X-UDeKAM: An Intelligent Method for Acquiring Declarative Structured Knowledge using Chatterbot

Maqbool Ali, Jamil Hussain, Sungyoung Lee*
Department of Computer Science and Engineering
Kyung Hee University
Yongin, South Korea
{maqbool.ali, jamil, sylee}@oslab.khu.ac.kr

Byeong Ho Kang
School of Engineering and ICT
University of Tasmania
Hobart, Australia
byeong.kang@utas.edu.au

Abstract—**Declarative knowledge has a key role in health-care domain, which is hidden in multiple clinical-related knowledge resources. For acquiring and constructing the declarative structured knowledge from unstructured knowledge resources, *UDeKAM* methodology was proposed that lacks the intelligent human dialoguing mechanism for validating and filtering the important concept relations. This paper is the extension of our previous work and presents an extended version of UDeKAM, called *X-UDeKAM*, that includes *Chatterbot* along with *text preprocessing* as well as *controlled natural language processing* techniques. For realization of X-UDeKAM, a diabetes scenario is explained through example.**

Keywords—*declarative knowledge; text processing; controlled natural language, chatterbot; unstructured knowledge;*

I. INTRODUCTION

Knowledge is a key to compete that solves real-world problems and provides help in decision making. Declarative knowledge is a category of knowledge and expressed in the form of unstructured sentence. In health-care domain, there exists plenty of declarative knowledge which is very critical for quality of health management. Acquiring knowledge from unstructured knowledge resources is one of the key research area, where multiple systems exist that have been proposed for particular applications without support of controlled natural language and intelligent human dialoguing mechanism [1-4]. Keeping in view these facts, an extended version of UDeKAM [1], called *X-UDeKAM* is proposed. The architecture of the proposed methodology is depicted in Fig. 1, where *text preprocessing, Chatterbot,* and model construction using *controlled natural language* phases are elaborated.

II. METHOD

A. Text Preprocessing

The first step is to collect the domain resources i.e. *diabetes* documents. After collection process, this phase applies basic *Natural Language Processing* (NLP) techniques [4] and extracts the *concept relations*. For example, following are the two relations of word '*symptom*' that are extracted from diabetes standard guidelines:

1. symptom/feeling/somesthesia/unpleasant_person/negative_stimulus/hurt
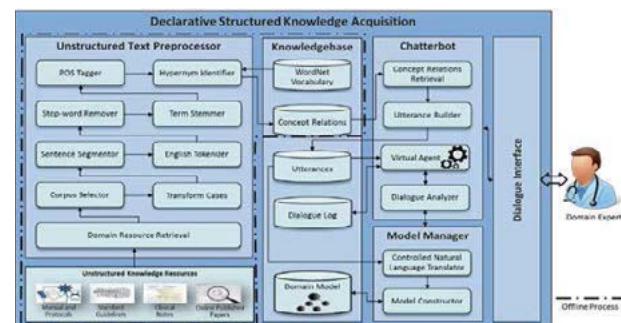2. symptom/blood_disease

Fig. 1.   X-UDeKAM architecture.

B. Chatterbot

A chatterbot is an intelligent computer program, which helps to simulate human conversation via textual methods [5]. Once concept relations are obtained, then there is a need for validation and filtration of important relations. This phase utilized intelligent dialoguing mechanism with the help of *Artificial Intelligence Markup Language* (AIML). For example, for validation of "blood_disease" relation, here is a very simple conversion between human and virtual agent is:

```
<topic name="diabetes-symptoms">
<category>
    <pattern>What are symptoms of diabetes patient ?</pattern>
    <template>Possible symptoms are <bot name="symptom"/>. </template>
</category>
<category>
    <pattern>Is blood_disease a symptom of diabetes?</pattern>
    <template>Yes. </template>
</category>
</topic>
```

C. Model Construction

After validation of concept relations, there is a need to construct domain model, which is consist of declarative knowledge. This phase used *Attempto Controlled English* (ACE) and OWL for constructing declarative structured knowledge [6, 7], which is partially shown in Fig. 2.
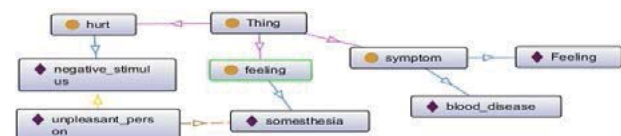

Fig. 2.   A partial view of domain model.

III. Conclusion

This paper demonstrated an extended version of declarative structured knowledge acquisition methodology with support of intelligent human dialoguing mechanism.

Acknowledgment

References

[1] M. Ali, S.Y. Lee, and B.H.Kang, "UDeKAM: A Methodology for Acquiring Declarative Structured Knowledge from Unstructured Knowledge Resources," Proceedings of the 2016 International Conference on Machine Learning and Cybernetics, Jeju, South Korea, 10-13 July, in press.

[2] A. Houser, "FrameMaker: Structured or Unstructured?", http://www.writersua.com/articles/frame/, Group Wellesley, 2004.

[3] R. Jindal and S. Taneja, "U-STRUCT: A Framework for Conversion of Unstructured Text Documents into Structured Form", In Advances in Computing, Communication, and Control, 2013, pp. 59-69.

[4] C. Friedman, L. Shagina, Y. Lussier, and G. Hripcsak, "Automated encoding of clinical documents based on natural language processing", Journal of the American Medical Informatics Association, vol. 11, no. 5, 2004, pp. 392-402.

[5] F.A. Mikic, et al., "CHARLIE: An AIML-based Chatterbot which Works as an Interface among INES and Humans", EAEEIE Annual Conference, 2009, pp. 1-6.

[6] K. Kaljurand, "ACE View—an Ontology and Rule Editor based on Attempto Controlled English", InOWLED, Oct 2008.

[7] R. Denaux, "Intuitive ontology authoring using controlled natural language", Ph.D. dissertation, School of Compt., Uni. of Leeds, 2013.