

Clinical Ontology Mapping

Toward Automatic Care Plan Recommendation

Khai Nguyen¹, Kaisei Reio² and Ryutaro Ichise¹

¹National Institute of Informatics, Tokyo, Japan

²Metaps Inc., Tokyo, Japan

Keywords: Clinical Ontology, Ontology Mapping, Care Plan Recommendation, MDS, ICF.

Abstract: In this paper, we share a sketch of an automatic care plan recommendation system in Japan. After that, we describe our proposed method and experience in the first step: clinical ontology mapping. We discuss the difficulties, method, preliminary results of a case study, which is to find corresponding mappings between two ontologies, the Minimum Data Set 3.0 (MDS)^a and the International Classification of Functioning, Disability and Health (ICF)^b.

1 INTRODUCTION

Personalized health care is being developed as a focus of modern healthcare systems. The motivation of such development is the superiority of providing caring solutions meeting the specific needs of an individual. For years in Japan, personalized care plans have been helping numerous people, mainly disease suffering patients, elders, and the disabled. These plans include the short-term and long-term support for persons having difficulty in daily activities. One of existing problems is that the care plan personalization is a complicated process due to the wide range of involved factors, from personal screening results to inter-personal relationships, together with environments, technological resources, etc. Therefore, it costs a lot of effort and time to make a care plan manually.

1.1 Automatic Care Plan Recommendation

For above reasons, developing automatic methods to support the care plan making is an important study. With an automatic system, it takes less time for the analyzer to design/determine a care plan. Currently, in the creation of care plan, an assessment sheet is given as one of required input. This sheet includes health screening results (e.g., body functions, dis-

ease diagnoses, medical history), personal situations (e.g., age, working status, financial capability), relationships (e.g., family, friends), environment factors (e.g., living condition, local policies), etc. A care plan analyzer will check the assessment sheet to identify the capabilities, difficulties, and needs of the person. Together with considerations of available supportive resources (e.g., technology, device, therapist, nursing), a care plan is suggested. We focus on automatic care plan recommendation from assessment sheets and supportive resources. Although the current research object is considered in the context of Japan, our long-term goal is a system widely applicable to different societies.

On the development of care plan recommendation system, we focus on methods whose model and result are interpretable. This is a conventional approach for most tasks whose interpretability of results is an essential requirement. One challenge of care plan recommendation is that the assessment sheets are not written in a well-structured format (e.g., semi-structured and free description). They are written using a collection of vocabularies derived from the Minimum Data Set (MDS). Although MDS is also organized as an ontology, its structure is shallow. Therefore, it is difficult to extract the information from those materials and thus, difficult to build a care plan recommendation system. The recent achievements in machine learning have demonstrated the capability of knowledge generalization from shallow-structured and unstructured data (LeCun et al., 2015; Deng et al., 2014). Applying this approach to our problem might

^a <https://www.cms.gov/>

^b <http://www.who.int/classifications/icf/en/>

be feasible. However, interpretability is a current drawback of many advanced machine learning methods.

To achieve the interpretability, we attempt to divide our study into two phases. In the first phase, we transform assessment sheets into a better-structured format. In the second phase, over the transformed data, we built an interpretable care plan recommender, such as a reasoner. Among various structured representations, ontology-based format is the best candidate due to its capability of knowledge representation and availability of many clinical ontologies. For example, most of up-to-date available ontologies are accessible via BioPortal¹ and UMLS meta-thesaurus². Before transformation of assessment sheets, the mappings of currently in-use ontology and the target higher quality ontologies are necessary. Therefore, ontology mapping is designed as the first most step. The remaining part of this paper describes the details and related issues of clinical ontology mapping.

1.2 Clinical Ontology Mapping

Ontology mapping is a well-studied problem as there are many independently created ontologies, unified ontology has become a need. Various ontology mapping systems have been proposed and have obtained certain achievements (Shvaiko and Euzenat, 2013). However, it is not trivial to apply an existing ontology mapping system to domain specified ontologies, such as clinical ontologies. General ontologies are conceptualized but clinical ontologies are not. That is, an item (i.e., class) in a general ontology is usually represented by a concept (e.g., ‘animal’ and ‘bird’). Meanwhile, an item in a clinical ontology is different because it is used to represent a clinical term, which may not be a single concept. For example, an item could be a functioning (e.g., ‘maintaining a sitting position’) or even a question (e.g., ‘does the president walk?’). Therefore, the mapping of clinical ontologies demands some specific techniques.

Years of effort in automatic clinical ontology mapping have resulted in a modest number of studies and leave many challenges. Previous studies focused on the recommendation of mappings rather than a fully automatic system. One reason for the missing of a fully automatic solution is the requirement of high-quality mappings in the domain. That is also why existing mappings between clinical ontologies have been done manually or semi-automatically.

(Fung and Bodenreider, 2005) attempted to use synonyms, mapping relations of UMLS concepts

¹<https://bioportal.bioontology.org/>

²<https://www.nlm.nih.gov/research/umls/>

to map SNOMED-CT³ and ICD9CM⁴. 86% of SNOMED-CT terms were mapped with 42% recall and 20% precision, compared to the gold standard. (Lomax and McCray, 2004) tried to map Gene Ontology⁵ to UMLS concepts using an automatic method together with manual curation. Such automatic method could suggest the mappings for 25% of Gene Ontology terms.

A common issue of previous studies is that they used simple lexical string matchings or synonym vocabularies. Such approaches do not capture semantic similarity of terms with complex differences. Moreover, they failed to suggest mappings for items which do not lexically match with any term in the target ontology or not exist in the vocabularies. We try to develop a mapping method capable of generating high-quality mapping candidates for any input.

1.3 Candidate Ontologies

As earlier mentioned, we currently use vocabularies derived from MDS for assessment sheets. However, the structural information of MDS is limited. Therefore, we aim to map MDS ontology to other better-structured ontology. In addition to structure, the coverage and recognition of target ontologies are important criteria for selection. Among available ontologies, there are two promising candidates: International Classification of Functioning, Disability and Health (ICF) and SNOMED Clinical Terms (SNOMED-CT).

One advantage of ICF is that it contains terminology of components similar to items in current assessment sheets. Those components are body function, body structure, environment, and activity capability. Furthermore, ICF is recognized by 191 member states of the World Health Organization. Therefore, by using ICF, it is easier to apply our solution in other countries in future. In the meantime, SNOMED-CT is considered as the most comprehensive terminology up to date. SNOMED-CT contains more than 320,000 concepts and has better coverage in specific areas compared to ICF (Tu et al., 2015). This ontology is being used in more than 50 countries.

There are already some mappings between SNOMED-CT and MDS (Tuttle et al., 2007), SNOMED-CT and ICF⁶. However, the mappings between MDS and ICF derived from those data is

³<http://www.snomed.org/snomed-ct/>

⁴<https://www.cdc.gov/nchs/icd/icd9cm.htm>

⁵<http://www.geneontology.org/>

⁶<https://bioportal.bioontology.org/ontologies/ICF/?p=mappings>

limited to 9 pairs. Therefore, currently, it is necessary to find more mappings between MDS and ICF.

Although it is feasible to manually map the medium-sized ontologies such like MDS and ICF, a support tool is helpful in future mapping tasks. For example, when checking and implementing undiscovered mappings between MDS and SNOMED-CT or other ontologies, or dealing with systems using other terminologies. The detail of our mapping method is described in the next section.

2 MAPPING MDS AND ICF - A CASE STUDY

2.1 MDS 3.0 Ontology

MDS is an ontology designed for nursing home screening. It is structured into 25 sections, of which three (Calc, Control, Filter) are for information related to the screening system. The 22 remaining ones are assessment sections. Each section differs from others by its unique screening purpose. For example, section 'C' is about 'cognitive patterns', section 'D' is about 'mood'. Each section contains a flat list of assessment items. That is, the structural information is limited to only sibling and section-item relations.

In total, there are 1038 assessment items. Each item comes with string title, and only a few contain a short description. A string title could be a concept (e.g., disease names), a sentence (e.g., activity descriptions), or even a screening question (i.e., pre-defined interview questions). There are six types of item: text (e.g., social security number), date (e.g., birthdate), ICD code (International Classification of Disease⁷), number (e.g., height), checklist (yes/no), and code (e.g., code '1' refers to male, code '2' refers to female).

2.2 ICF Ontology

Different from MDS, ICF ontology is much better structured. From bottom to the top level is a path starting with component, going through a chapter, block, categories, and ending at a category. That is, each category may contain sub-categories and such relation is recursively defined. In a total of 1530 items, there are four components (i.e., body functions, body structures, activities and participation, and environmental factors) and 30 chapters, 37 blocks, and 1442 categories. Every item comes with a string title and a short description. A string title could be a concept

⁷<http://www.who.int/classifications/icd/>

(e.g., 'mobility' and 'light sensitivity') or a phrase (e.g., 'Lifting and carrying objects').

When mapping to ICF, only 846 MDS items of code and checklist type are used for mapping because ICF does not contain items of other types. However, all of 1530 ICF items are used because every of them can be matched.

2.3 Mapping Method

Given a MDS item x , we first estimate the similarity $sim(x, y)$ between x and every ICF item y . After that, we find the matching score $score(x, y)$ and apply a ranking technique based on those scores to identify most promising mapping candidates. For mapping all of 846 MDS items, the basic idea to perform the above procedure for each item. In other words, all pairwise mappings are examined. Because the number of mappings is about 1.3 million, considerably medium, it is possible to check all those mappings without applying advanced candidate generation techniques. The details of similarity, matching score, and ranking is described in the following subsections.

2.3.1 Similarity

The similarity of two items is estimated by comparing their string titles. Previous clinical ontology mapping systems faced the difficulty of similarity estimation for non-lexically related strings. Currently, it is possible to solve this issues thanks to the achievement of word embedding (Mikolov et al., 2013), which was empirically evaluated as effective methods for many NLP tasks, including string matching. By using word embedding, each string token is represented by a numerical vector and instead of comparing the word surfaces or using synonym dictionaries, the string matching is done by comparing the vectors.

There is some effort for clinical word embedding (Choi et al., 2016; Pyysalo et al., 2013). Among them, (Pyyalo et al., 2013) provide the vectors trained from millions abstract and articles of PubMed⁸ and PMC⁹. We also use this resource for extracting the word vectors.

Given an item z , we first extract all tokens in the title of item z (with stopword removal) and query their word vectors. Then, we calculate the average vector $v(z)$ to represent the title of z . Using the average vector, we estimate the similarity of two item x and y using the following equation.

$$sim(x, y) = \frac{1}{\exp[\text{EuclideanDistance}(v(x), v(y))]} \quad (1)$$

⁸<https://www.ncbi.nlm.nih.gov/pubmed/>

⁹<https://www.ncbi.nlm.nih.gov/pmc/>

Table 1: Example of relevant mappings.

MDS item	ICF item
B0100: Comatose	b1101: Continuity of consciousness
C0700: Staff asmt mental status: short-term memory OK	b1440: Short-term memory
G0300C: Balance: turning around while walking	d450-d469: Walking and moving
H0400: Bowel continence	b5253: Faecal continence
K0300: Weight loss	b530: Weight maintenance functions

2.3.2 Matching score

Based on the estimated similarities and structural information of ICF, we define the matching score as follows.

$$score(x, y) = \max[sim(x, y), \max_{z \in C(y)} score(x, z)] \quad (2)$$

where $C(y)$ are children items of y . In ICF, an item could contain children and such relation is recursively defined. Therefore, the meaning of above equation is that the matching score of item x and y is not only defined by their similarity or the similarity of x and children of y . It is defined by the similarity of x and all descendants of y . This definition comes from the assumption that if x matches to y , it will match to all ancestors of y . As a result, the matching score of a parent item is not less than any of its children.

2.3.3 Ranking

It is simple to just return the list of ranked ICF items based on their matching scores. The problem with this approach is that the top-ranked items can be very similar to each other. In the context of mapping suggestion, in which experts have to check the results manually, it is important to present the mappings selectively. Instead of only prioritizing mappings of highest scores, we minimize the similarity between top-ranked items. To this end, we implement a well-known re-ranking method in Information retrieval, the Maximal marginal relevance (MMR) (Carbonell and Goldstein, 1998). MMR provides a mechanism to control the balance of the relevance and the diversity of the retrieval results. In the context of our problem, MMR is defined as follows.

$$\text{MMR} \stackrel{def}{=} \operatorname{argmax}_{y \in ICF \setminus S} [\lambda score(x, y) - (1 - \lambda) \max_{z \in S} (strc(y, z))] \quad (3)$$

where S is the selected items from ICF, $strc$ is the similarity between ICF items, and $\lambda \in [0, 1]$ is a control factor (i.e., if $\lambda = 1$ only matching score is considered). By repeating the mechanism of equation (3), we collect K mapping candidates for each item x .

The $strc(y, z)$ measures the structural similarity between ICF items and is as follows.

$$strc(y, z) = \frac{2 \times depth(fcs(y, z))}{depth(y) + depth(z)} \quad (4)$$

where $depth(t)$ is the length of path to t from root item and $fcs(y, z)$ is the first common successor of y and z . Here we reuse the idea of (Wu and Palmer, 1994) due to its capability of capturing structural relevances.

2.4 Result

We implemented our mapping module in Java and tested on a desktop with Core i7 7700K CPU. We set λ to 0.9 after preliminarily checking different values, on an observation of five MDS items. We set $K = 5$ for MMR ranking as it is reasonable for manual checking. The time to complete the mapping process for 846 MDS items to ICF is quite fast, at 127 seconds.

We randomly pick up five MDS sections for checking, resulting in 548 mappings of 97 items. Among them, 43 mappings (9%) of 33 items (34%) are relevant. Table 1 lists random examples of relevant mappings from each selected section. According to that table, our method can detect the semantic equality between items of different terms.

3 DISCUSSION

There are still many aspects need investigations to improve the effectiveness. The first issue in our observation is the discrimination between matched and non-matched mappings. Using current matching score is insufficient to differentiate such cases. In many cases, the matching score of some relevant mappings is lower than those of the non-relevances. We intend to use multiple features to describe the similarity of items. For example, in addition to sim over full titles, we can include the similarity of the same part-of-speech, weighting scheme (e.g., TF-IDF, Google distance), or advanced measures on word vector (Kenter and De Rijke, 2015). By using multiple features, we

expect a possibility of the discrimination between relevant and non-relevant cases.

Making good use of existing mappings in the clinical domain (e.g., SNOMED-CT and other ontologies) is also important. The benefit from existing mappings includes detailed evaluations and characterization of mappings. In other words, it is feasible to generalize the mapping knowledge from a gold standard and apply the knowledge on different mapping tasks.

It is possible to compare all items of MDS and ICF exhaustively. However, for larger ontologies such as SNOMED-CT, such approach will be obstructed by computational issues. An investigation on candidate generation for clinical ontology mapping is also a need. It is also worthy to find a mapping method leveraging structural information when two ontologies are both well-structured.

As clinical ontology mapping requires curation by domain experts, utilizing checking results by live-feedback mechanisms is useful. A semi-supervised mapping method will help to improve the mapping quality.

Lastly, multiple ontologies are preferable because it increase the coverage of assessment sheets. However, such advantage comes with challenges, including ontology merging and deduplication. Furthermore, even when using more ontologies, it is not guaranteed a 100% coverage because assessment sheets are customized for a local society. Such problem happen when we work on other health care systems as well. Therefore, it is important to study how to define new items with sufficient logical relations.

4 CONCLUSION

We described an ongoing research project in care plan recommendation system in Japan and its preliminary step, which is the mapping of clinical ontologies. We proposed a straightforward yet reasonable method for mapping MDS and ICF ontologies. Although there are still many challenges to overcome as discussed, we envision an optimistic form of the direction. We hope our study could contribute to improving the quality of healthcare in different societies.

ACKNOWLEDGMENTS

This research is funded by Welmo inc.

REFERENCES

- Carbonell, J. and Goldstein, J. (1998). The Use of MMR, Diversity-based Reranking for Reordering Documents and Producing Summaries. In *Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 335–336.
- Choi, Y., Chiu, C. Y.-I., and Sontag, D. (2016). Learning Low-Dimensional Representations of Medical Concepts. *AMIA Summits on Translational Science*, 2016:41–50.
- Deng, L., Yu, D., et al. (2014). Deep learning: Methods and applications. *Foundations and Trends® in Signal Processing*, 7(3–4):197–387.
- Fung, K. W. and Bodenreider, O. (2005). Utilizing the UMLS for Semantic Mapping Between Terminologies. In *AMIA Annual Symposium*, volume 2005, pages 266–270.
- Kenter, T. and De Rijke, M. (2015). Short Text Similarity with Word Embeddings. In *ACM International Conference on Information and Knowledge Management*, pages 1411–1420.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature*, 521(7553):436–444.
- Lomax, J. and McCray, A. T. (2004). Mapping the Gene Ontology into the Unified Medical Language System. *Comparative and Functional Genomics*, 5(4):354–361.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv preprint arXiv:1301.3781*.
- Pyysalo, S., Ginter, F., Moen, H., Salakoski, T., and Ananiadou, S. (2013). Distributional Semantics Resources for Biomedical Text Processing. In *International Symposium on Languages in Biology and Medicine*, pages 39–43.
- Shvaiko, P. and Euzenat, J. (2013). Ontology Matching: State of the art and Future Challenges. *IEEE Transactions on Knowledge and Data Engineering*, 25(1):158–176.
- Tu, S. W., Nyulas, C. I., Tudorache, T., and Musen, M. A. (2015). A Method to Compare ICF and SNOMED CT for Coverage of US Social Security Administrations Disability Listing Criteria. In *AMIA Annual Symposium*, volume 2015, pages 1224–1233.
- Tuttle, M. S., Weida, T., White, T., and Harvell, J. (2007). Standardizing the MDS with LOINC® and vocabulary matches. Report, at U.S. Department of Health and Human Services.
- Wu, Z. and Palmer, M. (1994). Verbs Semantics and Lexical Selection. In *Annual meeting on Association for Computational Linguistics*, pages 133–138.