Domains in Cognitive Metaphor Theory and Metaphor Processing

KJ Nabeshima
Kansai University
Osaka, JAPAN
spiralcricket@gmail.com

ABSTRACT: Computational approaches to metaphor processing and identification are gaining popularity. In this paper, I will discuss the nature of a domain from a Cognitive Linguistic perspective, the kinds of domains used in sample computational approaches, and how they arrive at them.

Keywords: Metaphor, Domain

Received: 4 March 2017, Revised 2 May 2017, Accepted 24 May 2017

© 2017 DLINE. All Rights Reserved

1. Introduction

In this paper, a domain or frame is discussed from a Cognitive Linguistic perspective and domains in computational metaphor approaches are reviewed. In Section 2, Cognitive Metaphor Theory (CMT) and its treatment of a domain are summarized. In Section 3, examples of metaphor domains are listed and domain categories are discussed. Section 4 discusses three major origins of domains from psychology, sociology and brain science. In Section 5, several existing papers on computational metaphor processing are selected and discussed, with special attention to how they treat domains. Section 6 is the summary and conclusion. A domain is a crucial concept in CMT and I believe paying close attention to this concept will help advancement of computational research on conceptual metaphors.

2. Domains in CMT

Since Lakoff and Johnson's first book, the Cognitive Metaphor program has been flourishing (Lakoff and Johnson, 1980, 1999; Grady, 1997; Pragglejaz Group 2007). Meanwhile, its theoretical backbone may not necessarily be understood to its full details. Therefore, it is useful to recapture where it comes from and what theoretical structure it has. According to Lakoff and Johnson, metaphor is a mapping between two domains (Lakoff and Johnson 1999). Domains and several other technical terms in Cognitive Linguistics (ICMs, frames, cognitive domains) are known to capture the same nature of knowledge structure: Words do not exist alone; They come as surface manifestations of interconnected knowledge of words, context, and inferences (Clausner and Croft 1999). A domain is a background knowledge structure which includes related words, concepts and inferences. Two domains
involved in metaphor are asymmetric and called the source domain and the target domain: The source domain is a concrete domain which is used to discuss the current main topic, and the target domain is an abstract domain which is currently discussed.

In this section, how Cognitive Metaphor Theory sees the nature of a domain is discussed, along with the definition of a metaphor as a mapping between the source domain and the target domain.

3. Examples of domains

<table>
<thead>
<tr>
<th>T-Category</th>
<th>Target</th>
<th>Source</th>
<th>S-Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>Aspects</td>
<td>Components</td>
<td>SCHEMATIC</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Effects</td>
<td>Scars</td>
<td>BODY</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Functional</td>
<td>Healthy</td>
<td>BODY</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Repeating</td>
<td>Circle</td>
<td>SCHEMATIC</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Imperfect</td>
<td>Irregular</td>
<td>SCHEMATIC</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Action</td>
<td>Motion</td>
<td>SCHEMATIC</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Manner of Action</td>
<td>Manner of Motion</td>
<td>SCHEMATIC</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Properties</td>
<td>Possessions</td>
<td>ONTOLOGICAL</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Good</td>
<td>Forward</td>
<td>ORIENTATIONAL</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Existence</td>
<td>Up</td>
<td>ORIENTATIONAL</td>
</tr>
<tr>
<td>CULTURAL</td>
<td>Adversity</td>
<td>Bad Weather</td>
<td>ONTOLOGICAL</td>
</tr>
<tr>
<td>CULTURAL</td>
<td>Research</td>
<td>Exploration</td>
<td>CULTURAL</td>
</tr>
<tr>
<td>CULTURAL</td>
<td>Arguments</td>
<td>Paths</td>
<td>SCHEMATIC</td>
</tr>
<tr>
<td>CULTURAL</td>
<td>Investment</td>
<td>Pouring</td>
<td>ONTOLOGICAL</td>
</tr>
<tr>
<td>EMOTIONAL</td>
<td>Emotion</td>
<td>Liquids</td>
<td>ONTOLOGICAL</td>
</tr>
<tr>
<td>EMOTIONAL</td>
<td>Laughter</td>
<td>Substance</td>
<td>ONTOLOGICAL</td>
</tr>
<tr>
<td>EMOTIONAL</td>
<td>Dislike</td>
<td>Cold</td>
<td>PERCEPTUAL</td>
</tr>
<tr>
<td>EMOTIONAL</td>
<td>Lust</td>
<td>Heat</td>
<td>PERCEPTUAL</td>
</tr>
<tr>
<td>EMOTIONAL</td>
<td>Desire</td>
<td>Hunger</td>
<td>PHYSIOLOGICAL</td>
</tr>
<tr>
<td>MENTAL</td>
<td>Theories</td>
<td>Clothes</td>
<td>CULTURAL</td>
</tr>
<tr>
<td>MENTAL</td>
<td>Premises</td>
<td>Source</td>
<td>SCHEMATIC</td>
</tr>
<tr>
<td>MENTAL</td>
<td>Belief</td>
<td>Held Object</td>
<td>ONTOLOGICAL</td>
</tr>
<tr>
<td>MENTAL</td>
<td>Remembering</td>
<td>Recollecting</td>
<td>ONTOLOGICAL</td>
</tr>
<tr>
<td>SOCIAL</td>
<td>Communication</td>
<td>Showing</td>
<td>BODY</td>
</tr>
<tr>
<td>SOCIAL</td>
<td>Helping</td>
<td>Raising</td>
<td>ORIENTATIONAL</td>
</tr>
</tbody>
</table>

Table 1. Metaphors domains and their categories

Table 1 indicates a random set of 25 metaphors extracted from Master Metaphor List (Lakoff et al. 1991). You can see a variety of domains for both the target and the source.

In addition to the names of the source domains and the target domains, I added domain categories (S-category for source domain categories and T-category for target domain categories). Five target domain categories are used: ABSTRACT, CULTURAL, EMOTIONAL, MENTAL, and SOCIAL. Eight source domain categories are used: ABSTRACT, BODY, CULTURAL, ONTOLOGICAL, ORIENTATIONAL, PERCEPTUAL, PHYSIOLOGICAL, and SCHEMATIC.
I will explain each category briefly. I believe CULTURAL, EMOTIONAL, MENTAL and SOCIAL of the target domain categories go without much explanation. CULTURAL includes activities and artefacts particular to human species such as Research, Arguments, Investment, Exploration, and Clothes. EMOTIONAL includes concepts such as Emotion, Dislike, Desire, Lust and activities and phenomena related to emotion such as Laughter.

MENTAL includes mental activities particular to human species such as Belief and Remembering. I also put Theories here because it pertains to our ability to think. Premises is also here because it is a part of theory making. I categorized interpersonal activities such as Communication and Helping under the SOCIAL domain.

ABSTRACT includes domains such as Aspects, Effects, Functional, Repeating, Action, Properties and Existence. They are abstract in that they are intangible and you can’t form a consistent mental images. Instances of EFFECT may include a splash of water when you throw a stone, or a smile of a person when you say nice things to her. They can be very physical and concrete, but grasping all these instances as EFFECT is a generalization and raises its categorical status much higher than a tangible and imagistic level. Imperfect, Good, Adversity are included in ABSTRACT because they bring strong evaluative component and they are not tangible or imageable.

ONTLOGICAL includes Object and Substance. Weather is also under this category because it is a type of Natural phenomena, which has its own domain inferences and ontological status. ONTOLOGICAL should also include Living, Animate and Human. ORIENTATIONAL includes UP-DOWN, FRONT-BACK and LEFT-RIGHT. These are building-blocks of constructing the space. SCHEMATIC is similar to ABSTRACT in that it is idealized, schematic and topological, but different in that it originates in images and retain imagistic characteristics. I used the term SCHEMATIC after Image-Schemas in Cognitive Linguistics. Categories such as CIRCLE, IRREGULAR, PATHS, MOTION, SOURCE and PARTS should all be included in so-called Image-Schemas in Cognitive Linguistics. Finally, I set up a category Body to include miscellaneous things related to body such as Showing, Healthy and Scar. They are disparate but it is true there’s a tendency that things related to human body form the source domain.

Interesting thing to note is that CULTURAL shows up both in the target domain and the source domain. This means that it is not possible to strictly differentiate the source and the target by domain categories.

4. Three major origins of domains

In Section 3, we discussed several domain categories. Here are three major origins for metaphor domain categories: theories, social frames and domains in the brain.

4.1 Theories in developmental psychology

The first one is called “theories” in the developmental psychology (e.g. Spelke, 1991) which is a candidate for motivation behind the category ONTOLOGICAL. These include OBJECT, SUBSTANCE, ANIMATE, LIVING, HUMAN, NATURAL PHENOMENA, as well as subcategories of these categories such as HORSE.

Spelke (1991) lists domain specific principles of the object such as boundedness, cohesion, contiguity, rigidity and no action at a distance. Ochiai (1999) exemplifies inferences coming out of combinations of these principles as (i) through (iv).

(i) An object can’t initiate movement. Unless moved by the force of others, it stays where it is.
(ii) An object can’t go through another object. For example, when an object hits the wall, it stops unless you make a hole in the wall.
(iii) An object can’t change its form unless it is forced by others.
(iv) The parts of an object move together; if you pick up a part of the object, the other parts will follow.

These set of principles form what is called folk physics and it tells you about domain specific theory of how objects behave. Such results indicate that ontological categories such as OBJECT are acquired early and consistently. Here is a list of several research results summarized in Ochiai (1999).

- Three years’ old children can distinguish living and non-living by features such as looks, activities, ability to sense, possession of organs such as the heart and the bones, ability to self-reconstruct, and so on (Gelman and Kremer 1991).
• When shown a picture of an unknown object and asked whether it can go up and down the hill on its own, children around 3 to 4 could infer correctly the possibility of the movability of the object. (Massey and Gelman 1988).

• Children at 3 years old know that growth of animals requires change in size. In other words, they expect animals to change along the passage of time. They also know that buttons of clothes do not fix itself. (Inagaki and Hatano 1996).

• Children have ontological categories of animal and non-living, as demonstrated by their understanding that a porcupine cannot changed into a cactus whereas a raccoon can be changed to a skunk, or you cannot change a stuffed dog to a real dog (Keil 1989).

• Children know that artifacts are man-made whereas natural species cannot be made (Gelman and Kremer 1991).

These findings show that knowledge regarding ONTOLOGICAL domain category is real and ONTOLOGIES such as OBJECT, LIVING and ANIMATE are acquired early and are good sources of domains. I would like to add HUMAN and SUBSTANCE (as opposed to OBJECT) to this group.

4.2 Frames of Bateson (1972) and Goffman (1974)
The second origin of the domain is socio-cultural frame in the manner of Bateson (1972) and Goffman (1974). Bateson, who is known for his theory of double bind proposes a concept called “frame”, which includes plays, movies, interviews, jobs and so on. (Bateson 1972:179). Goffman (1974) extended the concept of “frame” and defined the frame or the primary framework as below.

And of course much use will be made of Bateson’s use of the term “frame.” I assume that definitions of a situation are build up in accordance with the principles of organization which govern events—at least social ones—and our subjective involvement in them;0frame is the word I use to refer to such of these basic elements as I am able to identify. (Goffman 1974:10-11)

Indeed a primary framework is one that is seen as rendering what would otherwise be a meaningless aspect of the scene into something that is meaningful. (Goffman 1974: 21)

There are two types of frames, social and natural, and I basically refer to social frame when I use the term frame here. Many of the domains that CMT deals with are frames in this sense, so are many that computer researchers deal with as we will see.

Humans have developed a vast array of social and cultural artifacts and activities such as houses, clothes, marriage, sports, games, travel, science, research, money, finance, stock market, automation, and so on and all these and more have an intricate structure and supporting conceptual system of its own. Frames are the origin of the domain categories such as CULTURAL and SOCIAL and to some extent MENTAL.

4.3 Regions or Cortices of the brain

Figure 1. Simulation (Pulvermüller 2008:86)
The third origin is a domain, in the sense of regions or cortices of the brain (Barsalou, 2008, Pulvermüller, 2008, Damasio, 2010). For example, Pulvermüller (2008) indicate Figure 1, in which different regions activated by different verbs are shown.

As the research on the brain progresses, we know more about how the brain is made up and its function. Humans has visual cortex, auditory cortex, sensory cortex, motor cortex, and so on and these so called domains may just be domains of CMT. PERCEPTUAL, and PHYSIOLOGICAL for sure and EMOTIONAL and MENTAL to a certain extent, come from this origin.

4.4 What holds different domains together
We can position these three origins as different levels in analogy theory: Ontology at the level of elements, frame at the level of relations and systems and finally brain regions at the level of properties of an element. Other major domain categories such as ABSTRACT and SCHEMATIC can be thought of as abstraction, evaluation and systematization of ontologies and properties.

What holds them together as a unified category? I would like to propose that domains form a loosely united category whose members have in common 1) lexical redundancy and 2) inferences. By lexical redundancy I mean that they have two or more lexical items in the domain. For example, LIQUID domain has terms such as pour, splash, flow and so on that are typical to that domain. I would like to adapt the term “anchor” from Shaikh et al. (2014) to express words particular to the domain. In that sense, pour, splash, flow are anchors to the LIQUID domain. Another example is neigh for the HORSE domain. Yet another example is cold, warm, and hot for WARMTH domain. All synonyms, antonyms, and closely associated words are anchors and they form a domain.

The other criterion of inference goes well with the first criterion. When something is UP, it is not DOWN. This is a type of an inference. When something is UP and there is no support, it falls. This is another inference. We know the way things fall, and we know about the resistance and the acceleration. These are all inferences as well. An inference may be propositional or perceptual.

These two loose criteria indicate that almost all concepts can form or belong to some kind of a domain. Although it may seem too loose, it does not contradict the idea of frame semantics advocated by a cognitive linguist Fillmore (Fillmore 1975).

Furthermore, these criteria of domains prove useful for assessing computational approaches. I have selected representative papers from computational linguistics to see what kind of domains are discussed and how they arrive at them, either manually or automatically.

In this section, I discussed domains in CMT. In the following Section, I will take up sample computational research on metaphor and discuss how they treat domains.

5. Domains in computational linguistics

In this section, five sample computational papers are introduced to discuss how computational metaphor approaches treat domains. They are Fass (1991), Mason (2004), Mohler et al. (2013), Strzalkowski et al. (2013) and Shaikh et al. (2014).

5.1 Fass (1991)
Fass (1991) uses a technic called semantic preferences, which is also known as selectional restrictions (Katz and Fodor 1963). For example, a verb drink takes [ANIMATE] as subject and [LIQUID] as object. The selectional restriction mechanism in Fass (1991) mainly deals with the domains I list as ONTOLOGICAL1. Figure 2 indicates the flow chart of how it distinguishes metaphor from metonymy.

Other types of domains are not discussed. Fass (1991) does not seem to automatically identify selectional restrictions. Also, as

---

1 An anonymous reviewer pointed out that “Fass indeed used only a few examples (…) but the selection restriction mechanism itself can deal with other domains.” I changed my wording from only to mainly and maintain that my claims applies to Fass (1991) without a problem. I also see problems in general with using selectionl restrictions to deal with conceptual (systematic) metaphors such as “We are spinning our wheels. (We are trying hard to better our repaship in vain.)” because many of these have more than two meanings or proverbial types such as “Rainy days never stay. (A bad situation change someday.)” because there’s no contradiction within the sentence itself.
it admits, it works only with “a small set of English sentences” (Fass 1991:86).

5.2 Mason (2004)
Mason (2004)’s CorMet is a corpus-based system for discovering metaphorical mappings between concepts. Mason (2004) uses what he terms as “domain key-words” and obtains domain specific documents by submitting queries to the Google search engine. For example, stocks, bonds, NASDAQ, Dow, investment, and finance are domain key-words for FINANCE domain. Queries such as (bonds AND Dow AND investment) or (NASDAQ AND investment AND finance) define FINANCE domain.

Another mechanism in Mason (2004) for finding a domain is to find important verbs in each domain. CorMet achieves this by: 1) obtaining a large sample of domain-relevant documents, 2) decomposing them into a bag-of-words representation, 3) and finding the ratio of occurrences of each word stem to the total number of stems in the domain corpus as shown in Figure 3.

The frequency of each stem in the corpus is compared to its frequency in general English. This method may enable us to pick out all the verbs frequently used for the domain, but risk including target domain words or abstract words. For example, Mason (2004) lists outsourc and forcast as characteristic stems for FINANCE domain, but they can be used more broadly in ECONIMICS and are probably not anchors (domain determiners) particular to FINANCE domain.

5.3 Mohler et al. (2013)
Mohler et al. (2013) collect what is called domain signature. Target domain signature starts with collecting target concept articles from Wikipedia. Then, all the terms in the articles are associated with WordNet senses. Later, these senses are clustered using the graph-based Chinese Whispers algorithm. By eliminating some words using npmi(normalized pointwise mutual information) and TF-IDF(term frequency inverse document frequency), you have the target domain signature. This process is shown in Figure 4.
Source domain signature is built by separating out the target domain signature. Although they talk much about the target domain signature, they hardly talk about the source domain signature. From the paper, it is hard to understand what kind of source domains they have and how they extract them.

The way they extend the target concept *governance* to a domain which includes terms such as *law, government, administrator* seems legit, and it seem to work for all domain categories. However, the current system is set up to find metaphors by “comparing sementic signature of a text with a set of known metaphors.” (Molher 2013:34) and does not detect the source domain automatically nor find new metaphor source domains.

5.4 Strzalkowski et al. (2013)

Strzalkowski et al. (2013) use Topical Structure tracking and imageability. Their system is fully automated. For example, for a target concept *federal bureaucracy*, the system returns sentences including the term and N number of sentences before and after the target sentence (N is set to 2 in the prototype). Topical Structure deals with cohesion (lining up a referential structure) which is made possible by using repetition, synonyms, hyponyms, and pronoun references from WordNet. An interesting assumption here is metaphorical words lie “outside” of the Topical Structure. If we take Topical Structure to be “what the speaker and hearer are really talking about”, then Topical Structure may be equated with the target domain. Also, this treatment eliminates the problem of confining metaphor search within one sentence, which is typical for a system using the technic of semantic preference only.

Another driver for this system is imageability scores, with the assumption that metaphorical words are more imageable. Imageability scores are taken from MRCPD imageability rating. Although the system is at its prototype level, the project seems to be promising.

However, there are two concerns from a point of view of domains. First is the nature of Topical Structure. Imagine there is an article about FRB raising interest rate. It is clearly of FINANCE domain. In the article, the author may describe its meeting, with details of how a member opened the door and how another member pounded the table. Will those descriptions be included in the Topical Structure? Alternatively, the author may discuss his past experience of going to the bank. In other words, the actual topic may involve several different domains. Will the Topical Structure include these other domains related to the main topic?

The other concern is about imageability. Perceptions other than visual can be the source domain as indicated in Section 2. How about olfactory domain? How about PHYSIOLOGY such as PAIN or HUNGER. Are they imageable or not? At this point, I do not know the structure and the method of imageability rating to decide whether it fits perfectly to discern what is metaphorical or not.

---

1 An anonymous reviewer pointed out that “This argument is wrong in that the imageability of a word is the ease to which people can form a mental image associated with that word, and thus it is related to all kinds of perceptions.” I hope that is the case but term “image” is often misunderstood by people as visual, and the way the survey is conducted can influence the rating.
5.5 Shaikh et al. (2014)
Shaikh et al. (2014) take up examples such as cure poverty and construct the source domain of such a metaphor. This paper stems out of IARPA’s metaphor program, which defines 50 candidates of the source domains and DISEASE is one of them. First, linguistic metaphors such as cure poverty are brought into the system. Then a linguistic pattern [cure [OBJ:X/nn]] is formed. Arguments matching variable X are identified using balanced corpus, and these terms are clustered into a semantic category DISEASE(proto-source domain). So called seeds for DISEASE domain are disease, cancer, plague. Seeds are used to discover relations within the domain. For example, nouns, verbs and verb phrases, adjectives that co-occur with seeds with sufficiently high frequency and sufficiently high mutual information are selected. The group of words and relations are together called a conceptual space. The frequency with which these relations can be found with seeds is called Relation Frequency (RF) and Inverse Domain Frequency (IDF) is calculated for all other 50 domains. Table 2 and 3 indicate these processes.

Table 2. Example of seeds in Shaikh et al. (2014)

<table>
<thead>
<tr>
<th>DISEASE</th>
<th>disease, cancer, plague</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABYSS</td>
<td>abyss, chasm, crevasse</td>
</tr>
<tr>
<td>BODY_OF_WATER</td>
<td>ocean, lake river, pond, sea</td>
</tr>
<tr>
<td>PLANT</td>
<td>plant, tree, flower, weed, shrub, vegetable</td>
</tr>
<tr>
<td>GEOGRAPHIC_FEATURE</td>
<td>land, land form, earth, mountain, plateau, island, valley</td>
</tr>
</tbody>
</table>

Table 3. Top ranking relations in Shaikh et al. (2014)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DISEASE</td>
<td>diagnose with</td>
<td>800</td>
<td>V</td>
<td>before</td>
<td>1.94</td>
</tr>
<tr>
<td>DISEASE</td>
<td>afflict with</td>
<td>85</td>
<td>V</td>
<td>after</td>
<td>1.67</td>
</tr>
<tr>
<td>DISEASE</td>
<td>afflict with</td>
<td>33</td>
<td>V</td>
<td>before</td>
<td>1.52</td>
</tr>
<tr>
<td>DISEASE</td>
<td>cure of</td>
<td>29</td>
<td>N</td>
<td>before</td>
<td>1.46</td>
</tr>
<tr>
<td>BODY_OF_WATER</td>
<td>dive into</td>
<td>49</td>
<td>V</td>
<td>before</td>
<td>2.01</td>
</tr>
<tr>
<td>BODY_OF_WATER</td>
<td>wade through</td>
<td>44</td>
<td>V</td>
<td>before</td>
<td>1.88</td>
</tr>
<tr>
<td>BODY_OF_WATER</td>
<td>wade into</td>
<td>42</td>
<td>V</td>
<td>before</td>
<td>1.84</td>
</tr>
<tr>
<td>BODY_OF_WATER</td>
<td>rinse in</td>
<td>41</td>
<td>V</td>
<td>before</td>
<td>1.80</td>
</tr>
</tbody>
</table>

POVERTY IS DISEASE metaphor may have expressions such as cure, which is almost invariably assigned to DISEASE domain, and these terms are called “anchors.” Terms such as spread are less specific. Furthermore, expressions such as “fight against POVERTY” might evoke ENEMY (31%) more than DISEASE (17%) and also evoke ANIMAL or MONSTER (less than 10%), but this expression may still be compatible with DISEASE metaphor.

In sum, Shaikh et al. (2014)’s system seems most complete and the use of anchors to identify domains seem promising.

6. Conclusion
In this paper, I discussed the nature of domains in Cognitive Metaphor Theory and went over some computational metaphor approaches to see what domains they deal with and how they arrive at the creation and identification of domains. An approach that identifies all types of domains should be an ideal conceptual metaphor identification system and I believe taking the idea of the domains seriously will help arrive at a yet better system for conceptual metaphor identification.

References


