Denotation of Concepts  
Taking the term ‘Community’ as an example

Lan-ting Tseng¹, Feng-yan Lin², Yi-liang Chen³  
¹ & ², Graduate Institute of Building and Planning, National Taiwan University, Taiwan  
², also with the Institute of Information Science, Academia Sinica, Taiwan  
³, Dept. Mathematical Statistics and Actuarial Science, Aletheia University, Taiwan

Abstract. In the process of co-design, there is no urban planner able to keep away from interdisciplinary collaboration. In order to negotiate with variant professionals, planners or designers usually spend so much energy to integrate information for their needs. Therefore, formalizing conceptual network among domain experts may be helpful to mine tacit knowledge. Based on cognitive semantics and ontological engineering, the research tries to identify concepts. An analytical mechanism is proposed to make terms comprehensible. The term ‘community’ is taken as the target artefact to illustrate possible contexts. Its senses and relationships would be explored in relevant knowledge domains.

Keywords: Conceptual network; denotation; encapsulation; ontology.

1. Introduction

Natural languages in our daily life are full of imprecise and ambiguous concepts. Many terms are used in a ‘conventional’ manner, where common understanding and background knowledge of terms are implicitly assumed, so that sentences and communication times can be largely shortened but still effective. However, language is evolutionary in its use and meaning. As a fact, words of colloquial speech have vague, abstract, and far-reaching meanings. They refer to various aspects of an object field and have shifting boundaries. This is an important issue regarding words and ideas as a rich media to a real world (Lin, 1998:769; Poggenpohl et al, 2004).

Ontology is a branch of philosophy as a systematic account on the nature and organization of reality (Simoff & Maher, 1998). By ways of the ontological approach, many researches contribute to decrease the cost of externalizing knowledge. For AI, the concepts of ontology are formal systems representing domain concepts and their linguistic realizations with basic elements (Tu et al, 2002).

Collaborative design needs communication. In particular, the phase of conceptual design before form-making involves a lot of verbal communications. During the stage of fermentation of an emerging concept, the researchers are interested in how to deliberate and capture the new ideas in a formal way (Lai, 2004)? It is argued that representing conceptual design should facilitate multiple interpretations of design elements, which may be modified in various ways (Emdanat & Vakalo, 1998).
Since speeches are rich fields to study design, the research takes the term ‘community’ as an example to depict what a concept implies. We use three descriptive texts as samples. These documents are about what interviewees have been through during the event of Severe Acute Respiratory Syndrome (SARS), which was a viral respiratory illness in Taiwan in 2003 (Lin et al, 2004).

2. Analysis

2.1 Sentences in well-define form

Traditional techniques of identifying key concepts are usually looking for nouns, pronouns, and noun phrases from narrative sentences by parsing their syntactical structure, one of whose simple examples is shown in (1).

Subject (Â) x Verb (Â) x Object (Â) -> Sentence (Â)       (1)

Where  is a language, Subject (Â), Verb (Â) and Object (Â) are sets of subjects, verbs, and objects of  with finite lexicons as their elements. However, there are shortcomings in the technique mentioned above. For example, many sentences of inappropriate forms are not easily to parse (Lee & Tepfenhart, 2001: 82–83). Instead of parsing their syntax, this research takes three phases, including rephrasing, inferring, and encapsulating, for transforming initial sentences to those of well-defined forms (Figure 1).

The transformation mechanism  can be shown as (2).

\[ \Psi = (S, \Sigma, \Phi, C, P) \]  

where  is the set of original sentences in a descriptive text;  is the set of sentences which are rephrased from the original sentences and become more comprehensible;  is the set of sentences of single idea;  is the set of tuples of well-defined forms;  is a set of three families of functions of rephrasing (ρ1), inferring (ρ2), and encapsulating (ρ3), whose signatures are shown below and in turn map original sentences in  to rephrased sentences in  single idea sentences in , and well formed tuples in .

\[ \rho_1: S \rightarrow \Sigma \]  
\[ \rho_2: \Sigma \rightarrow \Phi \]  
\[ \rho_3: \Phi \rightarrow C \]

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Sentence or Tuples</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_i )</td>
<td>On the day of Jun 2, people in the civil community said that one of the reported case of SARS is a granny.</td>
<td>A012</td>
</tr>
<tr>
<td>( s_i' )</td>
<td>On the day of Jun 2, a granny living in civil community was reported as a case of SARS.</td>
<td></td>
</tr>
<tr>
<td>( t_j )</td>
<td>A granny is a reported case of SARS.</td>
<td></td>
</tr>
<tr>
<td>( t_j )</td>
<td>The reported time is a date of June 2.</td>
<td></td>
</tr>
<tr>
<td>( \langle WT_{1b}, WR_{1b}, WS_{1b} \rangle )</td>
<td>The case lives in a civil community.</td>
<td>A01201</td>
</tr>
<tr>
<td>( \langle WT_{1b}, WR_{1b}, WS_{1b} \rangle )</td>
<td>(granny, is-a, reported case of SARS)</td>
<td>A01201</td>
</tr>
<tr>
<td>( \langle WT_{1b}, WR_{1b}, WS_{1b} \rangle )</td>
<td>(reported time, is-a, date of June 2)</td>
<td>A01202</td>
</tr>
<tr>
<td>( \langle WT_{1b}, WR_{1b}, WS_{1b} \rangle )</td>
<td>(case, live, civil community)</td>
<td>A01203</td>
</tr>
</tbody>
</table>

Table 1, the assumption of all symbols, for some i.
For example, in Table 1, let $s_i$ be the sentence “On the day of Jun 2, people in the civil community said that one of the reported case of SARS is a granny.” The code number of $s_i$ is happened to be given by A012. Using the functions of $\rho_1, \rho_2, \rho_3$, the following transformations are made.

$$\rho_1 (s_i) = s_i'$$ \hspace{1cm} (6)

where $s_i' = “On the day of Jun 2, a granny living in civil community was reported as a case of SARS.”$

$$\rho_2 (s_i') = t_{ij}, j=1,2,3$$ \hspace{1cm} (7)

where $t_{i1} = “A granny is a reported case of SARS.”$

where $t_{i2} = “The reported time is a date of June 2.”$

where $t_{i3} = “The case lives in a civil community.”$

$$\rho_3 (t_{ij}) = (WT_{ij}, WR_{ij}, WS_{ij})$$ \hspace{1cm} (8)

where $WT_{ij}, WR_{ij}$ and $WS_{ij}$ are lexicons in $t_{ij}$, and $WR_{ij} \in$ Verb (Â). In the above example, $WT_{i1} = ‘granny’, WR_{i1} = ‘is-a’, WS_{i1} = ‘reported case of SARS’, the function of encapsulating maps the single idea sentence “A granny is a reported case of SARS.” to the tuple (granny, is-a, reported case of SARS), whose code is A01201.

Finally, a database can be built according to the tuples, where $WT_{ij}$ is also called target concept, $WS_{ij}$ source concept, and $WR_{ij}$ operator for convenience. It is noted that the operator $WR_{ij}$ can also be interpreted as a relationship between target concept $WT_{ij}$ and source concept $WS_{ij}$. Thus, the operator $WR_{ij}$ belongs to one of the four basic relationships, namely generalization, aggregation, dependency, and association, in terms of Unified Modeling Language (UML; Fowler, 2004). For example, Table 2 illustrates a portion of database where the term ‘community’ shows up within the target or source concept.

<table>
<thead>
<tr>
<th>Code</th>
<th>Target concept</th>
<th>Source concept</th>
<th>Operators</th>
<th>Generalization</th>
<th>Aggregation</th>
<th>Dependency</th>
<th>Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>A10903</td>
<td>threats</td>
<td>Community contagion</td>
<td>Is_a</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A06301</td>
<td>community named</td>
<td>pattern of space</td>
<td>Has_a</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A07702</td>
<td>community</td>
<td>Inhabitants</td>
<td>Has_a</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A11601</td>
<td>community</td>
<td>Contagion case</td>
<td>Has_a</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A11103</td>
<td>experience of daily life in community</td>
<td>Activity of operation</td>
<td>Cause</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A05702</td>
<td>community</td>
<td>Case</td>
<td>Increase</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A11002</td>
<td>community</td>
<td>participation</td>
<td>Need</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A11002</td>
<td>community</td>
<td>participation</td>
<td>Need</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A07703</td>
<td>community</td>
<td>Volunteer</td>
<td>assemble</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A07002</td>
<td>community</td>
<td>Information</td>
<td>generate</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A08305</td>
<td>community</td>
<td>External control</td>
<td>accept</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A07301</td>
<td>community</td>
<td>Inhabitants</td>
<td>Mobilize</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A09101</td>
<td>community</td>
<td>People</td>
<td>Mobilize</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A08307</td>
<td>community</td>
<td>Community</td>
<td>Help</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A08303</td>
<td>community</td>
<td>The Interior of community</td>
<td>Self manage</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A08903</td>
<td>community</td>
<td>Network of Immunity</td>
<td>Run</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2, all the tuples with the concept ‘community’ are collected from sentences in the ‘A’ document.
3. Results

3.1 The Set of ‘Source Concept’

The extended senses of a target concept can be deduced using Table 2. For example, Table 3 makes a list of terms from the column of ‘Source concept’ in Table 2 where the lexicon ‘community’ appears in ‘Target concept’. Thus, the terms we collect are concepts as possible extending senses of the term ‘community’ described by interviewees ‘A’, ‘B’ and ‘C’.

3.2 Categorizing Extended Source Concepts

The extended source concepts can be further categorized. For example, all the terms relevant to ‘human being,’ which is one of extended source concept of target concept ‘community’ and shown in Table 3 with bold-faced alphabets, are shown in Figure 2.

Thus, the terms form a set named Class_ Source_‘human being’

\[
\{ \text{ws} | \text{ws is a term relevant to ‘human being’} \}
\]

\[
= \{ \text{case, contagion case, inhabitants, illegal inhabitants, people, people in need, professionals, volunteer, voluntary inhabitants} \}
\]

Based on (9), the extended source concept ‘people’ can be categorized into variant subclass in criteria of some certain attributes. Such as, the class ‘People’ can be judges by their addresses to
decide if they are inhabitants of a community. Also, the subclass 'Inhabitants' can find out suspicious contagious cases in term of body temperature. Figure 2 is a diagram with a tree structure.

3.3 Illustrating concepts by UML

The UML is a family of graphical notations which help in describing and designing software system, particularly for building in object-oriented style (Fowler, 2004). According to Table 2, Figure 3 is a diagram revealing the conceptual network of the term 'community'.

Obviously, there are variant elements drawn in the diagram which is constructed by graphical components with definitions given by UML. For the sake of its relatively open standard, this research applies its tools as outlining in a reverse-engineer direction. (Fowler, 2004: 1-3). Although it is not a goal to compile a program in this research, building a diagram like Figure 3 gives benefits of making something implicit visible.

4. Conclusive Discussions

In viewpoint of data processing, natural languages possess not only strings of syntactical structures but also senses embedded mentally. As an issue of knowledge representation, what a challenge is to make documents both machine-readable and comprehensive precisely as human being.

This research proposes an analytical method to construct a well-defined conceptual network, which is to build up a relational data model in form of OODA from decomposing descriptive texts. Furthermore, a schema for integrating data is the main requirement to explore concepts and their relationships. Hence, the steps of mechanism are not only reconstructing a sentence as an artefact, but also transforming into tuples to mine potential sources and operators within concepts with respect to a specific target term.

Based on results of analysing the term 'community', the database we obtain is helpful to com-
pare what constitutes a concept among variant descriptive texts, find out and categorize the set of extending senses in common, and visualize the implicit context by UML.

In term of semantics, the concept of ‘community’ is not a lexicon independent of other relevant senses. Even in an approach to ontology engineering, it can be observed that a term extends senses vary in some situations dynamically. After all, what obstruct knowledge representation most are issues about techniques of formalization. Behaving in a well-defined form remains the key point to build up a database. Therefore, finding more effective mechanisms, as denotation in the research, may go further at next phase.

5. Acknowledgements

For all descriptive texts we use in the research, this paper is dedicated to communities suffering from the attack of SARS in Taiwan.

6. Reference


