Abstract

Terminologies of any domain contain a substantial number of compounds or multi-word terms. Through the shared constituent elements of the compounds, terms form complex networks, which constitute a terminology. A variety of methods developed for the analysis of complex networks can therefore be used for exploring the structure of terminologies. In this paper we will apply a theoretically valid method for detecting communities in networks to the partitioning of terminologies or clustering terms, in order to explore the structure of terminologies. Experiments show the usefulness of the method for analysing the motivated structure of terminologies.

1 Introduction

Terminologies of any domain contain a substantial number of compounds. In Japanese terminologies, for instance, about 85% of the terms are compounds (Nomura and Ishii, 1988). Terms are thus connected with one another through shared constituent elements and form complex networks which constitute the terminology. The exploration of these networks, which reflect the motivatedness of terms and show the associative structure of terminologies, is an important theoretical issue for the formal analysis of terminology and of lexicons in general, with potential applications such as dictionary compilation, postprocessing of automatically extracted sets of terms, etc. It is complementary to the description and processing of the syntagmatic aspects of complex terms (e.g. Daille, et. al., 1996; Jacquemin, 2001) as well as to text-oriented terminology studies, including such applications as term extraction and clustering (e.g. Heid, 1999; Weeds, et al., 2005). While these latter types of studies have been actively pursued, the associative structure of terminologies has not been studied to the same extent, though important exceptions exist (Hamon and Nazarenko, 2001; Grabar and Zweigenbaum, 2004).

This paper explores the complex networks of terminologies, by applying a network community detection method based on statistical mechanics to carry out a divisive clustering or partitioning of terminologies (Scott, 2000). Our theoretical stance, which is described in section 2, is that there is a theoretical precedence of terminologies as a set over individual terms occurring in texts (Kageura 2002), and thus we need to explore the structural nature of terminologies fully, in such a way that the exploration methods properly reflect theoretical views on and the interpretative framework of terminologies, and enable us to qualitatively interpret the results. We argue for the appropriateness of a Potts spin glass model-based divisive clustering of complex networks as a theoretical model for exploring and visualising (not in the sense of visualisation but in the sense of exploring the hidden structure of) the terminological networks.

In the following, we will first clarify the theoretical status of terminological networks constructed by common constituent elements in the study of terminology. We will then argue that to explore these associative terminological networks, a divisive clustering method based on the Potts spin glass model allows a valid interpretations of the results. After applying the method to the mixed terminologies of three domains and observing its performance from the point of view of divisive
Table 1. A putative terminology consisting of 12 terms

<table>
<thead>
<tr>
<th>“text classification”</th>
<th>“document classification”</th>
<th>“information retrieval”</th>
<th>“library classification”</th>
</tr>
</thead>
<tbody>
<tr>
<td>“disease diagnosis”</td>
<td>“document information”</td>
<td>“medical information”</td>
<td>“medical diagnosis”</td>
</tr>
<tr>
<td>“medical aid system”</td>
<td>“automatic text classification”</td>
<td>“diagnosis record”</td>
<td>“disease image record”</td>
</tr>
</tbody>
</table>

Figure 1. A network constructed from the sample terminology in Table 1

clustering, we will analyse, both quantitatively and qualitatively, the characteristics of the network of documentation terminology, contrasting it with the network of documentation and artificial intelligence terminologies combined.

2 Terminological networks

Take, for instance, a putative terminology consisting of 12 terms (Table 1). This terminology forms the network shown in Figure 1. Formally, this network is an undirected weighted graph, with vertices representing terms and edges representing shared constituents between terms. Naturally:

\[
\text{degree}(v_i) \approx \sum_j \text{frequency}(c_{ij}) \tag{1}
\]

\[
\text{weight}(e_{ik}) = |\{c_{ij}\} \cap \{c_{kl}\}| \tag{2}
\]

where \(v_i\) is the vertex or term with index \(i\), \(c_{ij}\) is a \(j\)-th constituent of the term representing \(v_i\), and \(e_{ik}\), which is defined by the number of common constituent elements between the two terms \(v_i\) and \(v_k\), is the edge between \(v_i\) and \(v_k\). As we will see in section 4, even for a small- or medium-sized terminology consisting of one to two thousand terms, this network becomes very complex with tens of thousands of edges. As the degrees of vertices reflect the frequency distribution of constituent elements in the terminology, which follows the power law (Kageura, 1998), the network has a scale-free nature (Albert and Barabasi, 2002).

These networks have a theoretical status in the study of terminology. Firstly, we observe that terminology as a set or terminological space, as opposed to individual terms, constitutes an object of study in its own right. It is the concept of terminology as a set that enables us to recognise individual terms as empirical objects. Unlike sentences, terms cannot be arbitrarily constructed by individuals; terms become terms only when they are accepted by a community and become part of a terminology. This process is not determined by any person or group. This, combined with the fact that complex terms show a strong inclination towards motivatedness vis-à-vis the overall terminology (Nomura and Ishii, 1988; Sager, 1991), supports the argument not only for the theoretical importance but also for the reality of terminological space which is socially constructed and shared. This in turn is internalised — unlike individual terms, not consciously — in subject-specialists’ minds. It is a terminological equivalent of the lexicon — a concrete social product — that belongs to the Saussurian langue (Saussure, 1910-1).

This terminological space is governed by the conceptual system or field of the domain as well as by the system of linguistic representations of concepts (Sager, 1991). Terms are materialised within this terminological space, guided by the conceptual system and using linguistic items (Rey, 1985). The process can be postulated as: (i) concepts to be represented are identified within the overall conceptual system; (ii) in accordance with the nature of concepts to be materialised as terms within the relevant conceptual subsystem, linguistic items representing required conceptual features are consolidated; (iii) terms are formed using these linguistic items, following the dictates of the conceptual subsystem and the system of linguistic representation of concepts; and (iv) the terms, when appropriate, are incorporated in the terminology. The standard description of the formation of terms by means of the relation between a term and its constituent elements belongs to stage (iii). Within this model, a well-structured terminology with many motivated terms can be seen as constructed within a well-coordinated system of concepts and linguistic representations, and a less structured terminology with fewer motivated terms within a less
systematic or ‘hot’ and chaotic system of concepts and linguistic representations. Note here that there is a general correlation between the systematicity of concepts and the systematicity of linguistic representations.

Correspondingly, the network structure defined over the terminology through the shared constituents is not merely a convenient, practical or opportunistic relational representation of a group of individual terms. Rather, to the extent that the terminology of a domain reflects the nature of a conceptual system through its relatively motivated and partially systematic linguistic forms of terms, the terminological network thus constructed reflects the conceptual system that underlies the terminology and the system of linguistic representations. Detecting communities in terminological networks is therefore a means to explore these dual systems of concepts and linguistic representations.

3 Community detection algorithm

3.1 State of the art

A number of community detection or clustering methods have been proposed and used to date (throughout this paper we use ‘community detection’ and ‘clustering’ synonymously; we also use ‘partitioning’ when we emphasise the divisive aspect of community detection). The methods can be divided into divisive/partitional and agglomerative approaches (Jain, et al., 1999; Newman, 2004; Scott, 2000). Within our context, a divisive approach is the natural choice, because it regards the terminology as a whole as the basic object of study, while agglomerative approaches — including relatively new methods (Handl, et al. 2003) — regard individual terms as the basic object.

With the growth in research into complex networks (Albert and Barabasi, 2002; Masuda and Konno, 2005; Newman, 2003; Strogatz, 2001; Watts and Strogatz, 1998), a wide variety of divisive algorithms have been proposed and examined recently (Duch and Arenas, 2005; Fortunato, et al., 2004; Fujishige, 2002; Newman, 2004; Newman and Girvan, 2004), in addition to a traditional minimum cut algorithm (Cheng and Wei, 1991; Fujishige, 2002). From the point of view of clustering applications, it goes without saying that the method that performs best vis-à-vis an ideal result should be chosen. So the comparative examination of different methods has been carried out using common test data (e.g. Danon, et al., 2005). In clustering applications, the performance and the upper bound differs, depending on the extent to which the necessary features are present. The present framework, on the other hand, is different from that of common clustering applications. Firstly, it is common knowledge in the field of terminology that the structure of each underlying system of concepts is multifaceted and multidimensional, that the coherency of the conceptual system varies from one terminology to another, and that the system of linguistic representations is not totally motivated. The first two characteristics lead to ambiguity concerning what are ‘ideal’ clusters, which in turn leads to deconstruction of the concept of a ‘gold standard’. The third characteristic means that the features are only partially provided. In addition to these, as the essential aim of using divisive clustering in this study is to employ it as a guide to the qualitative diagnosis of the systematicity of concepts and the motivatedness of linguistic representations, the method to be adopted should fit the basic theoretical and interpretative framework of the universe of terminological networks. This contrasts with the quest for the method with the best clustering performance vis-à-vis some ‘gold standard’. Incidentally, applying graph-based methods to word clustering is not new (Aizawa and Kageura, 1998; Matsuo, et al., 2006), but this prior work is application-oriented and uses cooccurrences of words.

3.2 Clustering using the Potts spin glass model

After examining the interpretative relevance of various methods (Cheng and Wei, 1991; Duch and Arenas, 2005; Fortunato, et al., 2004; Fujishige, 2002; Newman, 2004; Newman and Girvan, 2004), we adopted the algorithm developed by Reichardt and Bornholdt (2006), which is based on an infinite range Potts spin glass model (Dorogovtsev, et al., 2004; Nishimori, 1999). A similar model is used in NLP by Takamura, et. al. (2005) for extracting the semantic orientations of words. We found that this method fits our interpretative framework admirably. But before presenting our argument for adopting this method, let us first briefly describe the Potts spin glass-based method.

Formally, the Potts model consists of a lattice of $N$ sites, on each of which is placed a spin that can take $q$-states. The Hamiltonian or energy function $H$ (which is used as a cost function to be min-
imised in the algorithm) of the Potts model is given as:
\[
H({\{s\}}) = - \sum_{(i,j)} J_{ij} \delta(s_i, s_j)
\]
where \(s_i\) and \(s_j\) represent the state of spins \(i\) and \(j\), and \(J_{ij}\) represents the strength of mutual influence between \(s_i\) and \(s_j\). \(\delta\) gives 1 when \(s_i\) and \(s_j\) are in the same state and 0 otherwise.

Here \({\{s\}}\) follows the Boltzman distribution:
\[
P({\{s\}}) = \frac{\exp(-\beta E({\{s\}}))}{\sum_i \exp(-\beta E({\{s\}}))},
\]
where the denominator is the normalisation factor called the partition function, \(E\) gives the energy and \(\beta\) is a constant called inverse temperature.

Starting from the assumption that the Hamiltonian or the quality function for detecting communities in a network should: (i) reward internal edges between vertices of the same group (or in the same spin state); (ii) penalise missing edges between vertices in the same group; (iii) penalise existing edges between different groups; and (iv) reward nonlinks between different groups, and with a few simplifying assumptions, Reichardt and Bornholdt (2006) proposed the following Hamiltonian:
\[
H({\{\sigma\}}) = - \sum_{ij} (W_{ij} - \gamma p_{ij}) \delta(\sigma_i, \sigma_j),
\]
where \(W_{ij}\) denotes the adjacency matrix of the graph with normalised weight for edges, \(p_{ij}\) denotes the null model that gives the baseline probability that a link exists between node \(i\) and \(j\) (in actual applications, we chose the uniform distribution for \(p_{ij}\)), and \(\gamma\) is the parameter for distributing the weight between the reward for internal links and the penalty for internal nonlinks. This represents a spin glass with mutual influence \(j_{ij} = W_{ij} - \gamma p_{ij}\).

This Hamiltonian can be transformed to the modularity \(Q\) defined by Newman and Girvan (2004), which compares the actual and expected fractions of links between vertices in the same group:
\[
Q = \sum (e_{ii} - a_i^2)
\]
where \(e_{ii}\) indicates the actual fraction of links that connect vertices in the group \(i\), and \(a_i^2\) indicates the corresponding expected value. The modularity takes a value between 0 and 1. The higher value corresponds to a lower Hamiltonian and vice versa. We will later use the modularity instead of \(H\) as a top-down quality index of community detection.

The solution of the divisive clustering is given by the assignment of spins in the ground state of the system. Reichardt and Bornholdt (2006) give a computationally tractable method of finding the ground state by simulated annealing, based on the heat bath update rule at a given temperature. An R library as well as a C library with a Python interface that carry out the computation is made available by Csardi (2006). We used R library for the analyses.

Defining the problem of community detection as finding the ground state for the Potts spin glass model gives a straightforward interpretation in terms of the systematicity of terminology. The ground state, in which the Hamiltonian is minimised and the modularity is maximised, is the most ‘natural’ state that appears when the ‘temperature’ of the system has cooled down. In the case of a terminological network, the physical concept of ‘low energy’ reflects a state in which the maximally motivated structure is retained. As such, quite apart from the results of the clustering, the minimum value of the Hamiltonian (or correspondingly, the maximal value of the modularity) can be interpreted as indicating the basic motivatedness of the terminology. The resultant clusters can be interpreted as the basic and maximally motivated state of the terminology when the contribution of ‘excited’ states is suppressed.

In short, we obtain not only the clustering result for each terminology network, but also an index that reflects the basic degree of systematicity or motivatedness of the terminology. Given that any terminology is only partially motivated, which, from the point of view of clustering, means that the features are not completely given and the upper bound of the clustering results cannot be determined with 100 per cent validity, having some parameters that correspond to the natural interpretation of the object under observation is particularly important for analysing the terminological data — more important than choosing a method that happens to give the highest performance vis-à-vis some sort of ‘gold standard’. The meaning of the Hamiltonian given in the Potts spin glass model and the correspondence between the Hamiltonian and the modularity provide us with such a parameter.
Table 3. Basic quantities of the mixed terminological data (plus DC)

<table>
<thead>
<tr>
<th></th>
<th># terms</th>
<th># dupl. terms</th>
<th># const.</th>
<th># dupl. const.</th>
<th># edges</th>
<th>max 6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC</td>
<td>1228</td>
<td>-</td>
<td>831</td>
<td>-</td>
<td>11251</td>
<td>1046</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>DC+AI</td>
<td>2918</td>
<td>35</td>
<td>1823</td>
<td>306</td>
<td>36479</td>
<td>2571</td>
<td>-</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>AI+GG</td>
<td>3502</td>
<td>3</td>
<td>2508</td>
<td>145</td>
<td>34299</td>
<td>2857</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>33</td>
</tr>
<tr>
<td>GG+DC</td>
<td>3008</td>
<td>0</td>
<td>2085</td>
<td>101</td>
<td>27521</td>
<td>2405</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td>ALL</td>
<td>4695</td>
<td>38</td>
<td>3001</td>
<td>483</td>
<td>54426</td>
<td>3942</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>30</td>
</tr>
</tbody>
</table>

This method also provides the following additional benefits: (i) it allows us to control the number of communities or clusters by using the parameter $\gamma$ or by explicitly setting the maximum number of spin states; (ii) it allows us to detect communities hierarchically, though we will not explore this aspect in this paper; and (iii) the distribution of the sizes of the detected communities is not overtly skewed, unlike the decomposition carried out by such methods as minimum cuts. In relation to point (iii), we also applied minimum cut algorithms to the experimental data introduced below. As the terminological networks have a scale-free nature, the minimum cuts only cut out small groups and do not give desirable results, which is in conformity with Cheng and Wei (1991).

4 Clustering experiments

4.1 Experimental setup

In order to obtain a general idea of how the Potts spin glass model-based divisive clustering works, we first carried out an experiment in which the performance of clustering could be quantitatively observed. For the experiment, we used Japanese terminologies of three domains as basic data: documentation (DC) (Wersig and Neveling, 1984), artificial intelligence (AI) (JSAI, 1990), and geography (GG) (JME, 1981). The sources for these terminologies were standard and/or widely used reference lists used from the 1980s to 1990s. As a domain, DC and AI are close to each other, while GG is distant from the other two. The basic quantities of the terminological data are given in Table 2.

In order to quantitatively evaluate the nature of the method from the point of view of clustering performance, we carried out an experiment that involved re-partitioning mixed terminological networks made from pairwise combinations of the terminologies, as well as a combination of all three terminologies.

In constructing the terminological networks, we did not consider symbols, numbers or 25 functional elements such as “∅” and “∪”. We did not distinguish heads from modifiers, for two main reasons: (1) as we are concerned with the conceptual system and the terminological space rather than the syntagmatic aspects of individual forms or formation of terms, we deliberately put aside this distinction as the first step; and (2) from the point of view of conceptual systems, there are many heads that do not bear a heavy conceptual burden, such as “method” in English and many single-Chinese character elements in Japanese, which may be misleading in terms of the role of individual constituents vis-à-vis the conceptual system (these elements exist even after removing functional elements in constructing the networks).

Table 3 shows the number of terms, duplicated terms, constituents and duplicated constituents, and the distribution of subgraphs of the four types of terminological data compiled from the three terminologies. The data for DC is also indicated, to be referred to in section 5. In the subcolumns under ‘subgraphs’, ‘max’ means the largest subgraph, and the values indicate the number of vertices in the largest subgraph. The columns ‘6’ to ‘1’ indicate the number of vertices in the subgraphs, and the values indicate the number of graphs with 6 vertices, 5 vertices, etc.

From Table 3 we can see that the graphs of all four domains are decomposed into a single very large graph and a number of subgraphs with at most six vertices. The task here is to partition the largest subgraphs in Table 3 into two for DC+AI, AI+GG and GG+DC and three for ALL, and evaluate to what extent the original terminologies of
different domains can be regrouped. Though the distributions of the subgraphs themselves reflect the motivated nature of the terminological structure, we will not deal with this aspect in this paper.

4.2 Results

Table 4 shows the results of the evaluation. Note that, except for GG+DC, there are duplicated terms, so the grand totals in Table 4 become the number of vertices in the largest subgraphs in Table 3 plus the duplicated terms in the largest subgraphs. Q indicates the modularity, and $E_c$ indicates the number of removed edges.

<table>
<thead>
<tr>
<th></th>
<th>Group size</th>
<th># groups</th>
<th>$Q$</th>
<th>$E_c$</th>
<th>Group size mean</th>
<th>max</th>
<th>min</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC</td>
<td></td>
<td></td>
<td>0.8357</td>
<td>1823</td>
<td>43.6</td>
<td>109</td>
<td>4</td>
</tr>
<tr>
<td>DC+AI</td>
<td></td>
<td>31</td>
<td>0.7726</td>
<td>26837</td>
<td>82.9</td>
<td>220</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 5. Basic quantities of the maximum-modularity partitions

5 Qualitative observations

5.1 Experimental setup

For the qualitative observation of the terminological network, we used the DC terminology, which is familiar to the authors. We also used DC+AI terminologies, for the purpose of comparison. We selected DC+AI because, as shown below, mixed terminologies of AI+GG and GG+DC are rather clearly partitioned at the topmost level. This implies that the subsequent partitioning will be carried out mostly within the terminologies of the original domain. We wanted a terminology of an ‘intermingled’ nature, in order to guide the qualitative characterisation of the DC terminology.

The DC row in Table 2, and the DC and DC+AI rows in Table 3, show the basic quantities of the data. For these data, we applied the method by changing the parameter of the maximum number of clusters, which corresponds to the maximum number of possible spin states. The clustering results with maximum modularity value were chosen for the qualitative observations.

Though the modularities (Q) are not high, the terminologies of distant domains are separated well. Although there are no exactly comparable studies, and the simple quantitative performance of the partition itself is not the main concern of this paper, the performance of groupings for AI+GG, GG+DC and ALL in relation to GG, which range from around 84% to 98%, are comparable to the results of many experiments in text or word clustering and classification. The terminologies of closer domains are, on the other hand, mixed, as expected. This indicates: (i) the terminological networks constructed through common constituent elements contain, as a mass, the structural information of terminologies, as we discussed in section 2 to a substantial degree; and (ii) the Potts model-based clustering, which gives a straightforward terminological interpretation, provides a useful means for exploring the terminological networks.
Figure 1. A network constructed from the sample terminology in Table 1. Top left: A (“indexing”); top right: B (“storage and retrieval”); bottom left: C (“bits, codes, characters and alphabets”); bottom right: D (“analysis, research, researchers, users”)

5.2 General tendencies

Table 5 shows the quantitative data for the clustering results with maximum modularities for the DC and DC+AI terminologies. The number of communities for DC is 24 and that for DC+AI is 31. Incidentally, for DC+AI, 12 communities consist of more than 80% AI terms, and 6 consist of more than 80% DC terms. The other 13 communities are highly hybrid. It is notable that, for DC, the modularity attains a high value of 0.8357 with the removal of less than 2,000 edges, while for DC+AI, the maximum modularity is attained by the removal of 26,837, nearly 15 times the number of removed edges for DC. The modularity value 0.7726 of DC+AI is still much lower than that of DC. These reflect the difference between the coherent terminology of DC and the hybrid set of terms consisting of DC+AI.

We visualised the in-community network using the Kamada-Kawai algorithm (Kamada and Kawai, 1989) in order to analyse the resultant communities. The following four types of community were identified for both DC and DC+AI terminologies:

A. More or less coherent cluster with a single center and possibly with some peripherals (less than 20%), to which we can assign a single generic label relevant to the domain;

B. More or less coherent clusters with more than one center and possibly with some peripherals, to which we can assign a generic label relevant to the domain;

C. Clusters with more than one center and some peripherals, to which we can assign a set of related labels relevant to the domain;

D. Clusters with multiple centers and peripherals, to which we can only assign either a set of unrelated labels or a very general, non-domain specific label.

An example of each type is shown in Figure 2.

Table 6 shows the classification of the communities according to A–D. For intuitive understanding we assigned labels to the communities. For DC+AI, we also added the label AI or DC for those which consist of mostly (more than 80%) AI terms or DC terms.
Table 6. The classification of communities detected in DC and AI+DC

<table>
<thead>
<tr>
<th></th>
<th>DC</th>
<th>DC+AI</th>
</tr>
</thead>
</table>

5.3 Interpretations and diagnosis

The modularity values in Table 5 and the distribution of the types of community in Table 6 clearly show that the DC terminology has a much more motivated structure, which shows the dual systematization of both the underlying concept system and also the way terms are materialised within the terminological space. A large number of type D communities for the DC+AI terminology can be interpreted as the mutual interference between terms that follow the AI system and those that follow the DC system.

To examine this further, we looked at the structure of the network within the communities. The major causes of ‘hybridness’ are as follows:

1. The existence of a few terms consisting of two or more highly meaningful and frequent elements. A typical example is “program language”. This term attracts all the terms containing “program” as a constituent as well as all the terms containing “language”. We thus have a hybrid community of “language, programs, concepts, robots” in the DC+AI terminology. In this case, even though terms belonging to different groups are classified in the same community, they tend to make chunks in the community. If the two constituents are conceptually close to each other, the resultant community will become type B (the group “symbols and symbolic representations” from DC is constructed by the existence of “representation symbol”). The latter case is expected to be the standard case for a systematic terminology of a single domain. This exists in the DC terminology as well as in the DC+AI terminology.

2. The existence of constituent elements which occur frequently but represent only secondary concepts. Examples are “automatic”, “method”, etc. This leads to a community such as “automatic, misc.”, in which we cannot identify conceptual consistencies. An interesting point is that for the DC terminology, this kind of element does not contribute to making hybrid communities. The elements with heavier conceptual content might have suppressed the effect of these elements in a systematic terminology. Though observations of a wider range of terminologies are necessary, it might well be the case that this sort of community is unlikely to be formed for a coherent terminology representing a coherent domain. Incidentally, for the analysis of terminological structures, it would be interesting to analyse the community structure exploratively by adding or subtracting these constituents. For the practical aim of obtain-
ing good clustering results — though it is not our main concern here — excluding these constituents by such measures as tfidf in the construction of the network might well give better results.

3. Hybrid communities the causes of which are not clearly identifiable. A typical example is the community “analysis, research, researchers, users”. It seems natural to regard this type of hybrid community as a “dust bin”, in which a small number of peripheral terms are chunked together through the community detection process. The number of this kind of community is low in DC (only one) but three in DC+AI (“space, problems”, “semantics and discourse, misc.”, “automatic, misc.”).

Going back to Table 6, it is interesting and important to observe the relationships between communities. For instance, although a substantial ratio of communities have high internal coherence in the DC terminology in its minimum modularity ground state, how larger communities are formed also matters for exploring the motivated structure of terminologies. It would be helpful to use the method hierarchically, restricting the number of communities that can be formed. We will cover this in the next step of our research.

6 Conclusion
Starting from defining the terminological networks constructed by shared constituent elements, we have (i) clarified the theoretical status of these networks for the study of terminology, (ii) adopted the Potts spin glass model-based community detection method to explore the networks and argued that the method allows a very natural interpretation of the conceptual characteristics of terminological structures and networks which is theoretically relevant; (iii) quantitatively evaluated the performance of the method from the point of view of clustering; and (iv) examined the motivated nature of the terminology of documentation, contrasting it with the combined terminologies of artificial intelligence and documentation.

We have shown that, through a relevant definition of terminological structure and a suitable method, it is possible to explore the terminological space bound by the conceptual system and represented by a set of terms, whose characteristics are elegantly postulated by Rey (1995):

To the extent that a terminological system, even if it matches a coherent conceptual system, is incapable of reflecting its internal relationships, terminology is autonomous with respect to epistemology. To the extent that a terminological system, even if it is formed from unmotivated and common language words, denotes a conceptual system and exists only for denoting it, terminology is autonomous with respect to linguistics.

The Potts spin glass-based model provides us with a means to explore this terminological structure in such a way that we can interpret the results in terms of the motivatedness of the conceptual-linguistic systems. We are currently carrying out in-depth analyses of the motivated structure of terminologies of several domains, with the conceptual nature and status of constituent elements also under consideration.

So far we have deliberately distanced ourselves from placing this work within the framework of clustering applications, for reasons which include that already discussed in section 3.2. This does not mean that what we have introduced in this paper is not relevant for practical applications. One possible application is in the selection of headwords for terminological lexicons. Of particular importance in this process is maintaining the coherency of the selected headwords as a set. This process depends to a great extent on the sense of the lexicographer, and little formalisation has been carried out so far. No automatic methods have been successful, either. Given this situation, it would be useful if the editors could check the structure and consistency of headwords en masse through explorative means. What we have discussed in this paper could potentially contribute in this regard.

References


