifting and disheartening when held up against the field of NLP. Wittgenstein's alternative view of language looks at everyday-language, meeting one of the previously mentioned, primary goals of modern NLP and many of its related fields, which is looking at language as a natural phenomenon, instead of trying to mould it into a more restricted, systematic, logical language. Unfortunately, the conclusions of PI also seem closely aligned with the scepticism expressed in the problem of natural language understanding. There is no longer an ideal language, and meaning is not fixed in language itself, but is instead relative to our understanding, implicating that we do not have a rigid ontology that permits language to describe the world in general or universal terms – leaving the question of how language relates to the world to be viewed through the subjective prism of social context.

Luckily, Wittgenstein does not leave us completely in the dark in that, even though he provides no general, logical, hidden systematicity of language that might enable an NLP system to handle natural everyday language, he does give us an important insight into the nature and workings of natural language itself, with his ideas of rule-following and common life-form. As we recall, world knowledge is general knowledge about the actual structure of the world that language users, or conversational agents, must have in common, in order to communicate with each other.

[77] 'A priori' knowledge, or justification is independent of or prior to experience (e.g.: 'All bachelors are unmarried').
[78] Wittgenstein - Investigations, p. 40.
[79] Wittgenstein - Investigations, p. 41.

For an NLP system to be able to play language-games, we must therefore try to find out how our common life-form, which enables us to form correct rule-following, is evident in the way we use language. In other words, in order for an NLP system to actually try to simulate the necessary life-form when constructing the correct form of rule-following needed for playing language-games, we must first look into the kinds of patterns in language and cognition that correlate with our common life-form.

7 Common Life-form as a Basis for Patterns of Meaning in NLP

In this chapter, the implications of Wittgenstein's language philosophical shift in terms of how best to approach the issue of world knowledge in NLP is pursued further, leading us to a rejection of some of the basic assumptions evident within the general field of cognitive science, as well as to an alternative theory of cognition, which is then held up against the later Wittgenstein's notion of common life-form.

7.1 Wittgenstein's Shift according to Cognitive Theory

To recap, earlier chapters of this thesis have shown that there is a general propensity towards low level, statistical language analysis within the general field of NLP, and argued that starting at the very highest level of language analysis, i.e. the level of world knowledge, may hold unrealized potential for overcoming the current slump in new advances within the field. As pointed out, this may even potentially be key to attaining the theoretically highest, achievable level of HCI using NLP – something which otherwise might be rendered unreachable.

The top-down approach to language analysis pursued in this thesis has then been to use an analysis of the two opposing, language philosophical stances taken in Wittgenstein's first major work "Tractatus" and his second and final, major work "Philosophical Investigations", held up against the field of NLP and the issue of world knowledge, to arrive at the rejection of a universally applicable language. Instead we end up with the notion of an extreme contextualization of the undercurrents of language use as meaning, in the form of language-games with a situational uniqueness, remedied only by the all-important notion of our common life-form.

The question is now: What do the lessons learned by Wittgenstein, which we see reflected in the language philosophical shift between his two major works, tell us about the way we should approach the issue of world knowledge in NLP? Perhaps one of the most interesting aspects of Wittgenstein's revised position on language and cognition is not so much the new alternative theories that he introduces, as it is the theories which he completely dismantles and ultimately rejects.

An important view rejected by the later Wittgenstein, is that language, in the form of symbols and their relations, describes an objective reality which is readily accessible to the human mind through the use of our senses.[80] Taking this rejection to heart has potentially far-reaching consequences in terms of how we approach the task of analysing and systematizing natural language, as well as the cognitive processes that facilitate it.

One significant implication is that an ideal approach to NLP will need to deviate completely from some very basic assumptions that traditionally have characterized the cognitive science program – namely that there is a clear division between the mind and the world, and that the human mind in this way enjoys an observer-independent existence.[81] In other words, the human mind is habitually viewed as something separate from the body which then plays no crucial role when it comes to analysing the natural language produced by such an 'disembodied mind'. This stands in direct opposition to the later Wittgenstein's notion of our 'common life-form' as something which is key to understanding natural language. It can be argued that he views the human mind and body as a crucial part of the very common ground which allows us to play language-games successfully, and which as such is bound to be evident in language.[82]

The notion of an observer-independent mind is also closely related to another dominant view within the field of cognitive science, which is that the mind can be seen as a machine or computer, which simply processes the sensory input (symbols, signs, language, etc.) provided by the senses. Often referred to as the 'mind-as-machine paradigm', or 'the algorithmic mind position', this view of the human mind considers every cognitive process algorithmic in nature in that thought is simply a matter of symbol manipulation.[83] As the Stanford Encyclopaedia of Philosophy states, in its most recent article on 'cognitive science':

"Most work in cognitive science assumes that the mind has mental representations analogous to computer data structures, and computational procedures similar to computational algorithms."[84]

[80] This positivism, which could be read into Tractatus, was taken to heart by logical positivist movements such as the 'Vienna Circle', but it is worth noting that the early Wittgenstein found the notion overly simplistic, as author R. Monk also describes it in his detailed biography of Wittgenstein: 'Ludwig Wittgenstein – The Duty of Genius', p. 243-244.
[81] http://en.wikipedia.org/wiki/Cognitive_science
[82] The question of *how* the human mind and body is evident in natural language is treated in more detail in chap. 8.
[83] Lakoff (1990), p. 338-352.
[84] http://plato.stanford.edu/entries/cognitive-science

Viewing thought and language as mere symbol manipulation taking place in a disembodied mind, is exactly what the early Wittgenstein envisioned in Tractatus, with his picture theory, logical atomism and unwavering belief in logic itself. However, the later Wittgenstein saw language as much more than the relation between simple names and simple objects, as he shifted his stance on truth in language to be defined as the contextual *use* of a language instance, instead of its objective truth conditions. In this way, the later Wittgenstein rejects, not only the notion of a disembodied mind, but also the idea of the human mind defined as simply manipulating symbols. Meaning is not found in the constellations of the symbols that make up natural language, but in their highly contextualized *use*.

The sketched out implications of Wittgenstein's language philosophical shift go against some very basic objectivist views and assumptions on how the human mind, natural language and the world relate to each other, evident within the general field of cognitive science. Returning to the question of how these conclusions are to guide the way we undertake the issue of world knowledge in NLP, it seems that the top-down approach to language analysis pursued in this thesis calls for an alternative theory of cognition.

7.2 Embodied Cognition

A good candidate, which is compatible with the language philosophical stance taken by the later Wittgenstein, is the theory of embodied cognition. Drawing largely upon research carried out within the last 10 years in the fields of linguistics, cognitive science, AI, robotics and neurobiology, this alternative theory of cognition maintains that the nature of the human mind is, to a great extent, determined by the form of the human body. As it is stated in the 'Embodiment Thesis':

"Many features of cognition are embodied in that they are deeply dependent upon characteristics of the physical body of an agent, such that the agent's beyond-the-brain body plays a significant causal role, or a physically constitutive role, in that agent's cognitive processing."[85]

This means that all aspects of cognition are somehow shaped by human, bodily features such as our perceptual system, the intuitions that underlie our ability to move, activities and interactions with and within our environment, as well as the naive understanding of the world that is built into both the body and the brain.[86] This alternative theory of cognition not only challenges dominant views in philosophy of mind and cognitive science, which consider the body peripheral to understanding the nature of mind and cognition itself; it even goes on to question whether the human mind can really be said to be restricted to the brain at all.

[85] http://plato.stanford.edu/entries/embodied-cognition
[86] Lakoff (1990), p. 12-15.

The theory of embodied cognition displays great potential when it comes to interpreting the later Wittgenstein's notion of our common life-form, as the very thing that allows us to successfully play language-games, even though the rules of these language-games are both unique and relative to our understanding. As human beings, we have the same body parts and perceptual systems and we are also bound by the same physical laws when interacting with our environment. Our experiences are thus 'shaped' by these common circumstances, which allow for comparable patterns in cognition that correlate with our common life-form.

Further supporting Wittgenstein's notion of common life-form and further undermining the notion of a disembodied mind, are some recent findings within the field of cognitive neuroscience which suggest that the human brain is a priori 'hardwired' towards certain patterns of cognition related to our physicality, which influence the way we acquire and use language.

To be specific, during the 1990s a group of scientists at the University of Parma, Italy discovered something remarkable when they placed electrodes in the ventral premotor cortex of macaque monkeys. It seemed that the activation of special 'mirror neurons' in a monkey who carried out a complex action was identical to that of a monkey simply *observing* another performing that same complex action.[87] Later research soon showed the same to be true for humans performing complex actions, and functional MRI studies have reported finding areas homologous to the monkey mirror neuron system in one of the hypothesized language regions of the human brain. This, in turn, has led to suggestions that human language evolved from a gesture performance/understanding system implemented in mirror neurons, and that we are as such 'hardwired' for empathy and cooperation.

So even if the mind is indeed restricted to the brain itself, the notion of a cool, symbol manipulating algorithmic mind, would seem defeated by this recent development within the field of cognitive neuroscience. Instead, we find further evidence supporting the later Wittgenstein's notion of language-games as an innate ability more basic to our capacity for language than even the very acquisition of language itself.

Even though there seems to be both scientific basis as well as a recognized theory of cognition to support the views on language and meaning expressed by the later Wittgenstein, we still lack a clear idea of how the language philosophical lessons learned by Wittgenstein are to guide the way we approach the issue of world knowledge in NLP. The reason for this is that while we have indeed looked into Wittgenstein's notion of our common life-form and the kind of patterns in cognition that correlate with it, we have yet to determine how these patterns are evident in language itself. This will, however, be the focus of the following chapter.

[87] Rizzolatti and Craighero.

8 The Metaphorical Systematicity of Natural Language

This chapter introduces the theory of metaphor as an intricate part of our cognition and natural language usage. Along with some simple examples, the different types of metaphors and how they relate to the cognitive patterns sketched out in the previous chapter, as well as to the language philosophical outset of the later Wittgenstein, are then presented, to arrive at an account of language and cognition that allows for a degree of linguistic systematicity, without removing language from its everyday natural use.

8.1 Our Conceptual System as Metaphorical in Nature

When focusing on the cognitive and the linguistic aspects of the embodied mind thesis, two of its most important contributors must be said to be George Lakoff and Mark Johnson who wrote "Metaphors We Live By" (MWLB), first published in 1980. Not only is it considered a landmark publication among works that substantiate the notion of embodied cognition[88], it also provides some direct answers to the question of how embodied cognition is evident in natural language.

In short, Lakoff & Johnson claim to have found that most of our ordinary conceptual system is *metaphorical* in nature – a claim for which they offer a great amount of linguistic evidence. To them, the term 'metaphor' is to be understood as a metaphorical *concept* and should be differentiated from the traditional notion of 'poetic metaphor', which in turn can be seen as a mere artefact of language. These metaphorical concepts reveal themselves as subconscious conceptual structures that govern the way we think, act, communicate and live, and although generally present only on a subconscious level, they can be explored through the medium of natural language.[89]

8.2 Types of Metaphor

When presenting their linguistic evidence for these claims in MWLB, an initial examination is made of the structured and systematic way that metaphorical concepts work. 'Structural metaphors' are metaphors where one concept (often more abstract) is metaphorically structured in terms of another (often more concrete), yielding certain entailments. For example, the structural metaphor [Time is money][90] entails that time is a limited resource, which again entails that time is a valuable commodity, and so on.[91] Since it is nonsense to say that time really *is* money, this structuring is of course always *partial*. Moreover, an important feature of metaphor is that it highlights some aspects of the concept and hides others. For example, with [Time is money], time as a commodity is highlighted, while the notion of [Time as a landscape we move through] (as in: "*my 50th birthday is looming up on the horizon*") is hidden.

---

[88] http://plato.stanford.edu/entries/embodied-cognition
[89] Lakoff & Johnson, p. 1-6.
[90] An example of this metaphorical concept could be the sentence: "*We can not afford to lose more time!*".
[91] Lakoff & Johnson, p. 7-13.

Yielding further evidence to support the claim that our conceptual system is metaphorical in nature and providing an even clearer idea of how this relates to the Embodiment Thesis and Wittgenstein's notion of common life-form, are the kind of metaphorical concepts that Lakoff & Johnson call 'orientational metaphors'. Orientational metaphors are organized around spatial orientations like: up-down, front-back, in-out, and so on. An example of an orientational metaphor could be [Happy is up], which provides us with expressions such as 'Cheer up', 'I'm feeling low today', 'her spirits rose', etc.[92]

There is an internal systematicity to each orientational metaphor. An example of this provided by Lakoff & Johnson is how [Happy is up] defines a coherent system, and not just a number of isolated cases; an example of an incoherent system would be one where "I'm feeling up" would mean "I'm feeling happy", while "Her spirits rose" could mean "She became sadder".[93] Additionally, there is an external systematicity among the various orientational metaphors, which defines a certain degree of coherence among them. For example [Good is up] assigns 'up' orientation to general well-being, and is as such coherent with [Happy is up], [Health is up], etc.[94]

Besides structural and orientational metaphors, we also have 'ontological metaphors', which instead of spatialization are based on our experience of physical objects and substances (including our own bodies), and which permits us to view events, activities, emotions, ideas, etc., as entities and substances. For example, a race can be viewed as a container object, yielding expressions such as "he is out of the race" and "they are halfway into the race now".[95] Personification, where something nonhuman is seen as something human, is perhaps the most obvious type or extension of ontological metaphor, with examples like: "The facts argue against your claim", or "Death finally caught up with him".[96]

Especially with orientational metaphors, but also with ontological and structural metaphors, we seem to have a serious answer to the question of how our physical and cultural experience, which is certainly comparable to the later Wittgenstein's notion of common life-form, is evident in language in the form of recognizable patterns of meaning. Even the most basic notion of common life-form as an aspect of our mutual, experiential outset in physical likeness, is seen in natural language and in the way it conveys our understanding of abstract concepts rendered according to the internal and external systematicity of orientational metaphors.

---

[92] Lakoff & Johnson, p. 14-21.
[93] Lakoff & Johnson, p. 15-16.
[94] Examples of [Health is up; Sickness is down] could be: 'He *fell* ill' or 'He came *down* with the flu'.
[95] Lakoff & Johnson, p. 25-32.
[96] Lakoff & Johnson, p. 33-34.

As Lakoff & Johnson conclude, following their account of orientational metaphors: "*Most of our fundamental concepts are organized in terms of one or more spatialization metaphors.*"[97] As such, these metaphors are deeply rooted in language, cognition and our physical everyday experience, and serve as vehicles of understanding concepts, strictly by virtue of their own experiential basis[98] – i.e. by virtue of the perhaps most tangible and basic interpretation of the later Wittgenstein's notion of common life-form, as well as the cognitive and linguistic implications of the Embodiment Thesis.

Lakoff & Johnson do not argue that our conceptual systems are metaphorical in nature, due to our common physical basis alone. In fact, they claim that all experience is *cultural*, when they go on to say that we *"…experience the 'world' in such a way that our culture is already present in the very experience itself."*[99] Again, this ties in well with the later Wittgenstein's notion of common life-form, since even though our culture will make up an essential part of the contextual setting that surrounds any language-game, it is still detachable from a particular context or individual, in that it, pr. definition, spans across several individuals and settings.

The only concepts that are perhaps understood directly without metaphor are human, physical, body-oriented, spatial concepts like [up-down], [front-back], [near-far], [in-out], etc., which are 'directly emergent'[100], and since Lakoff & Johnson maintain that *all* experience is cultural, it will suffice to say that the before mentioned concepts are just more sharply delineated. In addition, because there are systematic correlates between these more sharply delineated concepts and our less sharply delineated *emotional* experiences, we form the basic orientational metaphors like [Happy is up], in order to conceptualize our emotions. Subsequently, structural metaphors, which to a higher degree are grounded in our non-physical experience, since both source and target domains are non-physical, non-orientational abstracts, employ concepts that are *metaphorically emergent*.

8.3 The Philosophy of Metaphor

Whether metaphorical concepts are grounded in our physical or cultural experience, or directly or metaphorically emergent, they all share the trait of being somehow structured by it, and in the same way, the conceptual structure of natural language, which is metaphorical in nature, is grounded in this experience. We understand the world through our *interaction* with it and as such, any concept of truth is relative to our cultural conceptual systems.[101]

---

[97] Lakoff & Johnson, p. 17.
[98] This also means that when Lakoff & Johnson say "happy is up", the "is" should be understood as denoting some set of experiences on which the metaphor is based, and in terms of which we understand it.
[99] Lakoff & Johnson, p. 57.
[100] 'Emergent' is understood as categories that emerge naturally from our experience.
[101] Lakoff & Johnson, p. 194-197.

This definition of truth stands in direct opposition to the traditional notion of truth in Western philosophy, and in the second half of MWLB, Lakoff & Johnson are mainly concerned with rejecting the dichotomy of the polarized debate between subjectivism and objectivism in Western philosophy and linguistics.[102] As their own alternative, Lakoff & Johnson come up with a third choice which they name 'experientialism'. The experientialist synthesis allows not only for a bringing together of the objectivist demand for absolute truth, with the subjectivist call for unconstrained imagination, but also explains how certain concepts that one could argue cannot be fully understood (such as emotions, aesthetics, spirituality, etc.) can be understood *partially* through 'imaginative reality' facilitated by the workings of metaphorical concepts.[103] Interestingly, these partially comprehensible concepts seem to be just the kind of language occurrences that the early Wittgenstein would like to see excluded from proper language use altogether.

In fact, Lakoff & Johnson's theory of experientialism ties in *so well* with the later Wittgenstein's rejection of objectivism, disembodied mind and logical atomism,[104] as well as his revised account of language and meaning, that the author of this thesis found it peculiar that it receives so little attention in MWLB. One of the few quotes that do accredit Wittgenstein the mention he deserves is perhaps the following comment made by Lakoff & Johnson on their account of experientialism:

"Our view also accords with some key elements of Wittgenstein's later philosophy: the family-resemblance account of categorization, the rejection of the picture theory of meaning, the rejection of a building block theory of meaning, and the emphasis on meaning as relative to context and to one's own conceptual system."[105]

To sum up, in MWLB, Lakoff & Johnson provide us with a richly documented, cognitive semantic theory of language that adheres almost perfectly to the later Wittgenstein's language philosophical outset, which according to the notion of approaching NLP design from the highest level of language analysis, was to guide the way we handle the issue of world knowledge in NLP.

Directly following from this, a question that naturally arises is to what degree (if any) a computational model will be able to take advantage of the metaphorical systematicity of language when put to the task of endowing an NLP system with world knowledge.[106] Before we can make any such evaluation, it would, however, be prudent first to acquire a rudimentary understanding of how an actual implementation of world knowledge might be construed. This is the intermediary goal of the following chapter.

[102] Lakoff & Johnson, chap. 24-30.
[103] Lakoff & Johnson, p. 193-203.
[104] Lakoff & Johnson's rejection of logical atomism is most apparent in their account of certain cognitive and linguistic primitives, such as causation, rendered as 'experiential gestalts' – a notion which is not pursued further in this thesis.
[105] Lakoff & Johnson, p. 182.
[106] This question is addressed directly in chap. 10, which explores the possibility of using artificial neural networks to serve this purpose, by featuring the implementation of a proof-of-concept.

9 World Knowledge through Ontology Engineering

This chapter introduces the basic concept of ontology engineering, as a possible answer to how we might implement world knowledge in NLP design. First, a basic explanation of the term is given. Then, one of two main types of ontologies used in NLP – domain specific ontologies – is introduced, in the form of a real world example of its practical application in the private, Danish company *Ankiro*. Hereafter, the second of the two main types of ontologies – top-level ontologies – is described, exemplified by the government funded *WordNet*. Finally, these two main types of ontologies are held up against each other and compared, in order to start off an initial discussion of how we might achieve an implementation of world knowledge, which takes an outset in the later Wittgenstein's language philosophical stance, as well as Lakoff & Johnson's cognitive semantic theory.

9.1 World Knowledge as Ontology

To recall, world knowledge is the highest level of language analysis and "…*includes general knowledge about the structure of the world…*" needed to successfully use language.[107] Now, if an NLP system is to be endowed with world knowledge, it must facilitate some kind of knowledge base, which can hold general information about the structure of the world. In other words, it would need to incorporate an 'ontology'.

As a philosophical term, an ontology is, in its broadest sense, the study of what exists in the world.[108] As a more tangible and practical term, ontologies are widely used in Knowledge Engineering, AI and Computer Science, and are applied to knowledge management, NLP, e-commerce, information retrieval, database design, education, and much more.[109] For each field and application, the term takes on new form and meaning. Throughout this thesis, the term will, however, be used solely as understood in the context of its various NLP applications.

Commonly, ontologies aimed at NLP consist of a network of subtype and subsumption relations, also called a taxonomy or hierarchy, where the relation is transitive so that each level introduces new features that in turn can be inherited by default, as illustrated by the simple example ontology in figure 9.1.[110]

[107] Allen, p. 10.
[108] A good example hereof, is the basic ontology of Tractatus described in chap. 5.1.
[109] Gómez-Pérez, Fernández & Chorcho, p. 1-3.
[110] Mitkov, p. 467.

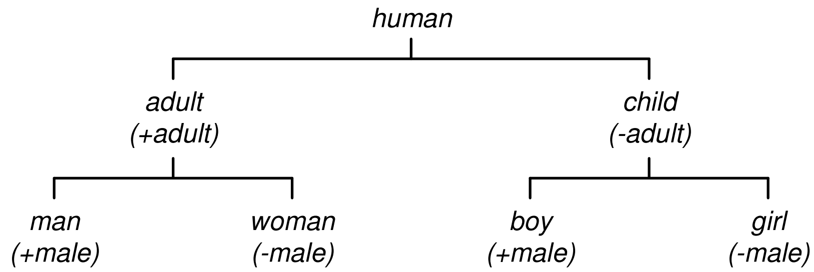


FIGURE 9.1

But even within the field of NLP, the definition of 'ontology' varies according to implementation framework, degree of formality and especially in terms of its conceptual spectrum.[111] Ontologies can support a knowledge base within an extremely specialized domain, or they can facilitate the relative ordering of the most general and abstract concepts. Most often, NLP systems are very limited in that they are restricted to a special domain – an example could be concepts relevant to booking an airline ticket. As Dr. Piek Vossen puts it in his article on ontologies and their use in NLP, in the most recent edition of 'The Oxford Handbook of Computational Linguistics': "*Most current NLP techniques hardly ever perform full natural language understanding.*"[112] As an illustrative example of this trend, the following subchapter takes a look at a real world example of Ontology Engineering.[113]

9.2 An Example of Applied Ontology Engineering: *Ankiro*

One of the few companies that have NLP and Ontology Engineering as their core competencies, is the Danish situated *Ankiro*, at which the author of this thesis was himself employed from 2003 to 2005. Ankiro is officially a research company with an outset in the further development of their advanced ontology technology, and it has always had strong ties to the research community.[114] The company works mainly with ontology-driven search tools – in fact, Ankiro was founded in 1998 to create the search technology employed by "Jubii" – a search engine which in its early years was the dominant alternative for browsing the Internet in Denmark.[115]

[111] Gómez-Pérez, Fernández & Chorcho, p. 9-25.
[112] Mitkov, p. 473.
[113] 'Ontology Engineering' refers to the set of activities that concern the ontology development process, the ontology life cycle, and the methodologies, tools and languages for building ontologies.
[114] http://www.ankiro.dk/content/dk/forskning
[115] http://www.ankiro.dk/content/dk/om_ankiro/historie

To enquire further into Ankiro's ontology architecture and discuss some of the basic problems and challenges associated with practical ontology engineering, the author of this thesis scheduled a meeting with Ankiro's Chief Technology Officer Steen Bøhm Andersen and Head of Language Department Louise Bie Larsen.[116] When asked about their stand on generality versus language domain specialization in ontology architecture, Mr. Andersen commented that even with small, specialized ontologies, the possible uses of language that can arise from the human side of the HCI presented an overwhelming scope of different problem areas that must be handled before the ontology driven search functionality can even begin to be exploited, and Mrs. Bie Larsen agreed.[117]

Such problem areas are bound to expand in conjunction with the widened scope of the applied domain of an ontology, which is expanded to allow for an increased degree of generality, but even when setting these linguistic difficulties aside, there is also the issue of how great a part of the totality of language that the ontology engineering process adopted in Ankiro will be able to handle. The problem is that every word and classification that is incorporated into the ontology, requires a certain degree of manual labour. This was explained by Mrs. Bie Larsen, as she demonstrated Ankiro's ontology training interface, detailing how an ontology is developed. She provides the following real-world example:

Ankiro has developed an ontology for a client organization that wants to allow its members to do an intelligent search in a large database of job ads. The way this was done, was to first analyse the content of the database and extract a collection of a few thousand high frequency words. Once these were in place, a linguist looked through each of them and entered them into what Mrs. Bie Larsen calls a 'Multi-inheritance Taxonomy'. Basically it is similar to the simple ontology seen in fig. 9.1, except for that instead of inherited properties always following the hierarchical structure of the taxonomy, they can be inherited from more than one super class.

She provides the following simple illustration of how this ontology is able to facilitate what can be termed an 'intelligent search'. Say, for instance, that a user has entered the names of two desirable geographical locations (**L1 AND L2**) and two job titles that are very similar and both of which hold interest for the user (**JT1 AND JT2**). A regular search would then consider the occurrence of all of these individual space-separated words for each job ad in the database, but because the ontology driven search engine is able to relate the classes of each of these words in the ontology, it can perform the search based on the following logical form for each job ad in the database:

(JT1 AND L1) OR (JT1 AND L2) OR (JT2 AND L1) OR (JT2 AND L2)

[116] The meeting took place at Ankiro's own offices on October 14, 2011. See also Appendix C.
[117] Some examples were the lexical ambiguity of words, and the syntactic ambiguity of certain constructions / phrases.

Of course, the ontology is also used for other things such as prioritizing and ordering collections of search results, but the point is that, in any case, the size and scope of the ontology is restricted by the consequent cost in man-hours and workload with respect to the linguists that have to identify and manually arrange the words correctly into the multi-inheritance taxonomy that makes up the ontology being developed.[118]

To sum up, even though Ankiro's advanced ontology architecture is a result of years of development and strong ties to the research community, it is still completely geared towards domain specific application. At the same time, the scope of their ontology implementations are essentially limited by the amount of manual labour that these implementations require, in that they are, after all, a private, profit oriented company with limited time and resources set aside for each client project. In this respect, they make for yet another example of the result-oriented approach to NLP design depicted in earlier chapters of this thesis.

9.3 Achieving Generality: Top-level Ontologies

Unlike regular, domain specific ontologies, top-level ontologies, also sometimes referred to as 'upper ontologies' or 'upper-level ontologies', describe general, reusable concepts which are common across different knowledge domains, and which existing domain specific ontologies should be linked to. As such, a top-level ontology should be able to encompass regular, domain specific ontologies, and also relate them to each other. [119] However, the differences in scope and implementation frameworks of existing, domain specific ontologies represent a problem in this respect, as do the different, competing approaches to top-level ontology engineering, which as a field is characterized by controversy, politics, competing approaches and academic rivalry.[120]

Any further, general definition of top-level ontologies is in this way problematic, since little consensus resides within the field as a whole. The author of this thesis would, however, like to offer his own simple understanding of top-level ontologies in the following paragraph. It is based on a study of the term as portrayed by Gómez-Pérez, Fernández-López and Corch in their book "Ontological Engineering" from 2010, which instead of providing a detailed definition of the term, describes the basic structures of several prominent top-level ontology architectures.[121]

[118] It should, however, be noted that Ankiro has developed an advanced interface that allows for an efficient work process when it comes to manually entering new words into the ontology, but its precise nature is not discussed in any further detail here, in order to protect certain trade secrets related to Ankiro's unique ontology development process.
[119] Gómez-Pérez, Fernández & Chorcho, p. 71.
[120] Gómez-Pérez, Fernández & Chorcho, p. 107-108.
[121] Gómez-Pérez, Fernández & Chorcho, p. 25-106.

Definition: While a domain specific ontology, as the name implies, is focussed on the part of language relevant to a specific domain of application, a top-level ontology maintains its potential for complete generality, by ordering concepts that would span across most domain specific ontologies. In addition, a top-level ontology will usually exceed the limited number of types of semantic relations that characterizes the standard taxonomy of domain specific ontologies – namely hypernymy and hyponymy ('subclass-of' and 'super class-of') – also called 'subsumption relations'.

With this definition in mind, a good example of a large, prominent, government funded top-level ontology, which has been widely used in NLP research, is 'WordNet' – a substantial database for English created and maintained at the Cognitive Science Laboratory of Princeton University.[122] It attempts to organize lexical information in terms of word meanings, by making use of a large number of 'synsets', each representing an underlying lexical concept.[123] These are then interlinked via different kinds of semantic relations. In addition to subsumption relations, it also features meronymy and holonymy ('part-of' and 'has-a') and synonymy and antonymy. The lexicon is divided into categories of nouns, verbs, adjectives and adverbs – each of which can have different kinds of the before mentioned semantic relations.

There are several other examples of top-level ontologies[124], but it will suffice to say that the field is still so new and fraught with controversy and a general lack of consensus, that any further rendition of the term that would seek to elaborate on the simple definition provided in the above, would have to do so by describing the individual implementation frameworks and general design approaches of even more of the different top-level ontologies that exist today – a task beyond the modest scope of this thesis. The aim was here simply to provide adequate descriptions of the two main types of ontologies, so that they, in the following subchapter, could be held up against each other and related to the language philosophical and cognitive semantic outset inferred by previous chapters' analyses.

9.4 Domain-specific vs. Top-level Ontologies

With the recently acquired overview of domain specific and top-level ontologies in place, we can now ask the question: What kind of ontology would be more feasible to use, when trying to endow an NLP system with world knowledge? If we return to the language philosophical outset of this thesis, one could argue that the later Wittgenstein's notion of language-games and the extreme degree of contextuality that characterizes individual instances of language use, could speak in favour of highly specialized world knowledge, realized by a very domain specific ontology.

[122] Gómez-Pérez, Fernández & Chorcho, p. 79-80.
[123] WordNet's latest version is 3.1, as of June 2011. As of 2006, the database contains 155,287 words organized in 117,659 synsets for a total of 206,941 word-sense pairs, (*Source: http://wordnet.princeton.edu*).
[124] Gómez-Pérez, Fernández & Chorcho, chap. 2.2.

On the other hand, building an ontology for a single, unique instance of language use, seems both superfluous and rather pointless, since its potential domain of application would be equally narrow and specialized. Instead, world knowledge should perhaps be perceived as pertinent to the common life-form which allows us to overcome this contextuality in our repeated use of natural language. One point is to distinguish between what can be called 'situational knowledge' and 'general knowledge'. While the situational knowledge, which is strongly tied to a single, unique language instance, does indeed help dictate the rules and form of its underlying language-games, the shared conceptual framework around such a language instance would take the form of general knowledge shared across several, individual language users by way of their common life-form, in accordance with our definition of world knowledge.[125]

When using ontology engineering to endow an NLP system with world knowledge, one should therefore try to structure the ontology around the patterns of language related to our shared patterns of cognition, so far best exemplified by the metaphorical systematicity of both natural language and human cognition found in Lakoff & Johnson's MWLB.[126] As the wording implies, this general, as opposed to situational, knowledge is best handled by a top-level ontology. This, however, does not rule out the use of domain specific ontologies. As we recall, the role of a top-level ontology, besides handling general concepts and additional types of semantic relations, can also be to encompass and integrate a number of underlying domain specific ontologies.

Regarding the issue of domain specific ontologies vs. top-level ontologies, the approach suggested here is thus to employ a combination of the two. One could, as an example, then imagine a top-level ontology built around the more basic metaphorical systematicity of language, such as orientational metaphors, which could then rely on domain specific ontologies to contextualize this general knowledge further, by handling subsystems of more complex metaphors related to pre-defined contexts.

However, before delving any further into the intricacies of how to approach the task of devising different candidate architectures for an ontology based on the language philosophical outset of the later Wittgenstein and the cognitive semantic theory presented by Lakoff & Johnson in MWLB, we will first return to the question posed in the previous main chapter – namely: "…to what degree (if any) a computational model will be able to take advantage of the metaphorical systematicity of language when put to the task of endowing an NLP system with world knowledge."[127] This is the focus of the following chapter.

[125] See chap. 3.5 + 11.1.
[126] See chap. 7 + 8.
[127] See end of chap. 8.

10 Proof-of-Concept for a Computational Model of Metaphor

In this chapter, a proof-of-concept (POC) is developed to show that the metaphorical systematicity of language can be handled by a computational model – in this case an artificial neural network (ANN), the potential application of which is also meant to be demonstrated. First, the overall purpose and limitations of the POC are explicated. Then the basics of ANNs are explained, and the implemented ANN is tested on a simple test case, before being applied to the problem represented by the POC. Besides POC design and implementation, this also involves coming up with a working solution to the problem of acquiring a sufficient amount of training data for the ANN. Finally the POC is evaluated and its different constraints and merits are analysed and discussed further.

10.1 Purpose and Limitations

When implementing a POC with the purpose of evaluating whether a computational model will somehow be able to take advantage of the metaphorical systematicity of language, it is important first to distinguish it from what is called a ‘prototype’ in computer science and software development. A prototype most often represents an early version of the final system, which simply has some features missing or which is only partly implemented. Some prototypes only fully implement a small part of the final system, while others might be so-called ‘beta-versions’, which just need a bit of testing and fixing before reaching the final stage of completion.[128] The amount of implemented functionality is not the issue. The point is that while a prototype represents some stage of development of the final product, a POC does not.

Instead, a POC does exactly what the name implies – it simply proves a concept, by acting as a ‘demonstration in principle’, the purpose of which is to verify that some concept or theory has the potential of being used.[129] Accordingly, the POC developed in this thesis is in no way meant to represent an early version of a specific new type of top-level ontology – it should simply demonstrate that it is somehow possible to make a computational model which takes advantage of the metaphorical systematicity of language. This concept is, after all, essential to the evaluation of whether or not the cognitive semantic and language philosophical premise established so far, represents a serious, plausible approach to the issue of endowing NLP systems with world knowledge through ontology engineering.

[128] Kruchten, p. 190-192.
[129] Kruchten, p. 129+278.

10.2 Computational Model: Artificial Neural Networks

Since the metaphorical systematicity of language, which the computational model is supposed to handle, is based on certain patterns in cognition that are evident as patterns in language[130], the computational model should be able to perform some degree of pattern recognition. The idea will be to take linguistic input and interpret it in accordance with a system of metaphors – in other words, it will need to use its pattern recognition abilities to perform classification of linguistic parameters into metaphor related segments.

Because language is imprecise from a computational perspective, it will need to do this by functioning as a 'supervised regression algorithm', that will receive correct, but often noisy examples of input/output data sets, which will be used to predict real-valued labels.[131] A computational model that handles noisy data well, and which uses pattern recognition to predict real-valued, categorical labels, is an 'artificial neural network' (ANN)[132], and for this reason it has been chosen as the computational model for the POC featured in this thesis. The following subchapter elaborates on the ANN implementation specifics, but let us first take a quick look at how a basic ANN works.

The ANN is inspired by the way the human brain works, and how it contains a highly interconnected group of neurons which, based on their input from other neurons, may or may not 'fire', allowing for electrical impulses to propagate through connections between them. In an ANN, each neuron takes a number of inputs along weighted connections as seen in fig. 10.1, which shows a basic neuron with two inputs x_1 and x_2.

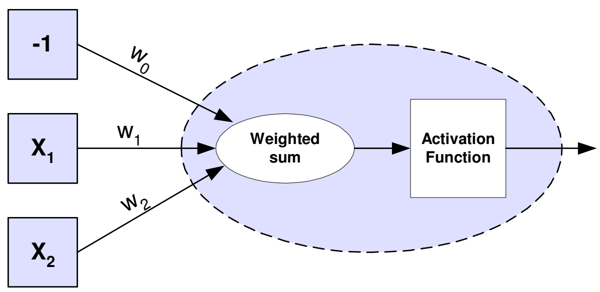


FIGURE 10.1

[130] See chap. 7 + 8.
[131] Heaton, p. 35-37.
[132] Ross, chap. 5.

The weighted sum of the values of the inputs, along with the bias weight with a fixed input of -1, is then used by the activation function to determine whether or not that neuron fires.[133] Using a bias in this way has the effect of giving each neuron a trainable threshold, equal to the value of the weight from the bias unit, which means that the neuron fires when the weighted sum of the 'real inputs' (in this case x_1 and x_2) exceeds the bias weight.[134] A commonly used activation function, which will also be used here, is the sigmoid function, which transforms any single input value between $-\infty$ and $+\infty$, into a reasonable value between 0 and 1. [135]

In an ANN, neurons are connected in a hierarchical structure of layers of neurons, where neurons within a layer only fire to neurons in the next layer, and never to neurons in that same layer, as seen in fig. 10.2. The example ANN seen here is also fully connected, meaning that every neuron can fire to every neuron of the next layer.

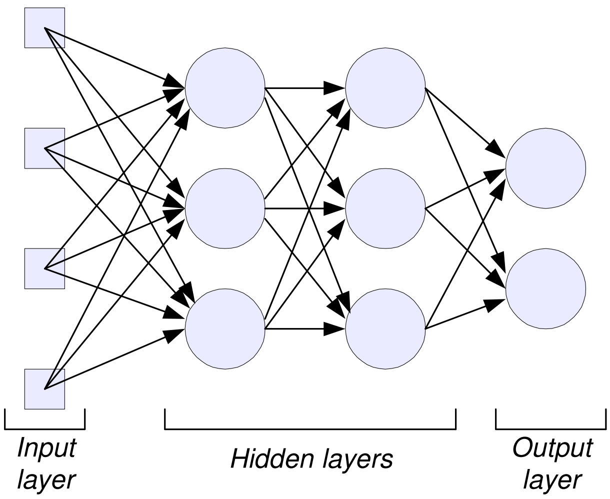


FIGURE 10.2

An ANN works on input data which is mapped to the input layer, and depending on the applied methodology, it may or may not have an expected output from which to create a model. The ANN implemented in this thesis will make use of supervised learning, in that it gradually creates a model that matches patterns of input mapped to the input layer, to an expected output which is compared to the values of the output layer.[136]

[133] Negnevitsky, p. 165-185.
[134] Russell & Norvig, p. 737-738.
[135] Heaton, p. 169.
[136] Heaton, p. 162-166.

For each input/output-pair provided, the ANN first propagates the input through its layers in a feed-forward manner to produce an output based on the current weights of the ANN, which upon initialization are assigned randomly. When the input does not match the expected output, the error is determined and used to back-propagate through the ANN, adjusting the weights, so that the error in the output for a similar pattern will be reduced. These steps are usually repeated until the average error is small enough for the ANN to yield satisfactory results. This is how an ANN can be used to generalize upon data through pattern recognition and classification.[137]

The implementation of the ANN itself is done in JAVA[138] and involves two classes: `Neuron.java` and `NeuralNetwork.java`.[139] The first is a simple class which holds functionality that enables it to calculate the weighted sum of its inputs, and set its own value to the sigmoid of that sum. The core functionality of the ANN is found in the `NeuralNetwork.java` class, which organizes lists of neurons in layers and holds functionality to set the input of the input layer, activate the fully connected feed-forward firing to produce an output, get the ensuing output from the output layer and apply the backpropagation algorithm, which adjusts the weights of the ANN by comparing the actual output to the expected output.[140]

Once the ANN was fully implemented, a simple test class was made, to ensure that the ANN implementation functioned properly, before attempting to apply it in the more advanced context of metaphorical systematicity. The Java class `NN_XOR_funcTest.java`,[141] tests the ANN on an XOR gate, which is not linearly separable, and as such will require a hidden layer and thereby also require the more intricate part of the ANN implementation to function properly, in order for the ANN to converge.[142] Fig. 10.3-A shows the input/output data sheet of the XOR function, while figure fig. 10.3-B shows the gradual convergence of the ANN in relation to the number of training iterations.[143]

[137] Ross, p. 40-46.
[138] The decision to use Java was based upon the authors own programming experience and the fact that the applied literature – namely Jeff T. Heaton's "Introduction to Neural Networks with Java" – also used Java for code examples.
[139] See 'JavaDoc' and 'Source Code' on the enclosed CD for more information.
[140] For further details, please see 'JavaDoc' and in-code comments in 'Source Code' on the enclosed CD
[141] See 'JavaDoc' and 'Source Code' on the enclosed CD for more information.
[142] Single-unit perceptrons are only capable of learning *linearly separable patterns*; in 1969 in a famous monograph entitled 'Perceptrons' Marvin Minsky and Seymour Papert showed that it was impossible for a single-layer perceptron network to learn an XOR function. This is therefore a good test for the multilayered neural network implemented here.
[143] Each pair of 0/1 combinations, together with the result of XOR on that pair, can be seen as one data set instance, meaning that the ANN actually is trained on the same 4 data set instances.

A

Input		Output
A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0

B

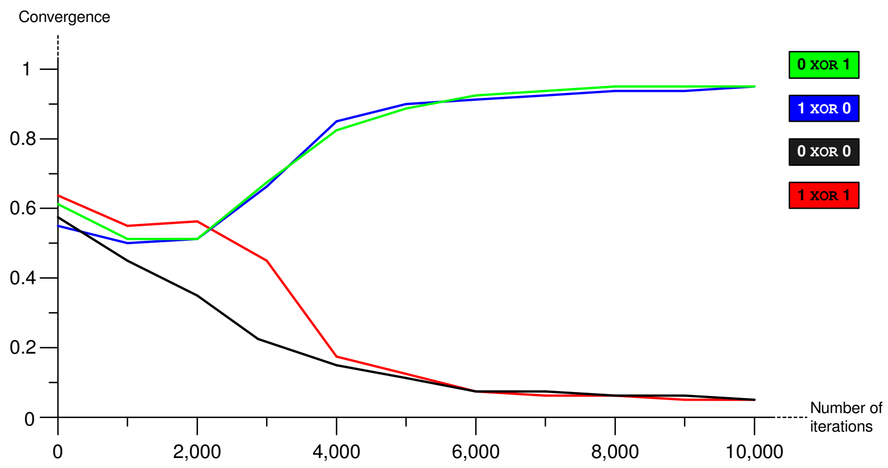


FIGURE 10.3

Besides demonstrating that the ANN is working properly, this small example also leads us to another critical issue; namely that of the amount of data sets needed to sufficiently train an ANN. As seen from fig. 10.3-B, even an extremely simple function like XOR needs more than 5,000 training iterations, or data sets, to reach a satisfactory degree of convergence. It is not hard to imagine that an ANN based on the metaphorical systematicity of language is likely to require even more training data sets. This is an issue, which the author of this thesis is aware of and will return to shortly.[144] For now, the overall design features of the POC first need to be laid out.

10.3 Design and Implementation

The POC will map linguistic input to one of three orientational metaphors: [*Happy is up; Sad is down*], [*More is up; Less is down*] and [*Conscious is up; Unconscious is down*]. Their physical basis is very straightforward in the descriptions by Lakoff & Johnson. The first is based on the fact that "*…a drooping posture typically goes along with sadness and depression, erect posture with a positive emotional state.*"[145] The second is based on how "*…if you add more of a substance or of physical objects to a container or pile, the level goes up.*"[146] The third is based on how "*…humans and most other mammals sleep lying down and stand up when they awaken.*"[147]

144 See chap. 10.4.
145 Lakoff & Johnson, p. 15.
146 Lakoff & Johnson, p. 16.
147 Lakoff & Johnson, p. 15.

All three metaphors are as such based on the most basic interpretation of the later Wittgenstein's notion of common life-form: our mutual physical experience – in this case pertaining to gravity and verticality. The output neurons of the ANN will each have a real-valued label between 0 and 1, which represents one of the three chosen orientational metaphors. The overall, functional purpose of the ANN, once it has been fully trained, will in this way be to tell us which of these three metaphors, if any, the linguistic input most strongly points to.

The input layer of the ANN will have 14 neurons, each representing a key-word, which in combination with one of a selected range of 'triggers' associated with physical verticality, is likely to form one of the three chosen metaphors. The triggers are: { up, down, under, high, low, fall, fell, rise, rose, drop, dropped }. The 14 key-words which each represents one of the three metaphors, when combined with the correct triggers are: { spirit, spirits, feel, felt, cheer, cheered, go, went, income, prices, wake, woke, asleep, hypnosis }.[148] The following overview illustrates the combinations that evaluate to one of the three metaphors (triggers are underlined):

HAPPY IS UP; SAD IS DOWN

[spirit, spirits] [<u>fall, fell</u>]
[spirit, spirits] [<u>rise, rose</u>]
[<u>high</u>] [spirit, spirits]
[<u>low</u>] [spirit, spirits]
[feel, felt] [<u>down</u>]
[feel, felt] [<u>up</u>]
[feel, felt] [<u>high</u>]
[feel, felt] [<u>low</u>]
[cheer, cheered] [<u>up</u>]

MORE IS UP; LESS IS DOWN

[go, went] [<u>up</u>]
[go, went] [<u>down</u>]
[income] [<u>rise, rose</u>]
[income] [<u>drop, dropped</u>]
[prices] [<u>rise, rose</u>]
[prices] [<u>drop, dropped</u>]
[prices] [<u>fall, fell</u>]

CONSCIOUS IS UP; UNCONSCIOUS IS DOWN

[wake, woke] [<u>up</u>]
[<u>fall, fell</u>] [asleep]
[<u>under</u>] [hypnosis]

[148] The observant reader will notice that only the simple present and simple past forms of the triggers and key-words that are verbs, have been included. This has been done to simplify the model and eliminate the issue of having to handle different auxiliary verbs and participles to do with more complex types of verb form and tense.

The way triggers and key-words are combined is simply by syntactical association. Specifically, if a trigger has a key-word immediately preceding or following it, it will be mapped to the metaphor associated with it. For example, the trigger ‘spirits’ will map to a [*Happy is up; Sad is down*] metaphor if it is preceded or followed by any of the key-words { fall, fell, rise, rose, high, low }. Textual examples could be ‘He was in high spirits.’, or ‘Spirits rose when rations arrived.’ The chosen configurations are based on the examples provided by Lakoff & Johnson in their treatment of each of the three metaphors in MWLB. In addition, since the ANN does not distinguish between its binary input values, it is crucial to keep a fixed order of these values to preserve forming patterns. The same goes for the output. To that end, fig. 10.4 both illustrates and defines the fixed order of both the input and output of the ANN.

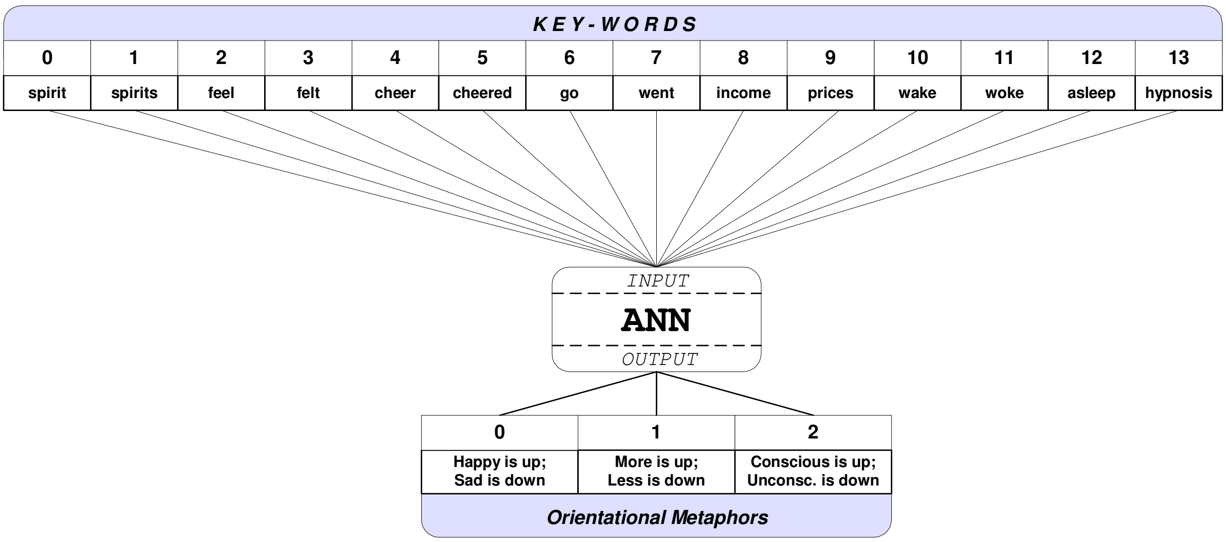


FIGURE 10.4

This leaves us with the problem of deciding on the internal structure of the ANN that will map triggers and key-words to metaphors. In his book on neural network design ‘Introduction to Neural Networks with Java’[149], AI researcher Jeff Heaton identifies two main design tasks, when implementing an ANN. One is deciding on the number of layers between the input- and output layer, also called ‘hidden layers’. The other is deciding on the number of neurons in each hidden layer. [150] The function represented by the POC ANN is potentially very complex, since it involves natural language. Nevertheless, two hidden layers should suffice in that such an ANN should be able to represent functions with any kind of shape, and there is currently no theoretical reason to use more than two layers.[151] This implementation will therefore feature two hidden layers.[152]

[149] Heaton, p. 128-131.
[150] See also fig. 10.2.
[151] Heaton, p. 128.

Deciding on the number of neurons in each of those layers is also a theoretical grey area for which Heaton provides a few rules of thumb, while encouraging an explorative approach. These rules of thumb suggest that the number of neurons in the hidden layers should be somewhere between the size of the input- and the output layers'– perhaps around 2/3 of the input layer's size plus the output layer, and finally that they should not be more than twice the size of the input layer.[153]

This should be seen in relation to the fact that too many neurons in the hidden layers will result in 'overfitting', where the information processing capacity of the ANN is inflated so that there is not enough training data to train all these neurons – a pressing matter for this implementation. At the same time, too few neurons in the hidden layers will decrease the scope and increase the time it takes to train the network, making it very difficult to train efficiently.[154] On that basis, the final design of this implementation will feature 10 hidden neurons in each of its two hidden layers.

10.4 Training Data

As mentioned in subchapter 10.2, it would seem that when a small ANN needs several thousand training iterations to learn a simple function like XOR, it serves as a strong indication that an ANN based on the systematicity of natural language is likely to require even more training iterations. Since we do not know the precise nature of the underlying function of language that we want the ANN to learn, such an ANN should ideally receive a new language instance for every training iteration, so that it can learn by example without having to reuse any data.

This is, however, far beyond of the scope of this thesis, due to the time and manpower that would be required to collect enough data to accumulate such a large corpus. The approach taken in this thesis, has instead been to use the limited amount of empirical data provided by Lakoff & Johnson, which is illustrated by the overview in chapter 10.3, for designing a 'simulated corpus' that will be able to train the ANN, and in this way avoid the pitfall of overfitting.

In systems dealing with natural language, the use of simulated corpora is becoming more and more popular, since the annotation of sufficient data for training NLP systems that employ artificial intelligence programming approaches is a problem which is becoming increasingly prevalent. An example of this trend is seen in the research article 'Comparing Real-Real, Simulated-Simulated, and Simulated-Real Spoken Dialogue Corpora' by Hua Ai and Diane Litman from Pittsburgh University, which describes how user simulation can be used to generate large, simulated corpora aimed at reinforcement learning.[155]

---

[152] It should be noted that there is actually no theory on the subject of how many hidden layers an ANN should have.
[153] Heaton, p. 129.
[154] Heaton, p. 129.
[155] Ai and Litman, p. 1-4.

Due to the limitations of the POC being developed here, instead of training a model to perform complex simulation, the more simple approach of using dynamic programming to develop a data generation function is taken. This will then provide the data sets needed to train the ANN, which will simply call the data generation function at every training iteration. In this way, the data provided by the data generation function will consist of an array of 14 input- and 3 output values between 0 and 1. Even though this also means that the best performance achievable by the ANN will be an approximation of that same function, the approach still has merit, in that it will show how an ANN can be used to perform pattern recognition and classification on natural language data with respect to metaphor, given a strictly empirically based corpus, accumulated through a more advanced type of metaphor representation and identification, than just simple, syntactic association.

10.5 Performance and Experiments

With the data generation function in place, it is finally possible to train the ANN using the data sets it provides. The Java class `NN_Trainer.java`[156], which creates the ANN with the specified number of layers and neurons, and loads the `DataManager.java`[157] class, which among other things, includes the data generation function. Training is then performed through the sequence illustrated by fig. 10.5, with the cycle continuing for a predefined number of training iterations. Once the ANN has been trained, its updated weights are written to the file `network.ann`[158] as real numbers, so that it can be saved outside of the Java runtime environment, and used again for further testing.

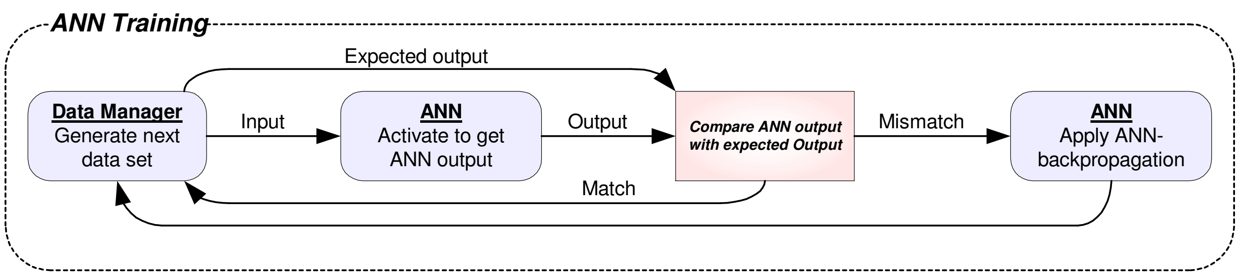


FIGURE 10.5

Once training has completed, it is possible to feed the ANN new data and evaluate its output, but before doing so, a test is performed to evaluate whether or not the trained ANN converges at all, and if so, at what rate in relation to the necessary number of training iterations. To perform this test, additional code was added to the `NN_Trainer.java` class, which stops the training cycle at every 1,000'th training iteration and calculates the prediction rate of the ANN in its current state.

---

[156] See 'JavaDoc' and 'Source Code' on the enclosed CD.
[157] See 'JavaDoc' and 'Source Code' on the enclosed CD.
[158] See 'JavaDoc' and 'Source Code' on the enclosed CD.

This is done by activating the ANN with new input 1,000 times *without* doing any backpropagation on mismatches, and then dividing the number of successful predictions with the total number of tries. The raw output, which can be seen in Appendix A, is illustrated in fig. 10.6, which shows the ANN's overall degree of convergence over 10,000 training iterations for 4 different test runs.

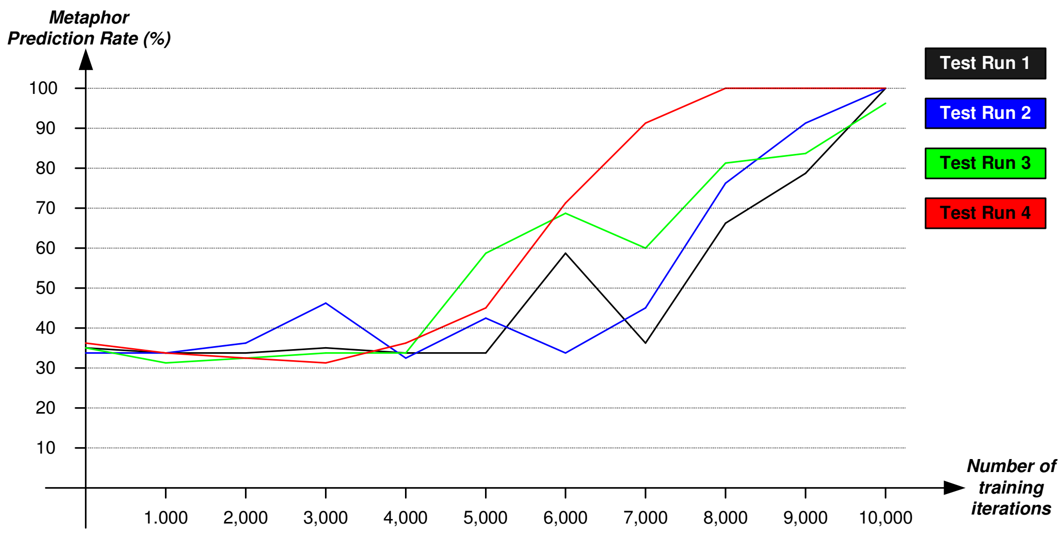


FIGURE 10.6

As expected, the initial metaphor prediction rate before training is very close to 33.3% for each test run, simply because the weights of the ANN are initialized to random values, and as a result there is an even chance of predicting one of the three possible metaphors to begin with. However, after around 4-5,000 training iterations, prediction rates start to increase and after 10,000 training iterations, three out of four prediction rates have reached 100%.[159] We can thus verify that the ANN is able to converge on linguistic data in the form of a simulated corpus, and thereby perform pattern recognition and classification through metaphor prediction. [160]

Such a high degree of convergence is of course uncharacteristic of computational models that perform pattern recognition and classification on data which takes an outset in something as complex, noisy and diverse as natural language data, but again this is due to the fact that the ANN implemented here, does little more than approximate the pattern of the data generation function, which in itself is relatively simple.[161]

---

[159] In addition, after 15.000+ training iterations, every test run becomes stable at 100% degree of convergence.
[160] The reader is encouraged to run the `NN_Trainer.java` class and see the full output for him- or herself. See the .txt-file 'Source Code\Instructions.txt' on the enclosed CD for further instructions on how to load and run the code.
[161] In fact, superficial testing with no hidden layers showed it to be linearly separable.

A test-class `NN_Tester.java`[162] was also developed to test the trained ANN on some authentic, linguistic data. The data chosen here is a collection of Hans Christian Andersen's fairy tales, containing around 230,000 words.[163] 89 language instances were found to be one of the three chosen orientational metaphors, of which 30 were wrong, for a precision rate of about 66%.

Sample Outputs

```
[More is up; Less is down]: ...it all went up in flames The ...
[Con. is up; Unc. is down]: ...the prince woke up and was not ...
[Happy is up; Sad is down]: ...kept his spirits up and smeared his ...
[More is up; Less is down]: ...if I went down to the bottom ...
[Con. is up; Unc. is down]: ...But she could fall asleep and sleep ...
```

FIGURE 10.7

Fig. 10.7 shows a few sample outputs of orientational metaphors identified in the H.C. Andersen text by the trained ANN. Though it is interesting to see some real examples, these results of course tell us little more than how well the intuition behind the design of the simulated corpus and its simplified metaphor representation is able to handle the lexical ambiguity of valid trigger/key-word combinations, by such syntactic association.[164]

10.6 Evaluation and Discussion

In demonstrating that the ANN is able to converge on the XOR gate, it was established that the ANN Java implementation developed here is fully functional. With a working ANN in place, the strategy of developing a simulated corpus was applied to the problem of acquiring a sufficient amount of ANN training data. To this end, a data generation function which randomly mapped 1 of 14 key-words, which could be matched with a range of orientational triggers, to 1 of 3 orientational metaphors, was developed. Though it is based directly on the linguistic data provided by Lakoff & Johnson in their chapter on orientational metaphors in MWLB[165], it should be obvious that using this kind of syntactic association as representation for metaphorically based patterns in natural language, not only lacks a strong theoretical foundation, but also makes for a gross simplification of the underlying systematicity of language that those patterns should represent.

---

[162] See 'JavaDoc' and 'Source Code' on the enclosed CD.
[163] See also 'Full Texts\HC Andersen - Fairy Tales.txt' on the enclosed CD.
[164] The complete output, which `NN_Tester.java` prints to console, when run, can be seen in Appendix: 2 Trained Neural Network Test Output.
[165] Lakoff & Johnson, p. 14-21.

However, as also stated in subchapter 10.1 on the overall purpose and limitations of the POC, the ANN developed here is not a prototype, but a ‘proof of *concept*’, the aim of which is simply to demonstrate that it is somehow possible to make a computational model which takes advantage of the metaphorical systematicity of language. Testing showed this modest goal to have been achieved. However, as mentioned earlier, one could still argue that the trained ANN, when applied to linguistic data, does little more than approximate the pattern of the data generation function, which itself is overly simplistic, as is also indicated by the unusually high degree of convergence displayed by the trained ANN.

The point is that even though the complex issue of metaphor representation has been greatly simplified in order for the POC to show concrete results, by going from the highest level of language analysis, i.e. world knowledge in the form of metaphorical systematicity, all the way down to the level of words and characters, skipping every language analysis level in-between, this does not take away from the seeming potential of using ANNs, once a better representation of metaphor is found. It is true that the POC goes against the very approach to NLP advocated in this thesis by grocely simplifying the highest level of language analysis in order to make it easier to jump straight down to the lowest language level, thus enabling the development of a functional POC implementation, but again; the goal has been to prove the concept of using a computational model to handle the metaphorical systematicity of language, and to show the potential of ANNs in this respect – not to develop an early prototype.

This potential of using ANNs illustrated by the POC is, however, an area prone to misconceptions concerning its different modes of application. To be specific, ANNs are often thought of as a way of directly simulating how the human brain works. Literature on ANNs even tends to leave out the ‘artificial’ and simply call them ‘neural networks’.[166] In mathematical terms, ANNs are only simplified abstractions of the complexities of the human brain, in the way they acquire knowledge through learning, and in the way this knowledge is stored within inter-neuron connection strengths. If an ANN can be said to serve as a simulation of the human brain, this is therefore on an extremely abstract and simplified level. This characteristic is important, since George Lakoff in his work following MWLB, grounds his cognitive semantic theory of metaphor in neurobiology.[167]

[166] Heaton, p. 31-32.
[167] Lakoff & Johnson - 2003 Afterword, p. 254-264.

The example of his neural theory of metaphor, which he often uses when giving talks on the subject, takes an outset in the orientational metaphor [More is up; Less is down] also used in the POC in this thesis. He explains that through a subconscious ordering of experience occurring in the brain, the metaphor becomes a structure of neural connectivity in the brain itself. For example, when an infant sees a cup being filled with milk again and again, the experience of quantity is, through co-occurrence, associated with the experience of verticality, in that the level of milk rises as more milk is added. In other words, neural connections between the two parts of the brain which have to do with verticality and quantity form as a direct result of these types of experiences.[168]

When viewing an ANN as a simulation of the human brain, this quickly leads to intuitions on using it to structure metaphors in a similar way. The problem is that an ANN is *not* a human brain – it is in crude terms a function approximator, which can be trained if fed correct sets of input- and output data. If we take the orientational metaphor [More is up; Less is down], which was just shown to be an actual structure in the human brain, how would an ANN hold this information? Having one input for verticality and one for quantity, and an output which describes their degree of connectedness as a real number between 0 and 1, tells us very little about how this connection relates to natural language. Furthermore, simulating the human brain by building a type of ANN which can represent *all* the necessary connections, or trying to combine all the ANNs required, seems to be a daunting task of overwhelming complexity. This is not to say that such an approach to achieving world knowledge in NLP is completely infeasible; it is just not the use of ANNs suggested in this thesis.

Instead, the use of ANNs advocated here is to work on linguistic data, drawing upon our understanding of relevant cognitive patterns, in order to perform metaphorical classification which can hopefully be used as abstracted world knowledge, to be utilized within the field of NLP.[169] The reasoning behind this approach, is that since patterns in cognition structure and define the world knowledge stored in our brains as neural connections, as well as the patterns in language through which we are able to analyse these underlying patterns of cognition, we should be able to use this knowledge to find a way to represent abstractions of the systematicity of language that would serve as input to an ANN, which could output classifications endowed with meaning on that basis. In this way, ANNs would be able to approximate the underlying systematicity of language (cognitive patterns) on the basis of linguistic data. Fig. 10.8 provides an illustration of this idea.

[168] To view the relevant part of the lecture, please see the 'Lecture with Lakoff'-folder on the enclosed CD.
[169] For more on these patterns, see chap. 7 + 8.

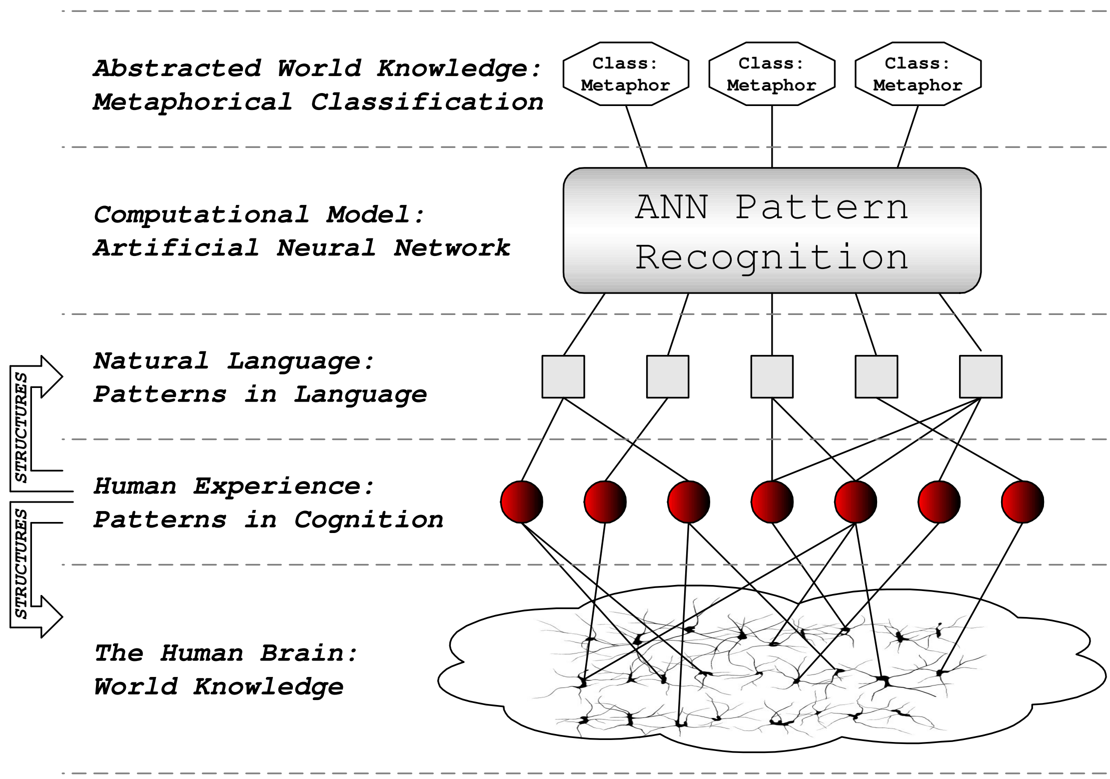


FIGURE 10.8

The way in which the ANN in fig. 10.8 would acquire a sufficient amount of training data would be similar to the way this problem was attacked in Ankiro.[170] Linguists, who in this case would need to have a thorough understanding of cognitive semantics and the cognitive theory of metaphor, would manually map ANN input to output, creating valid data sets in accordance with their own understanding. These manually validated data sets could then be compiled into a corpus which the ANN would use for training, allowing for an abstraction of the world knowledge contained in the minds of those linguists. This could then be used within the field of NLP where the trained ANN could generalize upon this knowledge, performing pattern recognition on noisy, linguistic data.

However, there are still two areas where additional work is needed before such a model could even begin to work. For one, the gross simplification of metaphor representation which in the POC of this thesis is reduced to syntactic association, would need to be developed further to better reflect the complexities of this cognitive semantic theory, while still ensuring that it could somehow be represented as an ANN input-format of binary or real numbers between 0 and 1. Secondly, further research into the scope of the metaphorical systematicity of language would need to be done, since MWLB only provides cognitive semantic examples and trends of metaphor in natural language use. They do not present a complete linguistic body of evidence in favour of a system that can account for the totality of natural language – an issue which will be treated further in the following chapter.

[170] See chap. 9.2.

11 Building upon Metaphor: Multi-level Contextuality

Besides demonstrating that it is possible to make a computational model which takes advantage of the underlying metaphorical systematicity of language, as well as elucidating the potential of ANNs when it comes to attaining abstracted world knowledge in the form of metaphorical classification through pattern recognition on noisy linguistic data, the POC also raises questions about the theoretical application and scope of metaphorical classification itself. In this chapter, Lakoff & Johnson's cognitive semantic theory of metaphor is investigated further by looking into some of the most pressing issues and problems surrounding its practical application.

To start with, the problem of NLU is re-evaluated with an outset in the cognitive semantic and language philosophical premise established in this thesis. Hereafter, further attention is paid to Lakoff & Johnson's scientific method and their theories' overall linguistic scope. Then the issue of contextual segmentation is treated in detail, and followed by a short overview of Lakoff's own answer to this problem – his theory of 'Frames'. Finally, an investigation into the compositional nature of metaphors and their possible representations prompts the introduction of Fauconnier's highly compatible Mental Space and Blending Theory.

11.1 The Problem of Natural Language Understanding – Revisited

When attempting to establish a computational model based on the systematicity of any aspect of natural language, the problem of NLU and its inherent criticism, discussed in chapter 3.5, is likely to resurface. In the academic world, it is perceived as a serious problem pertaining to any attempt to achieve a complete model of language[171], so let us briefly revisit the issue.

While the later Wittgenstein's notion of common life-form, as realised by Lakoff & Johnson's cognitive semantic theory, does imply a promising, systematic take on both language and cognition, one could argue that it is still subject to the problem of NLU, which states that because world knowledge is expressed in natural language – a system for describing human perceptions – and because human perceptions are intrinsically imprecise, natural language is itself inherently imprecise.[172] However, in light of previous, language philosophical analyses of truth and meaning, it should now be possible to see how the problem of NLU is actually biased towards an objectivist, disembodied view.

[171] Zadeh, 164-166.
[172] See chap. 3.5.

To start with, one should be wary of any kind of criticism based on the fallibility of the human perceptual system, since the logical conclusion is sure to be, that there is very little to be said about anything at all. To be specific, its logical conclusion amounts to the methodological scepticism expressed by the French philosopher René Descartes, who argued that since the senses can betray us, they cannot be trusted fully, and therefore the only thing we *can* be certain of, is that there is a mind which thinks/doubts – best known under the famous wording: 'Cogito ergo sum' ('I think, therefore I am'). Descartes does not give us anything, except this one-dimensional, ethereal point of 'doubting' – not even a body.[173] His mind-body dichotomy and dualism is the first, and perhaps purest, example of the disembodiment theory, and he plays a significant role in its pervasiveness in Western philosophy and the rationalist tradition.[174]

The bias also shows itself in the way the purpose of natural language seems to be to 'describe human perceptions' of the world in a precise way. Obviously, this kind of 'precision' entails hypothesising about a comparison of the description of these perceptions (language) and the things they describe or represent (objective reality) – much in accordance with Wittgenstein's picture theory in Tractatus – a theory which he later came to reject. If we, instead of taking the meaning of a natural language instance to be its truth conditions, choose to view its meaning as its *use*, as the later Wittgenstein did, we will see that any kind of 'precision' will be based on whether or not the language instance was able to fulfil its intended purpose within the conceptual framework that our common life-form facilitates.

Based on this analysis of the problem of NLU, and the language philosophical and cognitive semantic premise[175] established in earlier chapters of this thesis, it would seem that using natural language in a precise way is still theoretically possible, irrespective of the fallibility of the human perceptual system. With this conclusion in place, we now turn our attention to the issue of whether or not metaphorical systematicity is able to account for the totality of natural language use.

11.2 Natural Language: More than Metaphor

As hinted at in the final part of the evaluation and discussion of the POC developed in this thesis, the metaphorical systematicity of natural language as a complete solution model to the problem domain of designing a top-level ontology for use with NLP systems, seems to partly hinge on the ability of this cognitive semantic theory to encompass the totality of natural language. Discerning its true potential in this respect is, however, an extremely complex issue.

[173] Though Descartes did eventually re-establish a richer ontology, the ontological 'proof' which he utilizes is considered metaphysical in nature and is often criticised, since it simply rebuilds the physical world by introducing an a priori proof for the existence of God, which thereby enables the creation of the world and all things in it.
[174] http://plato.stanford.edu/entries/descartes
[175] This includes a rejection of the notion of absolute truth.

To be sure, Lakoff & Johnson provide a rich body of linguistic evidence in the form of many convincing examples which serve as strong indications in favour of their view that natural language is intrinsically metaphorical. The problem is that, as mentioned earlier, this is far from a complete system, able to account for the totality of natural language. This means that to start with, more work needs to be done before the cognitive semantic theory of metaphor can be said to encompass natural language, leaving the issue of whether or not it would be possible for it to do so at all, even with sufficient time and resources, as an open question.

The scientific, empirical method employed by Lakoff & Johnson in establishing their cognitive semantic account of language as metaphor only complicates the issue further. The reason is that it involves identifying metaphorical undercurrents of natural language use, and then digging up more of the linguistic occurrences that best exemplify them. One could therefore argue that not only is the identification of metaphors and types of metaphors prone to subjective intuitions, but also that the type of generalization involved in metaphorical type definitions, could be based solely on the very select few examples that were meant to serve as indications of a different or much more general feature of language. In their new book: "Quantitative methods in cognitive semantics: corpus-driven approaches" AI and Linguistics researchers Glynn & Fischer also touch upon the subject in their treatment of Lakoff's empirical method and approach to cognitive semantics:

"Lakoff argued that evidence for his proposals comes from the co-occurrence of linguistic phenomena. Reframed, his argument was for inductive analysis, a method of analysis which is the norm in the social sciences. Making generalisations based on a sample, then extrapolating those generalisations to the population is the basis of inductive scientific research, and the only viable method for the social sciences. Yet, without rigorous and testable techniques for establishing how representative or reliable a generalisation is, the 'results' remain merely hypotheses based on very small samples. Indeed...a sample was no more than the internal language knowledge of the linguist and, perhaps, a few colleagues."[176]

In Lakoff's defence, adopting empirical methods commonly used in the social sciences, is not entirely misplaced, in that the experiential basis of metaphors and language[177] have been shown to be not only related to our mutual, physical experience of the world, but also to the different levels of cultural context.[178] In fact, as stated in MWLB, in the chapter on how our conceptual system is grounded: "*It would be more correct to say that all experience is cultural through and through, that we experience our 'world' in such a way that our culture is already present in the very experience itself.*"[179] This does not, however, help the fact that the cognitive semantic system of metaphor is incomplete, and to some degree hypothetical, as also noted by Glynn & Fischer.

---

[176] Glynn & Fischer, p. 1-7.
[177] Think, for example, of the cultural connotations surrounding the later Wittgenstein's notion of common life-form.
[178] See chap. 8.2.
[179] Lakoff & Johnson, p. 57.

Another difficulty inherent to the practical application of the metaphorical systematicity of language to ontology engineering, is the unclear distinction between what Lakoff & Johnson call 'subcategorization' and 'metaphorical structuring'.[180] To be blunt, metaphorical structuring is a case of [x is y] as in [Happy is up], while subcategorization is a case of [x *is a kind of* y] as in [Civil war is a kind of conflict]. So while metaphorical structuring might potentially inspire altogether new types of ontology architectures, subcategorization implies a more traditional approach which would amount to a kind of specialised, taxonomy, as we also saw with the ontology architecture employed in the private company *Ankiro*.[181]

Nevertheless, it would still seem possible to correlate and classify these two types of structures into different subsystems of the same top-level ontology, but this requires us to be able to distinguish between the two. Herein lies the problem, which is made clear by Lakoff & Johnson when they say: "*The point is that subcategorization and metaphor are endpoints on a continuum.*"[182] This means that it can sometimes be difficult to determine whether a natural language instance amounts to a case of subcategorization, or to a case of metaphorical structuring.

Lakoff & Johnson dissolve this difficulty by introducing the notion of gestalts[183], stating that in unclear cases, an analysis of the elements entering an either sub-categorical or metaphorically structured relation should be analysed as gestalts to provide an answer to which of the two is the case. In addition, coherence within and between metaphors is also accorded the experiential basis of multidimensional gestalts by Lakoff & Johnson, who to that end argue that our concepts of objects, events and activities are characterizable as multidimensional gestalts whose dimensions emerge naturally from interaction with our environments.[184] This reliance on gestalt analysis for the disambiguation of subcategorization and metaphorical structuring, naturally complicates the practical application of the cognitive semantic theory of metaphor.

[180] Lakoff & Johnson, p. 83-86.
[181] See chap. 9.2.
[182] Lakoff & Johnson, p. 85.
[183] *Definition*: 'Gestalt'. A whole that, according to human understanding, is more basic than its parts.
[184] Lakoff & Johnson, p. 81-86.

11.3 The Contextual Granularity of Wittgenstein's Language Games

As per previous analyses,[185] the cognitive semantic theory of metaphor presented by Lakoff & Johnson was introduced in this thesis as an answer to how the Embodiment Thesis, as a theory of cognition, could apply to natural language in terms of fitting patterns of cognition to patterns of language – a theory which was well in tune with the later Wittgenstein's notion of common life-form. This notion of common life-form was, in turn, essential to achieving some degree of practical applicability to natural language, when taking an outset in the Wittgenstein's view of natural language as language games, since they implied such a high degree of contextuality that they would otherwise restrict the rules of natural language use to a single language instance, within its own specified social context.

The question is whether the overall systematicity of Lakoff & Johnson's cognitive semantic theory of metaphor applies equally well to all levels of context pertaining to a specific language instance. The problem faced when trying to answer this question is the contextual granularity implied by Wittgenstein's notion of language games. Our common life-form, which permits us to apply some degree of regularity to an otherwise unique set of rules pertaining to a particular language instance, applies to our physicality and the physical laws of our environment, as well as to our *culture* – a term which in itself is so broad, it could imply a number of different contextual levels, based on divisions by intrinsic group culture, social standing, political affinity, race, nationality and so on.

It may seem that a smooth scaling of context with no clearly defined transitional qualities is unproblematic. As mentioned earlier, Lakoff & Johnson even argue that all experience is essentially cultural, in spite of the way in which their classification of orientational metaphors seems to rely solely on our physical experience of the world. The problem is that one could argue that the contextual scope of different types of metaphors in language use varies greatly. For example, while many orientational metaphors would apply to almost any given cultural context, other more culture-sensitive types of metaphor would depend on the specific context of a given subculture.[186] Because the contextual scope of different types of metaphors varies in this way, so do their contextual levels of applicability in a top-level ontology based on their systematicity as evident in natural language use.

[185] Chap. 4 through 7.

[186] A simple example could be how the sentence "The wheels are really turning now!", could mean different things depending on whether the cultural setting is a high society group exchanging intellectual ideas, or a group of farm workers describing their physical work pace.

This smooth scaling of context goes all the way from our mutual, physical experience of existence to the private, deeply cultural experience of context in the individual. In order for an ontology to be able to structure its information so that it can be correctly applied, it would need some sensible, predefined levels of context which could help separate situational knowledge from the more general knowledge, as also argued in the previous chapter on domain-specific vs. top-level ontologies.[187] Lakoff & Johnson present no method for achieving such distinctions in MWLB. Lakoff does, however, come up with an answer of his own to how this kind of contextual segmentation might be achieved – something which is given further attention in the following subchapter.

As a side note, a problem of a similar nature is that even if we assume that the systematicity implied by the inductive analysis of very limited empirical data, carried out by Lakoff & Johnson in MWLB, is indicative of a system fully able to account for most of human cognition, there could still be more to the way we organize perceptions into world knowledge. Lakoff & Johnson's inability to distinguish clearly between subcategorization and metaphorical structuring without deeper gestalt analysis is a good example of this. It shows that there are structures other than metaphorical classification which are relevant to world knowledge – such as non-metaphorical subcategorization, which would also need to be handled by a top-level ontology, which aims at incorporating the language philosophical and cognitive semantic premise outlined in this thesis.

11.4 Contextual Segmentation using Frames

Professor George P. Lakoff has in recent years developed his cognitive semantic theory of metaphor into what is called "The Neural Theory of Language" (NTL). Together, Lakoff and Professor of Neurobiology Daniel Feldman, an associate of Lakoff's at Berkeley University, act as head of the "Center for the Neural Study of Language", where research is aimed at answering the question of how the physical brain, which is composed of neurons that function chemically, can give rise to human concepts and human language.[188]

As hinted at earlier, Lakoff aims to ground his work in cognitive semantics in neurobiology, and manages to do so with some success in his book 'Philosophy in The Flesh', published in 1999. Here he essentially identifies basic metaphors, which he calls 'primary metaphors', and shows how they are direct results of how the human brain works.[189] The [More is up; Less is down] metaphor mentioned earlier in this thesis, is a good example of this.[190]

---

[187] See chap. 9.4.
[188] http://icbs.berkeley.edu/natural_theory_lt.php
[189] Sowa.
[190] See chap. 10.3 + 10.6.

These primary metaphors are a result of subconscious neural learning, often during childhood, and are induced simply by one's functioning in the world.[191] They are, according to Lakoff, even independent of language. As such, they make for an important contribution to the issue of trying to discern the later Wittgenstein's notion of common life-form.[192]

The Neural Theory of Language has spawned several interesting research projects in several different areas, but especially relevant to the potential, contextual segmentation, the necessity of which was implied in the previous subchapter, is Lakoff's own theory of 'frames'. In a recent article from 2010, published in the journal "Environmental Communication", Lakoff provides the following overview and instructive example:

"One of the major results in the cognitive and brain sciences is that we think in terms of typically unconscious structures called 'frames'. Frames include semantic roles, relations between roles, and relations to other frames. A hospital frame, for example, includes the roles: Doctor, Nurse, Patient, Visitor, Receptionist, Operating Room, Recovery Room, Scalpel, etc. Among the relations are specifications of what happens in a hospital, e.g., Doctors operate on Patients in Operating Rooms with Scalpels. These structures are physically realized in neural circuits in the brain. All of our knowledge makes use of frames, and every word is defined through the frames it neurally activates. All thinking and talking involves 'framing.' And since frames come in systems, a single word typically activates not only its defining frame, but also much of the system its defining frame is in."[193]

Without going into too much detail concerning the intricacies of Lakoff's frame theory, it is possible to see how the notion of frames could be a good way of achieving some degree of contextual segmentation, since it holds the potential to delimit domains of language use and their related cognitive metaphorical concepts, without restricting specific words to one particular domain. Words mean different things in different contexts, and Lakoff's theory of frames recognizes this. In addition, the theory of frames as related to semantic roles and relations within and between frames, is in perfect accordance with the later Wittgenstein's understanding of meaning as *use*. Similarly, it correlates with the definition of world knowledge as including what each language user must know about the beliefs and goals of other language users.[194]

[191] The [Affection is warmth] metaphor is a good example of this. It is based on how the common experience of a child being held affectionately by a parent, results in the co-occurrence of increased temperature and affection.
[192] Lakoff & Johnson - 2003 Afterword, p. 256.
[193] Lakoff (2010), p. 70-81.
[194] See chap. 3.5.

Besides being one of the world's best known linguists and respected authorities in the area of Cognitive Science, Lakoff has also established himself as a respected, political commentator and critic, basing his arguments on his theory of frames, which he shows to be a crucial factor in the shaping of public opinion. Since 2002, he has consulted with the leaders of hundreds of advocacy groups on framing issues, lectured to large audiences across America, run dozens of workshops, spoken regularly on radio shows, TV shows and large political meetings, consulted with progressive pollsters and advertising agencies, and served as a consultant in major political campaigns.[195]

As a strictly academic endeavour, this thesis will, however, refrain from commenting on any social, environmental or political cognitive analyses undertaken by Lakoff.[196] Though, as an interesting side-note, one of the criticisms faced, when attempting to build a top-level ontology, is often that the pragmatism and cultural contextuality implied by the design and development of such a singular structure of representation does not allow for the exclusion of politics between persons or groups.[197]

11.5 Mental Space and Blending Theory

Having looked at the contextual segmentation made possible by Lakoff's theory of frames, which shows potential for the relative ordering of linguistic and cognitive elements – here among metaphorical concepts, it might also be a good idea to have take a closer look at the compositionally inherent qualities of metaphors. As we saw earlier, the structural properties of metaphors can vary in degrees of abstraction[198], but on the basic level, a metaphor amounts to understanding one thing in terms of another. However, according to Gilles Fauconnier's theory of conceptual blending, metaphors can be further deconstructed into mental spaces and mappings between these spaces, which results in one or more 'blends'.[199]

Mental spaces are domains of experience that are involved in cognitive meaning construction. They are constructed on the basis of linguistic expressions whose combination and pragmatic context determine the actual meaning construction. By this account, a linguistic expression does not have meaning on its own – it has meaning potential. When the grammatical information in it is applied to an actual cognitive configuration, it often yields several new possible configurations. One will be produced, creating a new step in the construction of the underlying discourse, generating meaning.[200]

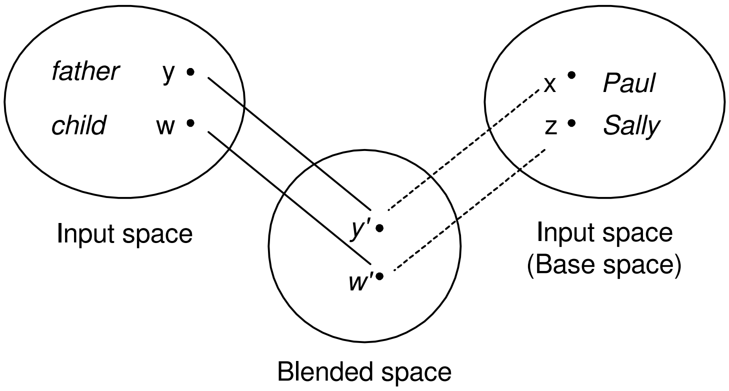
---

[195] http://www.chelseagreen.com/authors/george_lakoff
[196] His most recent book "Don't Think of an Elephant: How Democrats and Progressives Can Win: Solutions from George Lakoff" is a good example of Lakoff's political engagement.
[197] http://en.wikipedia.org/wiki/Foundation_ontology
[198] An example of this is the way structural metaphors are metaphorically emergent, based exclusively on abstract elements. See also chap. 8.2.
[199] Fauconnier & Turner, p. 39-58.
[200] Fauconnier & Turner, p. 5-20.

In this way, the meaning of a sentence is relative to its understanding – a view well in tune with the theories of Lakoff, as well as the later Wittgenstein. In fact, Lakoff and Fauconnier have even jointly published a paper with the title "On Metaphor and Blending", explaining how their cognitive semantic theories are largely compatible, and not to be seen as competing at any level.[201]

In short, Lakoff and Fauconnier have very similar language philosophical premises and views of the role and purpose of cognitive semantics in the analysis and interpretation of natural language. The difference between the two is that Fauconnier simply assumes many of the structures for which Lakoff argues, such as mappings from one domain to another, frames, metaphors, and so on.[202] Fauconnier then applies his theory of mental space blending to very particular linguistic examples on a very detailed level. This permits a deeper, cognitive analysis of metaphors. Fig. 11.1 shows an example of Fauconnier's instructive analysis of the simple declarative sentence "Paul is the father of Sally", taken from his most recent work: 'The Way We Think – Conceptual Blending and The Mind's Hidden Complexities' (2003), which he wrote in collaboration with Professor of Cognitive Science Mark Turner.[203]

FIGURE 11.1

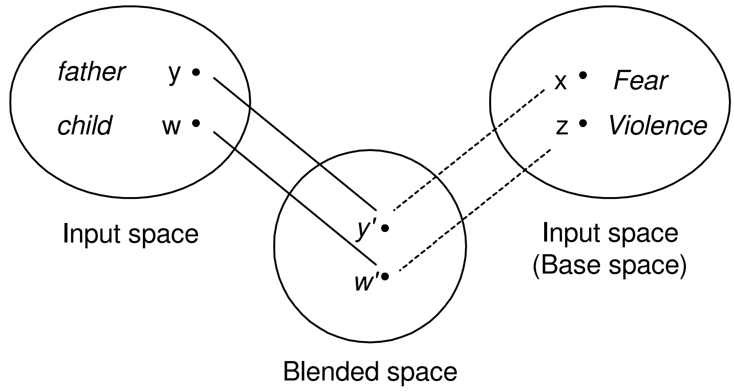


The base input space, consisting of 'Paul' and 'Sally', and the other input space with the roles 'father' and 'child' are linked in the blended space. The grammar of the sentence defines what are roles and actors, and the resulting blend is what Fauconnier calls an 'XYZ network', as in 'X is the Y of Z'. This is one of many examples of how he lays out rules on how blends we use to understand ideas are formed. This is obviously not a metaphor, but it illustrates the assignment of roles to actors, and the way they integrate in a blended space. An important part of Fauconnier's theory on conceptual blending is that it applies equally well to very literal meanings as illustrated in fig. 11.1 as it does to metaphorical meanings.

201 Fauconnier & Lakoff, p 1-5.
202 Lakoff & Johnson – 2003 Afterword, p. 261.
203 Fauconnier & Turner, p. 151.

For example, fig. 11.2 presents an analysis of the metaphorical sentence "Fear is the father of violence", which clearly shows how the analytical composition is identical to the previous example, which was based on a similar grammatical structure.

FIGURE 11.2



This is an important point and strength to Fauconnier's cognitive semantic theory. If an analysis of metaphors expressed in natural language can be reduced in this way, this also greatly reduces the number of difficulties involved in providing a deeper analysis of metaphors, which might be necessary when tackling the problem of rigid metaphor representation aimed at different types of computational models, and on top of that, the problem of disambiguating subcategorization and metaphorical structuring will also have been partly solved.[204] It is as Fauconnier himself puts it in his chapter on "Composition of Metaphoric Integration":

"*The language forms that lead to intuitively literal meanings can also give us intuitively metaphorical meanings that seem to belong to radically different kinds of thinking. Yet those identical forms are prompting for identical mapping schemes to guide those radically different constructions of meaning. And those mapping schemes compose in identical ways, regardlessly of whether the ultimate meanings are flatly literal, poetically metaphorical, scientifically analogical, surrealistically suggestive, or opaque.*"[205]

The complexity of the mapping schemes and resulting structures and blends of mental spaces successively introduced by Fauconnier, of course rises noticeably as he goes on to establish additional rules and facets which build upon his theory of conceptual blending, but the basic idea behind them is still essentially the same. However, it is beyond the scope of this thesis to go into further details on his work, which is to say the least, very extensive in its intricacies and complexities. The aim was here, simply to show how Fauconnier's cognitive semantic theory could potentially complement the representation, analysis and disambiguation of Lakoff & Johnson's work concerning the metaphorical systematicity of cognition and natural language.

[204] Fauconnier & Turner, p. 142-143 + 154-160.
[205] Fauconnier & Turner, p. 154.

12 A New Approach to Ontology Engineering and World Knowledge

In this chapter, the results, claims, ideas and suggested areas for further research set forth in this thesis are reviewed and juxtaposed. This is then followed by a brief note on their academic validity, as seen in light of various expert interviews and correspondences, coordinated by the author of this thesis.

12.1 Results and Suggested Areas for Further Research

With an analysis of both the potential and limitations of Lakoff & Johnson's cognitive semantic theory of metaphorical systematicity in place, it is time to review how these results can help guide the research and development of a top-level ontology capable of endowing NLP systems with world knowledge, which the author of this thesis, in the following, will refer to as a 'World Knowledge Ontology' (WKO).

From what we have seen so far, it seems that the metaphorical systematicity of language holds significant potential for creating a WKO which is able to handle the many metaphorical concepts which have been shown to be an inherent part of human cognition and natural language use. As an example of this potential, let us suppose that it was to be utilized by a dialogue system. The basic architecture of a dialogue system as described by Ginzburg and Fernández in their article on 'Computational Models of Dialogue' for 'The Handbook of Computational Linguistics and Natural Language Processing' from 2010, can be seen in fig. 12.1. The WKO would then replace what they call "Domain/task knowledge".

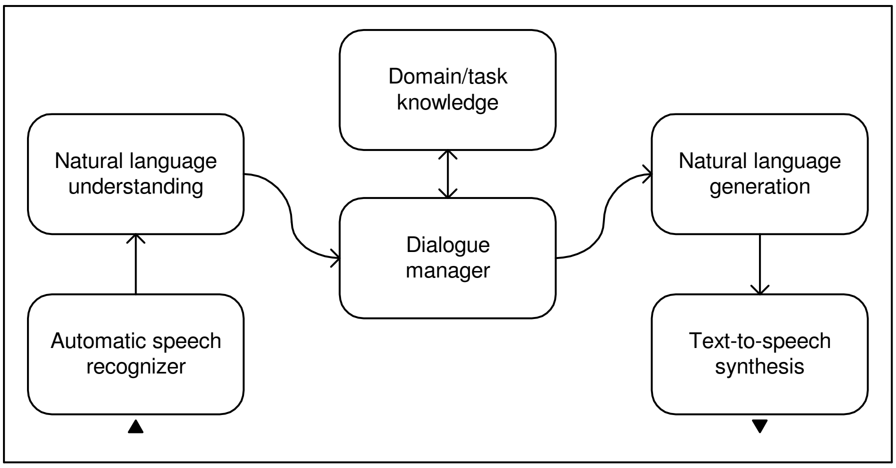


FIGURE 12.1

One could imagine the benefits of a dialogue system which was able to interpret the emotional state of a human conversation partner by way of metaphors like [Happy is up; Sad is down]. Such information would be relevant to interpreting not only the emotional state of the human counterpart, but also the intentionality with which it was potentially associated, since reason is often fuelled by and can not be separated from emotion, as Lakoff also argues in his book "Philosophy in The Flesh – The Embodied Mind and Its Challenge to Western Thought".[206] In a similar way, the underlying meaning of idiomatic expressions[207], which are otherwise difficult for NLP systems to handle, might also be extracted to better facilitate natural language understanding.

However, since such metaphorical systematicity is based on inductive analyses[208], ontology engineers should be careful not to view the cognitive semantic theory of metaphor as a complete solution to the systematization of natural language involved in WKO development. The notion is further supported by the way in which non-metaphorical subcategorization is likely to be an essential part of WKO design.[209] To this end, a separation of metaphorical and non-metaphorical content into different sub-systems of the overall WKO, as illustrated in fig. 12.2, could, in theory, prove useful.

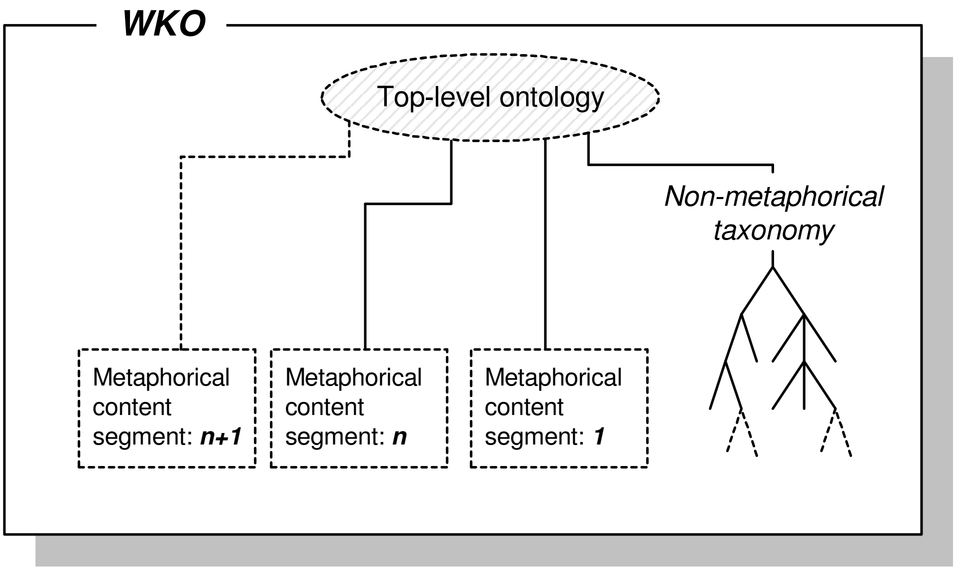


FIGURE 12.2

[206] Lakoff (1999), p. 414.
[207] In linguistics, idiomatic expressions are mainly figures of speech contradicting the principle of compositionality, i.e. whose sense means something different from what the individual words literally imply.
[208] As discussed in chap.11.2.
[209] Though the merits of complex metaphor interpretation are many, a WKO should also be able to classify a "cow" as a type of "mammal", a "car" as a type of "vehicle", and so on.

While the non-metaphorical content is perhaps handled by a taxonomy, which is the common approach in current, practical ontology engineering[210], the metaphorically structured content could be handled by other subsystems, possibly segmented according to contextual scope. As we recall, Lakoff even presents us with a possible strategy for achieving metaphorical segmentation, with his theory of frames, which also seems compatible with the notion of domain-specific ontologies serving as subsystems in a top-level ontology, when we think of his hospital example.[211]

However, Lakoff & Johnson's difficulties in disambiguating subcategorization and metaphorical structuring in a systematic way, poses a problem in terms of the divisions assumed possible by the hypothetical WKO model in fig. 12.2. One approach to resolving this issue is to further deconstruct metaphors into their inherent compositional properties in the form of mental spaces and blends, in accordance with Fauconnier's 'Mental Space and Blending Theory', which facilitates an analysis of metaphorical and non-metaphorical content using the same set of mapping schemes. In addition, it could also be an important tool when it comes to metaphor *representation* – an issue which the POC's use of an ANN as a computational model showed to be of prime importance.

What is more, the POC developed in this thesis also showed that even though the cognitive semantic theories of both Lakoff and Fauconnier are grounded in a neurobiological account of how the human brain works, this does not necessarily mean that devising 'brain-simulations' is the way forward, in that ANNs were found to be more apt to representing the underling *cognitive* patterns that account for natural language use.

Based on these findings, and in regard to the overall question of how to approach the task of developing a top-level ontology able to endow NLP systems with world knowledge, this thesis therefore recommends developing these ideas further, in order to come up with a candidate WKO architecture which is able to take advantage of the metaphorical systematicity of human cognition and natural language. This of course requires further research into the actual scope of the metaphorical systematicity of natural language, as well as an investigation into possible metaphor representations, using Fauconnier's Mental Space and Blending Theory, which could also hold potential for supporting different candidate WKO architectures and computational models.

---

[210] This was also seen with the ontology architecture employed at Ankiro – see chap. 9.2.
[211] More on this in chap. 11.4.

12.2 Academic Validation and Verification

In accordance with the general approach to NLP proposed earlier in this thesis, a WKO will then define the lower levels of language analysis, which should be implemented successively until the lowest level is reached, finally yielding a functional solution, which could potentially go beyond the limits in scope and lack of advancements which characterise the field of NLP at present. This notion is based on the theory set forth by the author of this thesis, that since the field of NLP has seen few significant improvements within the past years, the tendency towards starting at the lowest level of language analysis and advancing upwards may be suboptimal, and that in order to achieve significant improvements within the field of NLP, it may be necessary to 'rethink' NLP design strategies and instead start at the highest level of language analysis, i.e. world knowledge.[212]

In an interview[213] given to the author of this thesis, Prof. Bolette S. Pedersen from the Copenhagen Center for Language Technology (CST), who is also the project manager of the Danish version of WordNet and various other academic projects and associations concerned with NLP[214], agreed that the field of NLP seems to have "hit a wall". She also agreed that while there is a strong propensity towards low level, statistical language analysis within the general field of NLP, there have been few efforts towards approaching it from an outset in the higher levels of language analysis, such as world knowledge in the form of ontology engineering. She even added that she herself, on occasion, had thought of this, as an approach to NLP which could hold potential for new advances.

The true validation of this thesis which must be mentioned is, however, concerned with the very notion of developing a *top-level ontology* based on the cognitive semantic theory of metaphor and its outset in embodied cognition, inferred by the language philosophical analysis of Wittgenstein's two major language philosophical works. This notion is the central, but also the most controversial part of this thesis, in that it was not possible to find any current record of academic efforts towards this type of ontology engineering, when researching the subject. This prompted the author of this thesis to write an inquiry to Lakoff himself, asking if he considered "*…the cognitive theory behind (his) view of language and the human mind as largely metaphorical in nature, to be a solid ground from which it would be possible **to build a top-level ontology for use with natural language processing systems***".[215] To this Lakoff gave the following reply:

[212] See fig. 3.3.
[213] See Appendix C.
[214] http://www.cst.dk/bolette/my.html
[215] See also Appendix D.

> HI Erik,
>
> The answer is yes, and our group at ICSI at Berkeley has just gotten a grant to begin doing just that. We should start in January, and if the grant is continued and not cut by our Congress, we should have five years to do the initial work. It is, of course, a huge enterprise, enough to keep many people busy for many years.
>
> Welcome to the enterprise!
>
> I'm attaching a first draft of an overview paper I'm working on. Enjoy!
>
> George
>
> *Oct 27, 2011*

This serves as a serious validation of the approach set forth in this thesis. Not only does the author of the cognitive semantic theory of metaphor agree with the approach towards building a top-level ontology suggested here; he also discloses that considerable efforts towards researching "*just that*" have just recently attained a 5-year grant, and will commence within a few months at The International Computer Science Institute (ICSI) at Berkeley University. In his first draft overview paper, which he so generously attaches to his e-mail reply, he even goes on to elaborate on the relations between metaphor, frames and conceptual blending, further supporting some of the basic ideas set forth in this thesis.[216]

13 Conclusion

This thesis started out by establishing how there seems to be a general propensity towards low level, statistical language analysis within the general field of NLP, with statistical machine translation serving as a good example of this trend. A brief introduction to the different levels of natural language analysis was then used to substantiate this notion further, and to contextualize the highest level of language analysis: world knowledge – a term central to the first contention of this thesis; namely that world knowledge can be seen as a core of language analysis that affects all outer (lower) levels[217], and that it in this way could be an issue that must be dealt with first in order to know how to approach lower layers, when attempting an optimal design strategy, which could potentially lead to new advancements in NLP – a field currently characterized by stagnation.

---

[216] Since Lakoff's unpublished overview paper was marked "*FIRST DRAFT - DO NOT DISTRIBUTE*", the author of this thesis has not quoted it directly, but instead placed it on the CD accompanying the 1st edition of the printed version of this thesis, to be handed in at ITU on December 1, 2011. This thesis has been classified as *Confidential* to the extent that the contents of the CD are not to be distributed. This was done to ensure that this thesis, at a later point in time, could be published in its entirety, with the exclusion of said overview paper.

[217] See fig. 3.3.

As a way of taking steps towards such a 'top-down' approach to NLP by looking deeper into the issue of world knowledge, and also overcoming the scepticism expressed in the problem of NLU, the strategy of employing analytical language philosophy was suggested, based on the compelling relevance of its four central problems. To reduce the scope of this endeavour, this analysis was based exclusively on the two major works of Ludwig Wittgenstein, with each work representing contrasting views on language and meaning. Analyses of both works, as held up against the field of NLP, led to the rejection of objectivism as well as the disembodied view of cognition predominant in traditional cognitive science and Western philosophy, revealing the need for a method for analysing the patterns in language and cognition that correlate with our common life-form – a concept which Wittgenstein identifies as crucial to overcoming the hyper-contextuality of natural language use.

As a near perfect answer to this cognitive and language philosophical premise, Lakoff & Johnson's cognitive semantic theory of metaphor and the metaphorical systematicity of cognition and natural language was introduced, complementing and building upon the theories of the later Wittgenstein, thus forming the final language philosophical and cognitive outset and premise of this thesis, in which Lakoff's theory of frames also showed potential for achieving some degree of contextual segmentation. As a practical approach to endowing NLP systems with world knowledge, ontology engineering, specifically developing a top-level ontology, was suggested following a brief introduction to this relatively new field, which included a real-life example from the company *Ankiro*. The issue of developing a top-level ontology based on the established language philosophical and cognitive premise, motivated the development of a POC, which was to demonstrate the concept of getting a computational model, in this case an ANN, to handle the metaphorical systematicity of language.

Besides succeeding in this, the POC also elucidated the potential of ANNs in performing metaphorical classification through pattern recognition which could reflect world knowledge based on patterns in human cognition[218] – something for which ANNs seemed better suited than the kind of 'human brain simulations' that might intuitively follow from recent progress in the grounding of Lakoff's cognitive semantic theories in neurobiology. However, it also stressed the importance of metaphor *representation* – an issue further complicated by Lakoff & Johnson's difficulties in disambiguating subcategorization and metaphorical structuring. As a possible solution hereto, Fauconnier's highly compatible mental space and blending theory showed promise in its ability to deconstruct both metaphorical and non-metaphorical content, using the same mapping schemes.

---

[218] See fig. 10.8.

In closing, this thesis suggests a new approach to the field of NLP, where the highest level of natural language analysis – world knowledge – is addressed first, and then used to determine lower levels of language analysis successively, in a top-down fashion. A method of approach with this aim has been established, with an outset in analytical language philosophy and cognitive semantics, which together form the language philosophical and cognitive semantic premise which this thesis suggests be employed, when attempting to design a top-level ontology aimed at endowing NLP systems with world knowledge. The results of this thesis are not intended as an actual candidate architecture, which is in itself a huge undertaking, but merely as a theoretical framework for researching it. This framework may seem somewhat controversial in its originality, but it received strong validation from Lakoff himself in his revelation of upcoming research of precisely this nature at Berkeley University. In the future, it will be exciting to see if this research will lead to the implementation of a new type of top-level ontology which in turn, might potentially spur new advances within the general field of NLP.

References

Ai, Hua and Litman, Diana. Comparing Real-Real, Simulated-Simulated, and Simulated-Real Spoken Dialogue Corpora. Pittsburgh University, 2011.

Allen, James. Natural Language Understanding.
Redwood City: The Benjamin/Cummings Publishing Company, Inc., 1998

Clark, Fox & Lappin. The Handbook of Computational Linguistics and Natural Language Processing. Oxford: Wiley-Blackwell, 2010.

Fauconnier, Gilles and Turner, Mark. The Way We Think – Conceptual Blending and The Mind's Hidden Complexities. New York: Perseus Books Group, 2003.

Glynn, Dylan and Fischer, Kerstin. Quantitative Methods in Cognitive Semantics: Corpus-driven Approaches. Göttingen: Walter de Gruyter & Co., 2010.

Gómez-Pérez, Fernández & Chorcho. Ontological Engineering.
London: Springer-Verlag, 2004.

Heaton, Jeff. Introduction to Neural Networks with Java.
St. Louis: Heaton Research, Inc., 2005

Hofmann, Th. R. Realms of Meaning, An Introduction to Semantics.
New York: Addison Wesley, 1993.

Hutchins, W. John and Somers, Harold L. An Introduction to Machine Translation.
London: Academic Press, 1992.

Jurafsky & Martin. Speech and Language Processing.
New Jersey: Prentice-Hall, Inc., 2008.

Kruchten, Philippe. The Rational Unified Process: An Introduction, 3rd edition.
New York: Addison-Wesley, 2004.

Lackey, Douglas P. The Baruch Poll of Great Philosophy in the Twentieth Century.
The Philosophical Forum, 30: John Wiley & Sons, Inc., 2002.

Lakoff, George. Women, Fire, and Dangerous Things – What Categories Reveal about The Mind. Chicago: The University of Chicago Press, 1990.

Lakoff, George. Philosophy in the Flesh : The Embodied Mind and Its Challenge to Western Thought. New York: Perseus Books Group, 1999.

Lakoff & Johnson. Metaphors We Live By.
Chicago: The University of Chicago Press, 2003.

Lakoff, George. Why it Matters How We Frame the Environment.
Routledge: Env. Communication: A Journal of Nature and Culture, Volume 4, 2010.

Lübcke, Poul. Engelsk og Amerikansk Filosofi, Videnskab og Sprog.
Copenhagen: Politikens Forlag A/S, 2003.

Martinich, A. P. The Philosophy of Language.
New York: Oxford University Press, 2001.

Mitkov, Ruslan. The Oxford Handbook of Computational Linguistics.
Hampshire: Oxford University Press, 2005

Negnevitsky, Michael: Artificial Intelligence – A Guide to Intelligent Systems.
Essex: Addison-Wesley, 2005

Reusch, Bernd. Computational Intelligence – Theory and Applications.
Dortmund: Springer, 2006.

Rizzolatti, Giacomo and Craighero, Laila. The Mirror-Neuron System.
Annual Review of Neuroscience, Vol. 27: 169-192, 2004.

Ross, Peter. Neural Networks – An Introduction.
University of Edinburgh, 1999

Russell, Stuart and Norvig, Peter. Artificial Intelligence – A Modern Approach.
New Jersey: Prentice Hall, 2003.

Sluga, Hans and Stern, David G. The Cambridge Companion to Wittgenstein.
Cambridge: Cambridge University Press, 1999.

Sowa, John, F. Review of Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought. Computational Linguistics, vol. 25, no. 4, 1999.

Wittgenstein, Ludwig. Tractatus Logico-Philosophicus.
New York: Routledge, 2005.

Wittgenstein, Ludwig. Philosophical Investigations.
Oxford: Blackwell, 2005.

Zadeh, Lofti. From Search Engines to Question-Answering Systems: The Problems of World Knowledge, Relevance, Deduction, and Precisiation. Berlin: Springer, 2006.

Chelsea Green Publishing: “Our Chelsea Green Authors: George Lakoff” http://www.chelseagreen.com/authors/george_lakoff 2011.

“Cognitive Science”.
http://en.wikipedia.org/wiki/Cognitive_science 2011.

Fauconnier, Gilles and Lakoff, George. On Metaphor and Blending. The University of California, San Diego: The Department of Cognitive Science http://www.cogsci.ucsd.edu/~coulson/spaces/GG-final-1.pdf 2011.

Stanford Encyclopedia of Philosophy: “Cognitive Science, part 3: Representation and Computation” http://plato.stanford.edu/entries/cognitive-science 2011.

Stanford Encyclopedia of Philosophy: “Embodied Cognition” http://plato.stanford.edu/entries/embodied-cognition 2011.

Stanford Encyclopedia of Philosophy: “René Descartes” http://plato.stanford.edu/entries/descartes 2011.

“Upper Ontology (information science)” http://en.wikipedia.org/wiki/Foundation_ontology 2011.

“What’s New in Statistical Machine Translation”. http://people.csail.mit.edu/koehn/publications 2006.

Appendices

Appendix A: ANN Prediction Rates

```
Prediction rate at iteration 0: 0.346
Prediction rate at iteration 1000: 0.323
Prediction rate at iteration 2000: 0.33
Prediction rate at iteration 3000: 0.348
Prediction rate at iteration 4000: 0.321
Prediction rate at iteration 5000: 0.32
Prediction rate at iteration 6000: 0.592
Prediction rate at iteration 7000: 0.353
Prediction rate at iteration 8000: 0.664
Prediction rate at iteration 9000: 0.792
Prediction rate at iteration 10000: 1.0

Prediction rate at iteration 0: 0.331
Prediction rate at iteration 1000: 0.334
Prediction rate at iteration 2000: 0.364
Prediction rate at iteration 3000: 0.463
Prediction rate at iteration 4000: 0.329
Prediction rate at iteration 5000: 0.429
Prediction rate at iteration 6000: 0.336
Prediction rate at iteration 7000: 0.448
Prediction rate at iteration 8000: 0.763
Prediction rate at iteration 9000: 0.914
Prediction rate at iteration 10000: 1.0

Prediction rate at iteration 0: 0.347
Prediction rate at iteration 1000: 0.315
Prediction rate at iteration 2000: 0.321
Prediction rate at iteration 3000: 0.338
Prediction rate at iteration 4000: 0.337
Prediction rate at iteration 5000: 0.588
Prediction rate at iteration 6000: 0.685
Prediction rate at iteration 7000: 0.602
Prediction rate at iteration 8000: 0.817
Prediction rate at iteration 9000: 0.845
Prediction rate at iteration 10000: 0.961

Prediction rate at iteration 0: 0.363
Prediction rate at iteration 1000: 0.352
Prediction rate at iteration 2000: 0.327
Prediction rate at iteration 3000: 0.315
Prediction rate at iteration 4000: 0.348
Prediction rate at iteration 5000: 0.447
Prediction rate at iteration 6000: 0.717
Prediction rate at iteration 7000: 0.92
Prediction rate at iteration 8000: 1.0
Prediction rate at iteration 9000: 1.0
Prediction rate at iteration 10000: 1.0
```

Appendix B: Trained Neural Network Test Output

```
[More is up; Less is down]: ...Then they went down into the garden ...
[Con. is up; Unc. is down]: ...her She woke up and was quite ...
[Con. is up; Unc. is down]: ...pure bliss and fell asleep on the ...
[Con. is up; Unc. is down]: ...I finally do fall asleep I'm soon ...
[Con. is up; Unc. is down]: ...And she woke up right away and ...
[More is up; Less is down]: ...Claus and went up and knocked on ...
[More is up; Less is down]: ...if I went down to the bottom ...
[Con. is up; Unc. is down]: ...closed and he fell asleep with his ...
[Con. is up; Unc. is down]: ...Then he woke up and all the ...
[Con. is up; Unc. is down]: ...He didn't wake up until the sun ...
[More is up; Less is down]: ...was They went up to the hall ...
[Con. is up; Unc. is down]: ...tired Johannes woke up very early in ...
[More is up; Less is down]: ...as she went under the water the ...
[More is up; Less is down]: ...the sun went down Elisa saw eleven ...
[More is up; Less is down]: ...the sun went down they stood there ...
[Con. is up; Unc. is down]: ...when Elisa woke up She thought she ...
[More is up; Less is down]: ...the sun went down her brothers came ...
[Happy is up; Sad is down]: ...kept his spirits up and smeared his ...
[More is up; Less is down]: ...man Listen go down and ask him ...
[More is up; Less is down]: ...I'd better go down and see? ...
[More is up; Less is down]: ...they didn't go up to the top ...
[Con. is up; Unc. is down]: ...the foreigner woke up He was sleeping ...
[Con. is up; Unc. is down]: ...wide and woke up He leaped to ...
[More is up; Less is down]: ...and could go up there whenever they ...
[Con. is up; Unc. is down]: ...sea she woke up and felt a ...
[More is up; Less is down]: ...night she went down the wide marble ...
[Con. is up; Unc. is down]: ...or she'll wake up, said the ...
[Con. is up; Unc. is down]: ...little thing woke up quite early in ...
[More is up; Less is down]: ...a shiver go down his spine. What ...
[Con. is up; Unc. is down]: ...when I wake up I seem to ...
[Con. is up; Unc. is down]: ...here and then fall asleep That was ...
[More is up; Less is down]: ...the sun went down and a short ...
[Con. is up; Unc. is down]: ...while I woke up to find the ...
[More is up; Less is down]: ...it all went up in flames The ...
[Con. is up; Unc. is down]: ...the prince woke up and was not ...
[More is up; Less is down]: ...The sun went down The whole sky ...
[Con. is up; Unc. is down]: ...of it he fell asleep It was ...
[Con. is up; Unc. is down]: ...marble monument and fell asleep It was ...
[More is up; Less is down]: ...gallery They went up the same steps ...
[More is up; Less is down]: ...the sun went down He had also ...
[Con. is up; Unc. is down]: ...little elf couldn't fall asleep for it ...
[Con. is up; Unc. is down]: ...when she woke up Oh what salty ...
[Con. is up; Unc. is down]: ...the leaves and fell asleep The sun ...
```

(Continued on next page)

[More is up; Less is down]: ...snow and go down again to the ...
[More is up; Less is down]: ...child to go down to Italy Just ...
[More is up; Less is down]: ...bag and went up the mountain the ...
[More is up; Less is down]: ...the path went down towards the Rhone ...
[More is up; Less is down]: ...others They went down to Chillon the ...
[Con. is up; Unc. is down]: ...But she could fall asleep and sleep ...
[Con. is up; Unc. is down]: ...sounded Babette woke up The dream was ...
[More is up; Less is down]: ...had to go down many many steps ...
[More is up; Less is down]: ...seen and go down to the river ...
[More is up; Less is down]: ...the sun went down behind the big ...
[Con. is up; Unc. is down]: ...when she woke up in the morning ...
[Con. is up; Unc. is down]: ...When she woke up towards morning she ...
[More is up; Less is down]: ...The house went up in flames and ...
[More is up; Less is down]: ...should just go down to her and ...
[More is up; Less is down]: ...dare to go down there when she’s ...
[Con. is up; Unc. is down]: ...eyelids and I fell asleep I dreamed ...
[More is up; Less is down]: ...princess who went down to the bog ...
[More is up; Less is down]: ...the sun went down and the sky ...
[Happy is up; Sad is down]: ...He whose spirit rose above antiquity’s Gods ...
[More is up; Less is down]: ...The gentry went up to town The ...
[Con. is up; Unc. is down]: ...pond he woke up half-dead If he ...
[More is up; Less is down]: ...whole family went down by the canal ...
[Con. is up; Unc. is down]: ...The Portuguese woke up too shifted about ...
[Con. is up; Unc. is down]: ...the Portuguese woke up he was standing ...
[More is up; Less is down]: ...The sun went down and the hens ...
[Con. is up; Unc. is down]: ...respectable Then she fell asleep It was ...
[Con. is up; Unc. is down]: ...dead? Wake up wake up...
[Con. is up; Unc. is down]: ...Wake up wake up? the rooster ...
[Con. is up; Unc. is down]: ...shall one day fall asleep and die ...
[Con. is up; Unc. is down]: ...called I’ll wake up and be a ...
[Con. is up; Unc. is down]: ...awhile and then fell asleep Gracious What ...
[Con. is up; Unc. is down]: ...said Soon he fell asleep and dreamed ...
[Con. is up; Unc. is down]: ...dung beetle woke up he crept out ...
[More is up; Less is down]: ...The sun went down and the full ...
[More is up; Less is down]: ...I would go down and slide on ...
[More is up; Less is down]: ...and I went down to the basement ...
[Con. is up; Unc. is down]: ...the canary woke up and started to ...
[More is up; Less is down]: ...little boy went down right away to ...
[More is up; Less is down]: ...so I went down the drain, washed up ...
[More is up; Less is down]: ...of paper went down to the gate ...
[More is up; Less is down]: ...stairway that went up a long way ...
[Con. is up; Unc. is down]: ...if he woke up from a trance ...
[Con. is up; Unc. is down]: ...Dane didn’t wake up no matter how ...
[Con. is up; Unc. is down]: ...fresh wind and fell asleep which wasn’t ...
[More is up; Less is down]: ...the peddler went up to the manor ...
[More is up; Less is down]: ...seventy-year-old Excellency went down to the old ...

Appendix C: Interview Details

Interview 1

Date: August 11th, 2011.
Location: Centre for Language Technology, Njalsgade 140, DK-2300 Copenhagen S.
Topic: Current state of and developments within the field of NLP and ontology engineering.
Interviewer: Thesis Student Erik D. Johnson.
Interviewee: Prof. Bolette S. Pedersen from the Copenhagen Centre for Language Technology.
Contact Information: bspedersen@hum.ku.dk , http://www.cst.dk/bolette

Interview 2

Date: October 14th, 2011.
Location: Ankiro, Nybrogade 28, DK-1203 Copenhagen K.
Topic: The theoretical and practical application of ontology engineering.
Interviewer: Thesis Student Erik D. Johnson.
Interviewee: Head of Language Department Louise Bie Larsen.
Contact Information: lbl@ankiro.dk , http://ankiro.dk/content/dk/kontakt
Interviewee: Chief Technology Officer Steen Bøhm Andersen.
Contact Information: sba@ankiro.dk , http://ankiro.dk/content/dk/kontakt

Appendix D: Lakoff's e-mail Correspondence

On Thu, Oct 27, 2011 at 6:27 AM, George Lakoff <george.lakoff@gmail.com> wrote:
HI Erik,

The answer is yes, and our group at ICSI at Berkeley has just gotten a grant to begin doing just that. We should start in January, and if the grant is continued and not cut by our Congress, we should have five years to do the initial work. It is, of course, a huge enterprise, enough to keep many people busy for many years. Welcome to the enterprise!

I'm attaching a first draft of an overview paper I'm working on. Enjoy!

George

>*On Wed, Oct 26, 2011 at 4:29 AM, Erik David Johnson <erk@erk.dk> wrote:*
>*Dear Professor George P. Lakoff,*
>
>*My name is Erik David Johnson, and I am currently writing my thesis on*
>*Cognitive Semantics, Ontology Engineering and Advanced AI, at the*
>*IT-University of Copenhagen. I would very much appreciate if you could*
>*tell me your thoughts on a particular issue which is related to your*
>*theory of experientialism and the metaphorical systematicity of language.*
>
>*My question is: To what extent do you consider the cognitive theory behind*
>*your view of language and the human mind as largely metaphorical in*
>*nature, to be a solid ground from which it would be possible to build a*
>*top-level ontology for use with natural language processing systems?*
>
>*Any thoughts you could share with me regarding this issue would be deeply*
>*appreciated,*
>
>*Kind Regards,*
>
>*Erik David Johnson*