How Is Meaning Grounded in Dictionary Definitions?

A. Blondin Massé
Laboratoire de combinatoire et d'informatique mathématique
Université du Québec à Montréal
Montréal (QC), CANADA H3C 3P8
alexandre.blondin.masse@gmail.com

G. Chicoisne, Y. Gargouri, S. Harnad, O. Picard
Institut des sciences cognitives
Université du Québec à Montréal
Montréal (QC), CANADA H3C 3P8
chicoisne.guillaume@uqam.ca, yassinegargouri@hotmail.com
harnad@ecs.soton.ac.uk, olivierpicard18@hotmail.com

O. Marcotte
Groupe d'études et de recherche en analyse des décisions (GERAD) and UQÀM
HEC Montréal
Montréal (Québec) Canada H3T 2A7
Odile.Marcotte@gerad.ca

Abstract

Meaning cannot be based on dictionary definitions all the way down: at some point the circularity of definitions must be broken in some way, by grounding the meanings of certain words in sensorimotor categories learned from experience or shaped by evolution. This is the "symbol grounding problem". We introduce the concept of a *reachable* set — a larger vocabulary whose meanings can be learned from a smaller vocabulary through definition alone, as long as the meanings of the smaller vocabulary are themselves already grounded. We provide simple algorithms to compute reachable sets for any given dictionary.

1 Introduction

We know from the 19th century philosopher-mathematician Frege that the *referent* and the *meaning* (or "sense") of a word (or phrase) are not the same thing: two different words or phrases can refer to the very same object without having the same meaning (Frege, 1948): "George W. Bush" and "the current president of the United States of America" have the same referent but a different meaning. So do "human females" and "daughters". And "things that are bigger than a breadbox" and "things that are not the size of a breadbox or smaller".

A word's "extension" is the set of things to which it refers, and its "intension" is the rule for defining what things fall within its extension.. A word's meaning is hence something closer to *a rule for picking out its referent.* Is the dictionary definition of a word, then, its meaning?

Clearly, if we do not know the meaning of a word, we look up its definition in a dictionary. But what if we do not know the meaning of any of the words in its dictionary definition? And what if we don't know the meanings of the words in the definitions of the words defining those words, and so on? This is a problem of infinite regress, called the "symbol grounding problem" (Harnad, 1990; Harnad, 2003): the meanings of words in dictionary definitions are, in and of themselves, ungrounded. The meanings of some of the words, at least, have to be grounded by some means other than dictionary definition look-up.

How are word meanings grounded? Almost certainly in the sensorimotor capacity to pick out their referents (Harnad, 2005). Knowing *what to do with what* is not a matter of definition but of adaptive sensorimotor interaction between autonomous, behaving systems and categories of "objects" (including individuals, kinds, events, actions, traits and states). Our embodied sensorimotor systems can also be described as applying information processing rules to inputs in order to generate the right outputs, just as a thermostat defending a temperature of 20 degrees can be. But this dynamic process is in no useful way analogous to looking up a definition in a dictionary.

We will not be discussing sensorimotor grounding (Barsalou, 2008; Glenberg & Robertson, 2002; Steels, 2007) in this paper. We will assume some sort of grounding as given: when we consult a dictionary, we already know the meanings of at least some words,

Coling 2008: Proceedings of 3rd Textgraphs workshop on Graph-Based Algorithms in Natural Language Processing, pages 17–24
Manchester, August 2008

somehow. A natural first hypothesis is that the grounding words ought to be more concrete, referring to things that are closer to our overt sensorimotor experience, and learned earlier, but that remains to be tested (Clark, 2003). Apart from the question of the boundary conditions of grounding, however, there are basic questions to be asked about the structure of word meanings in dictionary definition space.

In the path from a word, to the definition of that word, to the definition of the words in the definition of that word, and so on, through what sort of a structure are we navigating (Ravasz & Barabasi, 2003; Steyvers & Tenenbaum, 2005)? Meaning is compositional: A definition is composed of words, combined according to syntactic rules to form a proposition (with a truth value: true or false). For example, the word to be defined w (the "definiendum") might mean $w_1 \,\&\, w_2 \,\&\, \ldots \,\&\, w_n$, where the w_i are other words (the "definientes") in its definition. Rarely does that proposition provide the full necessary and sufficient conditions for identifying the referent of the word, w, but the approximation must at least be close enough to allow most people, armed with the definition, to understand and use the defined word most of the time, possibly after looking up a few of its definientes d_w, but without having to cycle through the entire dictionary, and without falling into circularity or infinite regress.

If enough of the definientes are grounded, then there is no problem of infinite regress. But we can still ask the question: What is the size of the grounding vocabulary? and what words does it contain? What is the length and shape of the path that would be taken in a recursive definitional search, from a word, to its definition, to the definition of the words in its definition, and so on? Would it eventually cycle through the entire dictionary? Or would there be disjoint subsets?

This paper raises more questions than it answers, but it develops the formal groundwork for a new means of finding the answers to questions about how word meaning is *explicitly* represented in real dictionaries — and perhaps also about how it is *implicitly* represented in the "mental lexicon" that each of us has in our brain (Hauk et al., 2008).

The remainder of this paper is organized as follows: In Section 2, we introduce the graph-theoretical definitions and notations used for formulating the symbol grounding problem in Section 3. Sections 4 and 5 deal with the implication of this approach in cognitive sciences and show in what ways grounding kernels may be useful.

2 Definitions and Notations

In this section, we give mathematical definitions for the dictionary-related terminology, relate them to natural language dictionaries and supply the pertinent graph theoretical definitions. Additional details are given to ensure mutual comprehensibility to specialists in the three disciplines involved (mathematics, linguistics and psychology). Complete introductions to graph theory and discrete mathematics are provided in (Bondy & Murty, 1978; Rosen, 2007).

2.1 Relations and Functions

Let A be any set. A *binary relation on* A is any subset R of $A \times A$. We write xRy if $(x, y) \in R$. The relation R is said to be (1) *reflexive* if for all $x \in A$, we have xRx, (2) *symmetric* if for all $x, y \in A$ such that xRy, we have yRx and (3) *transitive* if for all $x, y, z \in A$ such that xRy and yRz, we have xRz. The relation R is an *equivalence relation* if it is reflexive, symmetric and transitive. For any $x \in A$, the *equivalence class of* x, designated by $[x]$, is given by $[x] = \{y \in A \mid xRy\}$. It is easy to show that $[x] = [y]$ if and only if xRy and that the set of all equivalence classes forms a partition of A.

Let A be any set, $f : A \to A$ a function and k a positive integer. We designate by f^k the function $f \circ f \circ \ldots \circ f$ (k times), where $\circ$ denotes the *composition of functions*.

2.2 Dictionaries

At its most basic level, a dictionary is a set of associated pairs: a *word* and its *definition*, along with some disambiguating parameters. The *word*[1] to be defined, w, is called the *definiendum* (plural: *definienda*) while the finite nonempty set of words that defines w, d_w, is called the set of *definientes* of w (singular: *definiens*).

Each dictionary entry accordingly consists of a definiendum w followed by its set of definientes d_w. A *dictionary* D then consists of a finite set of pairs (w, d_w) where w is a word and $d_w = \{w_1, w_2, \ldots, w_n\}$, where $n \geq 1$, is its definition, satisfying the property that for all $(w, d_w) \in D$ and for all $d \in d_w$, there exists $(w', d_{w'}) \in D$ such that $d = w'$. A pair (w, d_w) is called an *entry* of D. In other words, a dictionary is a finite set of words, each of which is defined, and each of its defining words is likewise defined somewhere in the dictionary.

2.3 Graphs

A *directed graph* is a pair $G = (V, E)$ such that V is a finite set of *vertices* and $E \subseteq V \times V$ is a finite set of *arcs*. Given $V' \subseteq V$, the *subgraph induced by* V', designated by $G[V']$, is the graph $G[V'] = (V', E')$ where $E' = E \cap (V' \times V')$. For any $v \in V$, $N^-(v)$ and $N^+(v)$ designate, respectively, the set of incoming and outgoing neighbors of v, i.e.

$$\begin{aligned} N^-(v) &= \{u \in V \mid (u, v) \in E\} \\ N^+(v) &= \{u \in V \mid (v, u) \in E\}. \end{aligned}$$

We write $\deg^-(v) = |N^-(v)|$ and $\deg^+(v) = |N^+(v)|$, respectively. A *path* of G is a sequence

[1] In the context of this mathematical analysis, we will use "word" to mean a finite string of uninterrupted letters having some associated meaning.

$(v_1, v_2, \ldots, v_n)$, where n is a positive integer, $v_i \in V$ for $i = 1, 2, \ldots, n$ and $(v_i, v_{i+1}) \in E$, for $i = 1, 2, \ldots, n-1$. A *uv-path* is a path starting with u and ending with v. Finally, we say that a uv-path is a *cycle* if $u = v$.

Given a directed graph $G = (V, E)$ and $u, v \in V$, we write $u \to v$ if there exists a uv-path in G. We define a relation $\sim$ as

$$u \sim v \Leftrightarrow u \to v \text{ and } v \to u.$$

It is an easy exercise to show that $\sim$ is an equivalence relation. The equivalence classes of V with respect to $\sim$ are called the *strongly connected components* of G. In other words, in a directed graph, it might be possible to go directly from point A to point B, without being able to get back from point B to point A (as in a city with only one-way streets). Strongly connected components, however, are subgraphs in which whenever it is possible to go from point A to point B, it is also possible to come back from point B to point A (the way back may be different).

There is a very natural way of representing definitional relations using graph theory, thus providing a formal tool for analyzing grounding properties of dictionaries: words can be represented as vertices, with arcs representing definitional relations, i.e. there is an arc (u, v) between two words u and v if the word u appears in the definition of the word v. More formally, for every dictionary D, its *associated graph* $G = (V, E)$ is given by

$$\begin{aligned} V &= \{w \mid \exists d_w \text{ such that } (w, d_w) \in D\}, \\ E &= \{(v, w) \mid \exists d_w \text{ such that } (w, d_w) \in D \text{ and } \\ &\quad v \in d_w\}. \end{aligned}$$

Note that every vertex v of G satisfies $\deg^-_G(v) > 0$, but it is possible to have $\deg^+_G(v) = 0$. In other words, whereas every word has a definition, some words are not used in any definition.

Example 1. *Let D be the dictionary whose definitions are given in Table 1. Note that every word appearing in some definition is likewise defined in D (this is one of the criteria for D to be a dictionary). The associated graph G of D is represented in Figure 1. Note that (not, good, eatable, fruit) is a path of G while (good, bad, good) is a cycle (as well as a path) of G.*

3 A Graph-Theoretical Formulation of the Problem

We are now ready to formulate the symbol grounding problem from a mathematical point of view.

3.1 Reachable and Grounding Sets

Given a dictionary D of n words and a person x who knows m out of these n words, assume that the only way x can learn new words is by consulting the dictionary definitions. Can all n words be learned by x through dictionary look-up alone? If not, then exactly what subset of words can be learned by x through dictionary look-up alone?

Word	Definition	Word	Definition
apple	red fruit	bad	not good
banana	yellow fruit	color	dark or light
dark	not light	eatable	good
fruit	eatable thing	good	not bad
light	not dark	not	not
or	or	red	dark color
thing	thing	tomato	red fruit
yellow	light color		

Table 1: Definitions of the dictionary D

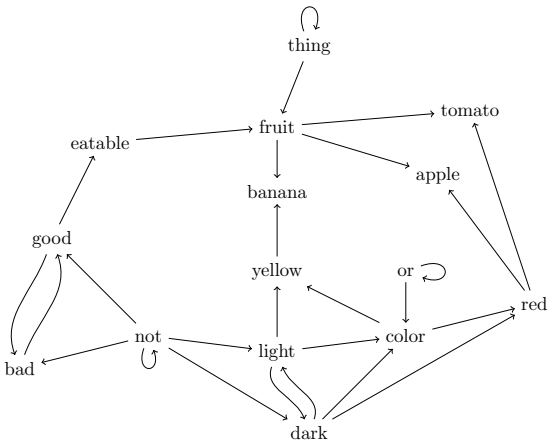

Figure 1: Graph representation of the dictionary D.

For this purpose, let $G = (V, E)$ be a directed graph and consider the following application, where 2^V denotes the collection of all subsets of V:

$$\begin{aligned} R_G : \quad 2^V &\longmapsto 2^V \\ U &\longmapsto U \cup \{v \in V \mid N^-(v) \subseteq U\}. \end{aligned}$$

When the context is clear, we omit the subscript G. Also we let R^k denote the k^{th} power of R. We say that $v \in V$ is *k-reachable from U* if $v \in R^k(U)$ and k is a nonnegative integer. It is easy to show that there exists an integer k such that $R^\ell(U) = R^k(U)$, for every integer $\ell > k$. More precisely, we have the following definitions:

Definition 2. *Let $G = (V, E)$ be a directed graph, U a subset of V, and k an integer such that $R^\ell(U) = R^k(U)$ for all $\ell > k$. The set $R^k(U)$ is called the* reachable set from U *and is denoted by $R^*(U)$. Moreover, if $R^*(U) = V$, then we say that U is a* grounding set *of G.*

We say that G is *p-groundable* if there exists $U \subseteq V$ such that $|U| = p$ and U is a grounding set of G. The *grounding number* of a graph G is the smallest integer p such that G is p-groundable.

Reachable sets can be computed very simply using a breadth-first-search type algorithm, as shown by Algo-

rithm 1.

Algorithm 1 Computing reachable sets

```
1: function REACHABLESET(G, U)
2:     R ← U
3:     repeat
4:         S ← {v ∈ V | N_G^-(v) ⊆ R} − R
5:         R ← R ∪ S
6:     until S = ∅
7:     return R
8: end function
```

We now present some examples of reachable sets and grounding sets.

Example 3. *Consider the dictionary D and the graph G of Example 1. Let $U = \{bad, light, not, thing\}$. Note that*

$$\begin{aligned} R^0(U) &= U \\ R^1(U) &= U \cup \{dark, good\}, \\ R^2(U) &= R^1(U) \cup \{eatable\} \\ R^3(U) &= R^2(U) \cup \{fruit\} \\ R^4(U) &= R^3(U) \end{aligned}$$

so that $R^*(U) = \{bad, dark, eatable, fruit, good, light, not, thing\}$ (see Figure 2). In particular, this means that the word "eatable" is 2-reachable (but not 1-reachable) from U and all words in U are 0-reachable from U. Moreover, we observe that U is not a grounding set of G ("color", for example, is unreachable). On the other hand, the set $U' = U \cup \{or\}$ is a grounding set of G, so that G is 5-groundable.

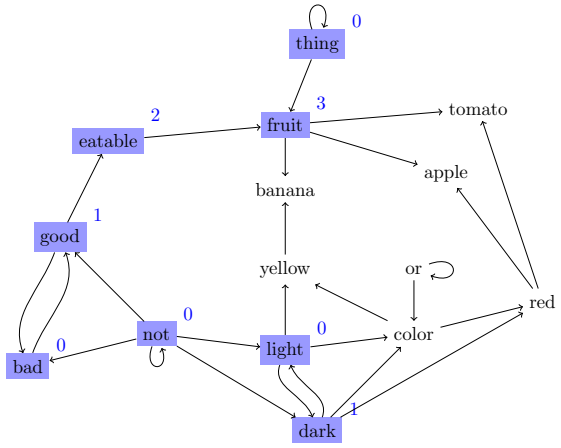

Figure 2: The set $R^*(U)$ (the words in squares) obtained from U

3.2 The Minimum Grounding Set Problem

Given a dictionary and its associated graph G, we are interested in finding minimum grounding sets of G. (Note that in general, there is more than one grounding set of minimum cardinality.) This is related to a natural decision problem: we designate by k-GS the problem of deciding whether G is k-groundable. We show that k-GS is closely related to the problem of finding minimum feedback vertex sets. First, we recall the definition of a feedback vertex set.

Definition 4. *Let $G = (V, E)$ be a directed graph and U a subset of V. We say that U is a* feedback vertex set *of G if for every cycle C of G, we have $U \cap C \neq \emptyset$. In other words, U covers every cycle of G.*

The *minimum feedback vertex set problem* is the problem of finding a feedback vertex set of G of minimum cardinality. To show that feedback vertex sets and grounding sets are the same, we begin by stating two simple lemmas.

Lemma 5. *Let $G = (V, E)$ be a directed graph, C a cycle of G and $U \subseteq V$ a grounding set of G. Then $U \cap C \neq \emptyset$.*

Proof. By contradiction, assume that $U \cap C = \emptyset$ and, for all $v \in C$, there exists an integer k such that v belongs to $R^k(U)$. Let ℓ be the smallest index in the set $\{k \mid \exists u \in C$ such that $u \in R^k(U)\}$. Let u be a vertex in $C \cap R^\ell(U)$ and w the predecessor of u in C. Since $U \cap C = \emptyset$, k must be greater than 0 and w a member of $R^{\ell-1}(U)$, contradicting the minimality of ℓ. □

Lemma 6. *Every directed acyclic graph G is 0-groundable.*

Proof. We prove the statement by induction on $|V|$.

BASIS. If $|V| = 1$, then $|E| = 0$, so that the only vertex v of G satisfies $N_G^-(v) = \emptyset$. Hence $R(\emptyset) = V$.

INDUCTION. Let v be a vertex such that $\deg^+(v) = 0$. Such a vertex exists since G is acyclic. Moreover, let G' be the (acyclic) graph obtained from G by removing vertex v and all its incident arcs. By the induction hypothesis, there exists an integer k such that $R^k_{G'}(\emptyset) = V - \{v\}$. Therefore, $V - \{v\} \subseteq R^k_G(\emptyset)$ so that $R^{k+1}_G(\emptyset) = V$. □

The next theorem follows easily from Lemmas 5 and 6.

Theorem 7. *Let $G = (V, E)$ be a directed graph and $U \subseteq V$. Then U is a grounding set of G if and only if U is a feedback vertex set of G.*

Proof. ($\Rightarrow$) Let C be a cycle of G. By Lemma 5, $U \cap C \neq \emptyset$, so that U is a minimum feedback vertex set of G. ($\Leftarrow$) Let G' be the graph obtained from G by removing U. Then G' is acyclic and $\emptyset$ is a grounding set of G'. Therefore, $U \cup \emptyset = U$ is a grounding set of G. □

Corollary 8. *k-GS is NP-complete.*

Proof. Denote by k-FVS the problem of deciding whether a directed graph G admits a feedback vertex set of cardinality at most k. This problem is known to be NP-complete and has been widely studied (Karp, 1972; Garey & Johnson, 1979). It follows directly from Theorem 7 that k-GS is NP-complete as well since the problems are equivalent. □

The fact that problems k-GS and k-FVS are equivalent is not very surprising. Indeed, roughly speaking, the minimum grounding problem consists of finding a minimum set large enough to enable the reader to learn (reach) all the words of the dictionary. On the other hand, the minimum feedback vertex set problem consists of finding a minimum set large enough to break the circularity of the definitions in the dictionary. Hence, the problems are the same, even if they are stated differently.

Although the problem is NP-complete in general, we show that there is a simple way of reducing the complexity of the problem by considering the strongly connected components.

3.3 Decomposing the Problem

Let $G = (V, E)$ be a directed graph and G_1, G_2, ..., G_m the subgraphs induced by its strongly connected components, where $m \geq 1$. In particular, there are no cycles of G containing vertices in different strongly connected components. Since the minimum grounding set problem is equivalent to the minimum feedback vertex set problem, this means that when seeking a minimum grounding set of G, we can restrict ourselves to seeking minimum grounding sets of G_i, for $i = 1, 2, \ldots, m$. More precisely, we have the following proposition.

Proposition 9. *Let $G = (V, E)$ be a directed graph with m strongly connected components, with $m \geq 1$, and let $G_i = (V_i, E_i)$ be the subgraph induced by its i-th strongly connected component, where $1 \leq i \leq m$. Moreover, let U_i be a minimum grounding set of G_i, for $i = 1, 2, \ldots, m$. Then $U = \bigcup_{i=1}^{m} U_i$ is a minimum grounding set of G.*

Proof. First, we show that U is a grounding set of G. Let C be a cycle of G. Then C is completely contained in some strongly connected component of G, say G_j, where $1 \leq j \leq m$. But $U_j \subseteq U$ is a grounding set of G_j, therefore $U_j \cap C \neq \emptyset$ so that $U \cap C \neq \emptyset$. It remains to show that U is a minimum grounding set of G. By contradiction, assume that there exists a grounding set U' of G, with $|U'| < |U|$ and let $U'_i = U' \cap V_i$. Then there exists an index j, with $1 \leq j \leq m$, such that $|U'_j| < |U_j|$, contradicting the minimality of $|U_j|$. □

Note that this proposition may be very useful for graphs having many small strongly connected components. Indeed, by using Tarjan's Algorithm (Tarjan, 1972), the strongly connected components can be computed in linear time. We illustrate this reduction by an example.

Example 10. *Consider again the dictionary D and the graph G of Example 1. The strongly connected components of G are encircled in Figure 3 and minimum grounding sets (represented by words in squares) for each of them are easily found. Thus the grounding number of G is 5.*

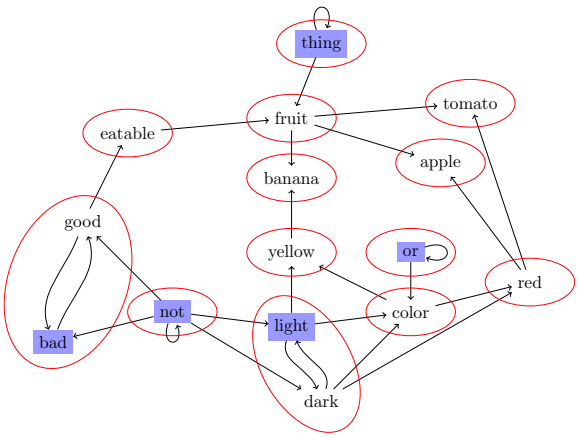

Figure 3: The strongly connected components and a minimum grounding set of G

3.4 The Grounding Kernel

In Example 10, we have seen that there exist some strongly connected components consisting of only one vertex without any loop. In particular, there exist vertices with no successor, i.e. vertices v such that $N_G^+(v) = 0$. For instance, this is the case of the words "apple", "banana" and "tomato", which are not used in any definition in the dictionary. Removing these three words, we notice that "fruit", "red" and "yellow" are in the same situation and they can be removed as well. Pursuing the same idea, we can now remove the words "color" and "eatable". At this point, we cannot remove any further words. The set of remaining words is called the *grounding kernel* of the graph G. More formally, we have the following definition..

Definition 11. *Let D be a dictionary, $G = (V, E)$ its associated graph and $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, ..., $G_m = (V_m, E_m)$ the subgraphs induced by the strongly connected components of G, where $m \geq 1$. Let V' be the set of vertices u such that $\{u\}$ is a strongly connected component without any loop (i.e., (u, u) is not an arc of G). For any u, let $N^*(u)$ denote the set of vertices v such that G contains a uv-path. Then the* grounding kernel *of G, denoted by K_G, is the set $V - \{u \mid u \in V' \text{ and } N^*(u) \subseteq V'\}$.*

Clearly, every dictionary D admits a grounding kernel, as shown by Algorithm 2. Moreover, the grounding kernel is a grounding set of its associated graph G and every minimum grounding set of G is a subset of the grounding kernel. Therefore, in studying the symbol grounding problem in dictionaries, we can restrict

Algorithm 2 Computing the grounding kernel

```
function GroundingKernel(G)
    G' ← G
    repeat
        Let W be the set of vertices of G'
        U ← {v ∈ W | N+_G'(v) = ∅}
        G' ← G'[W − U]
    until U = ∅
    return G'
end function
```

ourselves to the grounding kernel of the graph G corresponding to D. This phenomenon is interesting because every dictionary contains many words that can be recursively removed without compromising the understanding of the other definitions. Formally, this property relates to the *level* of a word: we will say of a word w that it is of *level* k if it is k-reachable from K_G but not ℓ-reachable from K_G, for any $\ell < k$. In particular, level 0 indicates that the word is part of the grounding kernel. A similar concept has been studied in (Changizi, 2008).

Example 12. *Continuing Example 10 and from what we have seen so far, it follows that the grounding kernel of G is given by*

$$K_G = \{\textit{bad, dark, good, light, not, or, thing}\}.$$

Level 1 *words are "color" and "eatable", level* 2 *words are "fruit", "red" and "yellow", and level* 3 *words are "apple", "banana" and "tomato".*

4 Grounding Sets and the Mental Lexicon

In Section 3, we introduced all the necessary terminology to study the symbol grounding problem using graph theory and digital dictionaries. In this section, we explain how this model can be useful and on what assumptions it is based.

A dictionary is a formal symbol system. The preceding section showed how formal methods can be applied to this system in order to extract formal features. In cognitive science, this is the basis of *computationalism* (or cognitivism or "disembodied cognition" (Pylyshyn, 1984)), according to which cognition, too, is a formal symbol system – one that can be studied and explained independently of the hardware (or, insofar as it concerns humans, the wetware) on which it is implemented. However, pure computationalism is vulnerable to the problem of the grounding of symbols too (Harnad, 1990). Some of this can be remedied by the competing paradigm of embodied cognition (Barsalou, 2008; Glenberg & Robertson, 2002; Steels, 2007), which draws on dynamical (noncomputational) systems theory to ground cognition in sensorimotor experience. Although computationalism and symbol grounding provide the background context for our investigations and findings, the present paper does not favor any particular theory of mental representation of meaning.

A dictionary is a symbol system that relates words to words in such a way that the meanings of the definienda are conveyed via the definientes. The user is intended to arrive at an understanding of an unknown word through an understanding of its definition. What was formally demonstrated in Section 3 agrees with common sense: although one can learn new word meanings from a dictionary, the entire dictionary cannot be learned in this way because of circular references in the definitions (*cycles*, in graph theoretic terminology). Information – *nonverbal* information – must come from outside the system to ground at least some of its symbols by some means other than just formal definition (Cangelosi & Harnad, 2001). For humans, the two options are learned sensorimotor grounding and innate grounding. (Although the latter is no doubt important, our current focus is more on the former.)

The need for information from outside the dictionary is formalized in Section 3. Apart from confirming the need for such external grounding, we take a symmetric stance: In natural language, some word meanings — especially highly abstract ones, such as those of mathematical or philosophical terms — are not or cannot be acquired through direct sensorimotor grounding. They are acquired through the *composition* of previously known words. The meaning of some of those words, or of the words in their respective definitions, must in turn have been grounded through direct sensorimotor experience.

To state this in another way: Meaning is not just formal definitions all the way down; nor is it just sensorimotor experience all the way up. The two extreme poles of that continuum are *sensorimotor induction* at one pole (trial and error experience with corrective feedback; observation, pointing, gestures, imitation, etc.), and *symbolic instruction* (definitions, descriptions, explanation, verbal examples etc.) at the other pole. Being able to identify from their lexicological structure which words were acquired one way or the other would provide us with important clues about the cognitive processes underlying language and the mental representation of meaning.

To compare the word meanings acquired via sensorimotor induction with word meanings acquired via symbolic instruction (definitions), we first need access to the encoding of that knowledge. In this component of our research, our hypothesis is that the representational structure of word meanings in dictionaries shares some commonalities with the representational structure of word meanings in the human brain (Hauk et al., 2008). We are thus trying to extract from dictionaries the grounding kernel (and eventually a minimum grounding set, which in general is a proper subset of this kernel), from which the rest of the dictionary can be reached through definitions alone. We hypothesize that this kernel, identified through formal structural analy-

sis, will exhibit properties that are also reflected in the mental lexicon. In parallel ongoing studies, we are finding that the words in the grounding kernel are indeed (1) more frequent in oral and written usage, (2) more concrete, (3) more readily imageable, and (4) learned earlier or at a younger age. We also expect they will be (5) more universal (across dictionaries, languages and cultures) (Chicoisne et al., 2008).

5 Grounding Kernels in Natural Language Dictionaries

In earlier research (Clark, 2003), we have been analyzing two special dictionaries: the Longman's Dictionary of Contemporary English (LDOCE) (Procter, 1978) and the Cambridge International Dictionary of English (CIDE) (Procter, 1995). Both are officially described as being based upon a *defining vocabulary*: a set of 2000 words which are purportedly the only words used in all the definitions of the dictionary, including the definitions of the defining vocabulary itself. A closer analysis of this defining vocabulary, however, has revealed that it is not always faithful to these constraints: A significant number of words used in the definitions turn out not to be in the defining vocabulary. Hence it became evident that we would ourselves have to generate a grounding kernel (roughly equivalent to the defining vocabulary) from these dictionaries.

The method presented in this paper makes it possible, given the graph structure of a dictionary, to extract a grounding kernel therefrom. Extracting this structure in turn confronts us with two further problems: *morphology* and *polysemy*. Neither of these problems has a definite algorithmic solution. Morphology can be treated through stemming and associated look-up lists for the simplest cases (*i.e.*, was → to be, and children → child), but more elaborate or complicated cases would require syntactic analysis or, ultimately, human evaluation. Polysemy is usually treated through statistical analysis of the word context (as in Latent Semantic Analysis) (Kintsch, 2007) or human evaluation. Indeed, a good deal of background knowledge is necessary to analyse an entry such as: "*dominant*: the fifth note of a musical scale of eight notes" (the LDOCE notes 16 different meanings of *scale* and 4 for *dominant*, and in our example, none of these words are used with their most frequent meaning).

Correct disambiguation of a dictionary is time-consuming work, as the most effective way to do it for now is through consensus among human evaluators. Fortunately, a fully disambiguated version of the WordNet database (Fellbaum, 1998; Fellbaum, 2005) has just become available. We expect the grounding kernel of WordNet to be of greater interest than the defining vocabulary of either CIDE or LDOCE (or what we extract from them and disambiguate automatically, and imperfectly) for our analysis.

6 Future Work

The main purpose of this paper was to introduce a formal approach to the symbol grounding problem based on the computational analysis of digital dictionaries. Ongoing and future work includes the following:

The minimum grounding set problem. We have seen that the problem of finding a minimum grounding set is NP-complete for general graphs. However, graphs associated with dictionaries have a very specific structure. We intend to describe a class of graphs including those specific graphs and to try to design a polynomial-time algorithm to solve the problem. Another approach is to design approximation algorithms, yielding a solution close to the optimal solution, with some known guarantee.

Grounding sets satisfying particular constraints. Let D be a dictionary, $G = (V, E)$ its associated graph, and $U \subseteq V$ any subset of vertices satisfying a given property P. We can use Algorithm 1 to test whether or not U is a grounding set. In particular, it would be interesting to test different sets U satisfying different cognitive constraints.

Relaxing the grounding conditions. In this paper we imposed strong conditions on the learning of new words: One must know all the words of the definition fully in order to learn a new word from them. This is not realistic, because we all know one can often understand a definition without knowing every single word in it. Hence one way to relax these conditions would be to modify the learning rule so that one need only understand at least $r\%$ of the definition, where r is some number between 0 and 100. Another variation would be to assign weights to words to take into account their morphosyntactic and semantic properties (rather than just treating them as an unordered list, as in the present analysis). Finally, we could consider "quasi-grounding sets", whose associated reachable set consists of $r\%$ of the whole dictionary.

Disambiguation of definitional relations. Analyzing real dictionaries raises, in its full generality, the problem of word and text disambiguation in free text; this is a very difficult problem. For example, if the word "make" appears in a definition, we do not know which of its many senses is intended — nor even what its grammatical category is. To our knowledge, the only available dictionary that endeavors to provide fully disambiguated definitions is the just-released version of WordNet. On the other hand, dictionary definitions have a very specific grammatical structure, presumably simpler and more limited than the general case of free text. It might hence be feasible to develop automatic disambiguation algorithms specifically dedicated to the special case of dictionary definitions.

Concluding Remark: Definition can reach the sense (sometimes), but only the senses can reach the referent.

Research funded by Canada Research Chair in Cognitive Sciences, SSHRC (S. Harnad)and NSERC (S. Harnad & O. Marcotte)

References

Barsalou, L. (2008) *Grounded Cognition*. Annual Review of Psychology (in press).

Bondy, J.A. & U.S.R. Murty. (1978) *Graph theory with applications*. Macmillan, New York.

Cangelosi, A. & Harnad, S. (2001) *The Adaptive Advantage of Symbolic Theft Over Sensorimotor Toil: Grounding Language in Perceptual Categories*. Evol. of Communication 4(1) 117-142.

Changizi, M.A. (2008) *Economically organized hierarchies in WordNet and the Oxford English Dictionary*. Cognitive Systems Research (in press).

Chicoisne G., A. Blondin-Massé, O. Picard, S. Harnad (2008) *Grounding Abstract Word Definitions In Prior Concrete Experience*. 6th Int. Conf. on the Mental Lexicon, Banff, Alberta.

Clark G. (2003) *Recursion Through Dictionary Definition Space: Concrete Versus Abstract Words*. (U. Southampton Tech Report).

Fellbaum, C. (1998) *WordNet: An electronic lexical database*. Cambridge: MIT Press.

Fellbaum, C. (2005) *Theories of human semantic representation of the mental lexicon*. In: Cruse, D. A. (Ed.), Handbook of Linguistics and Communication Science, Berlin, Germany: Walter de Gruyter, 1749-1758.

Frege G. (1948) *Sense and Reference*. The Philosophical Review 57 (3) 209-230.

Garey, M.R. & D.S. Johnson (1979) *Computers and Intractability: A Guide to the Theory of NP-completeness*. W.H. Freeman, New York.

Glenberg A.M. & D.A. Robertson (2002) *Symbol Grounding and Meaning: A Comparison of High-Dimensional and Embodied Theories of Meaning*. Journal of Memory and Language 43 (3) 379-401.

Harnad, S. (1990) *The Symbol Grounding Problem*. Physica D 42:335-346.

Harnad, S. (2003) *Symbol-Grounding Problem*. Encylopedia of Cognitive Science. Nature Publishing Group. Macmillan.

Harnad, S. (2005) *To Cognize is to Categorize: Cognition is Categorization*. In Lefebvre, C. and Cohen, H. (Eds.), Handbook of Categorization. Elsevier.

Hauk, O., M.H. Davis, F. Kherif, F. Pulvermüller. (2008) *Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging*. European Journal of Neuroscience 27 (7) 1856-1866.

Karp, R.M. (1972) *Reducibility among combinatorial problems*. In: R.E. Miller, J.W. Thatcher (Eds.), Complexity of Computer Computations, Plenum Press, New York, 1972, pp. 85-103.

Kintsch, W. (2007) *Meaning in Context*. In T.K. Landauer, D.S. McNamara, S. Dennis & W. Kintsch (Eds.), Handbook of Latent Semantic Analysis. Erlbaum.

Procter, P. (1978) *Longman Dictionary of Contemporary English*. Longman Group Ltd., Essex, UK.

Procter, P. (1995) *Cambridge International Dictionary of English (CIDE)*. Cambridge University Press.

Pylyshyn, Z. W. (1984) *Computation and Cognition: Towards a Foundation for Cognitive Science*. Cambridge: MIT Press.

Ravasz, E. & Barabasi, A. L. (2003) *Hierarchical organization in complex networks*. Physical Review E 67, 026112.

Rosen, K.H. (2007) *Discrete mathematics and its applications*, 6th ed. McGraw-Hill.

Steels, L. (2007) *The symbol grounding problem is solved, so what's next?* In De Vega, M. and G. Glenberg and A. Graesser (Eds.), Symbols, embodiment and meaning. Academic Press, North Haven.

Steyvers, M. & Tenenbaum J.B. (2005) *The large-scale structure of semantic networks: statistical analyses and a model of semantic growth*. Cognitive Science, 29(1) 41-78.

Tarjan, R. (1972) *Depth-first search and linear graph algorithms*. SIAM Journal on Computing. 1 (2) 146-160.