

Conversational Knowledge Teaching Agent that Uses a Knowledge Base

Kyusong LEE, Paul Hongsuck SEO, Junhwi CHOI, Sangjun KOO, Gary Geunbae LEE

Department of Computer Science and Engineering,

Pohang University of Science and Technology, South Korea

{Kysonglee, hsseo, chasunee, giantpanda, gblee}@postech.ac.kr

Abstract

When implementing a conversational educational teaching agent, user-intent understanding and dialog management in a dialog system are not sufficient to give users educational information. In this paper, we propose a conversational educational teaching agent that gives users some educational information or triggers interests on educational contents. The proposed system not only converses with a user but also answer questions that the user asked or asks some educational questions by integrating a dialog system with a knowledge base. We used the Wikipedia corpus to learn the weights between two entities and embedding of properties to calculate similarities for the selection of system questions and answers.

1 Introduction

Dialog is the most natural interaction between a mentor and mentee in the real world. Therefore, dialog-based intelligent tutoring systems (ITSS) have been widely studied to teach science (Jordan et al., 2013; Litman and Silliman, 2004; Graesser et al., 2004; VanLehn et al., 2002; Vanlehn et al., 2005), foreign language (Kyusong et al., 2014; Lee et al., 2010; Lee et al., 2011; Johnson et al., 2007), and programming language (Fossati et al., 2008; Lane and VanLehn, 2015) usually without intervention from a human teacher. However, previous dialog-based language learning systems mostly only play the role of a conversational partner using chatting like spoken dialog technology, and providing feedback such as grammatical error correction and suggesting better expressions.

However, in real situations, students usually ask many questions to indulge their curiosity and a tutor also asks questions to continue the conversation and maintain students' interest during the learning process. In science and programming language learning, mostly pre-designed scenarios and contents are necessary; these are usually handcrafted by human education experts. However, this process is expensive and time-consuming.

Our group is currently involved in a project called POSTECH Immersive English Study (POMY). The program allows users to exercise their visual, aural and tactile senses to receive a full immersion experience to develop into independent EFL learners and to increase their memory and concentration abilities to the greatest extent (Kyusong Lee et al., 2014). During field tests, we found that many advanced students asked questions that cannot be answered using only a dialog system¹. Recently, knowledge base (KB) data such as freebase and DBpedia have become publicly available. Using the KB, knowledge base question answering (KB-QA) has been studied (Berant and Liang, 2014); it has advantages of very high precision because it exploits huge databases. Hence, we proposed a dialog-based intelligent tutoring system that uses a KB, as an extension of POMY, POMY Intelligent Tutoring System (POMY-ITS). The main advantage is that the human cost to manually construct educational contents is eliminated. Moreover, the system chooses its response after considering information importance, current discourse, relative weights between two entities, and property similarity. The additional functions of the POMY-ITS are that it:

- 1) Answers user's question such as factoid questions, word meaning;

¹ http://isoft.postech.ac.kr/research/language_learning/db-call/videos/e3-1.mp4

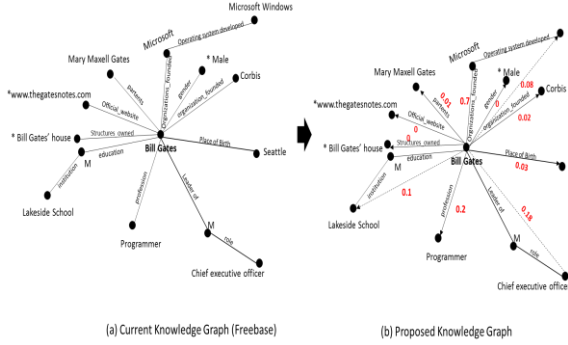


Figure 1: current knowledge graph is undirected graph and proposed knowledge graph is directed weighted graph. (* denote the weight is 0 which meant the tutor never asked about this question)

- 2) Generates questions to continue the conversation and to interest the user;
- 3) Uses entities and properties in freebase to generate useful information that might interest a user, and presents it in natural language.

To implement 1) the QA function, we used Parasempre (Berant and Liang, 2014) based KB-QA system as our QA system. However, in this paper, we focus only on 2) and 3) which are generating questions or informing by selecting appropriate entity and property in the KB; we do not present the detailed explanation or assess the accuracy of the QA system.

2 Intuition of the system

A user who asks about Bill Gates, may also be interested in Microsoft and Paul Allen, which are topics strongly related to Bill Gates. In the KB graph, the ‘Bill Gates’ entity is connected to many other entities. However, these connections present too much information, such as URLs of related websites, gender of Bill Gates, published books, music, and architecture. However, KB does not contain the entity importance or weighted relationship between entities and properties (Figure 1). This information can be useful to POMY-ITS to enable it to decide what to ask or talk about. When a system and a user are talking about Bill Gates’

Algorithm 1 : RuleBasedDA (U, S_{i-1})

- Require:** U : user utterance
Require: S_{i-1} : previous system action
- 1: **if** U contains WH questions **and** $IsEntity(U)$
 - 2: **then** $DA = U: Question$
 - 3: **else if** S_{i-1} is $S: Question$
 - 4: **then** $DA = U: Answer$
 - 5: **else**
 - 6: **then** $DA = U: others$
-

Algorithm 1: Generation Algorithm, $IsEntity$ returns true is when entity is detected in user utterance

Table 1: Example dialog and user dialog act and system action (S:system, U:user)

Utterance	Dialog Act
U:Hi, nice to meet you.	U:others
S:Hello, good to see you.	Matched Example
U:Who is Bill Gates?	U:question
S:Bill Gates is organization learner and programmer.	S:Answer
S:Do you know what company Bill Gates founded?	S:Question
U:Microsoft	U:answer
S: That’s right.	S:CheckAnswer
S: Bill Gates founded Microsoft with Paul Allen	S: Inform

wife’s name, the user may also want to know when they got married or who Bill Gates’ other family members are. Manual construction of the entity relationship or order of scenarios would be very expensive. Our system considers entity and property to decide automatically what to ask or to inform. To deploy the system, we used the Wikipedia corpus to learn property similarity, and weight between two entity pairs.

3 Method

The main role of the POMY-ITS is to give information that a user wants to know. The KB-QA technology will give the answer if the utterance is a ‘wh’ question, but often, a user does not know what to ask. Thus, the conversation must include initiative dialog. When the dialog between a tutor and a user stalls, the tutor should ask a relevant question to or give useful information related to the current context.

3.1 The Role of Dialog Management

First, the system should know whether a user utterance is a question, an answer, or has some other function (Algorithm 1). If the user utterance is a question, KB-QA will answer. If the utterance is an answer, the system will check whether or not the user utterance is correct. Otherwise, we used the example based dialog system which uses a similarity measure to find an example sentence in the example DB (Nio et al., 2014), and utters the sentence (Table 1). The following are the system actions such as *Answer*, *Question (entity, property)*, *Inform (entity, property, obj)*, *Check-UserAnswer*. To generate the next system utterance, we should select arguments such as entity, property, and object. For example,

- *Question (entity=”Bill Gates”, property=”organization.founded”)* will generate “Do you

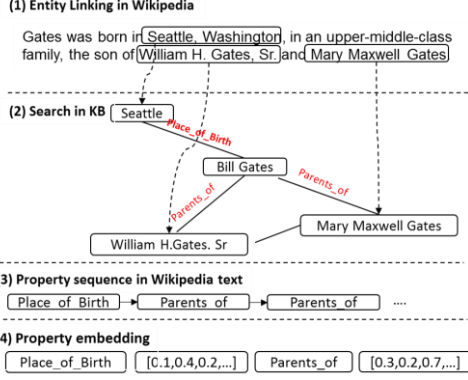


Figure 2: Procedure of property embedding

want to know the company Bill Gates founded?”

- *Inform(entity="Bill Gates", property="organization.founded", obj="Microsoft")* will generate “Bill Gates founded Microsoft”

In this paper, we mainly explore how to select the most appropriate entity and property for generating system utterances.

3.2 Weight between two entities

Freebase is stored in a graph structure. The entity ‘Bill Gates’ is linked to many properties and entities in ‘triple’ format. However, the edges are not weighted. When the system provides useful information to a user about Bill Gates, then his profession, or books that he wrote will be more interesting to a user than Gates’ gender or URL information. Moreover, the relationship between two entities can be represented as a directional graph. When we explain about Bill Gates, Basic programming language is important because he used it when he was programming. However, when we explain about Basic programming language, Bill Gates is not very important. Entities in Wikipedia are linked (Mendes et al., 2011) to obtain the weight information. Weight $w(v_t, v_j)$ is obtained as the follows when v_t is ‘Bill Gates’ and v_j is ‘Microsoft’; First, we need the number of occurrence of “Microsoft” entity in the “Bill Gates” Wikipedia page to get $Freq(v_j)_{v_t}$. Second, we search the shortest path from “Bill Gates” to “Microsoft” in Freebase KB graph, then count the number of properties to get $n(v_t, v_j)$.

$$w(v_t, v_j) = \alpha \frac{Freq(v_j)_{v_t}}{\sum_{v_k \in V_t} Freq(v_k)_{v_t}} + \beta \frac{1}{n(v_t, v_j)} \quad (1)$$

$Freq(v_j)_{v_t}$ denotes frequency of v_j in Wikipedia v_t page. $\forall V_t$ denotes all entities in the Wikipedia v_t page. $n(v_t, v_j)$ denotes # of hops between v_t and v_j (e.g., $n(\text{Bill Gates}, \text{Microsoft}) = 1$, $n(\text{Bill$

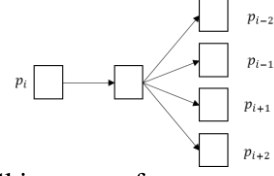


Figure 3 Skip-gram of property embedding

Gates, Microsoft Windows) = 2 in Figure 1-(a)) We eliminate edges that have $w(v_t, v_j) = 0$ and nodes where $n(v_t, v_j) > 2$ (a ‘more than 3 hop’ relationship). α and β are currently set to 1.

3.3 Property Embedding

The intuition of property-embedding similarity is as follows: when a user is talking about Bill Gates’ professional achievement, POMY-ITS’s best option would be to explain something related to professional achievement. However, designing all possible replies manually would be too expensive. When a user asks about Bill Gates’ parents, POMY-ITS’s best option would be to explain or ask the user about Gates’ other family members. To determine that the “people.person.parents” property is more similar to “people.person.children” than “people.person.employment_history” (Figure 5), property-embedding vectors are generated to compute the similarity between two properties. We first obtain the sequence of the property from the Wikipedia corpus (Figure 2), then we use Skip-gram to train the vectors (Figure 3). The training objective of the Skip-gram model is to find word representations that are useful to predict the surrounding (Mikolov et al., 2013). We used skip-gram to predict the next property r given the current property as the following equation:

$$\frac{1}{T} \sum_{t=1}^T \sum_{-2 < j < 2, j \neq 0} \log p(r_{t+j} | r_t) \quad (2)$$

where r_t denotes current property. The basic Skip-gram formulation uses the soft-max function to define $p(r_{t+j} | r_t)$:

$$p(r_o | r_i) = \frac{\exp(v'_{r_o} v_{r_i})}{\sum_{r=1}^R \exp(v'_{r_o} v_{r_i})} \quad (3)$$

where v_r and v'_r are, respectively, the input and output vector representations of r , and R is the number of properties in Freebase.

3.4 System Utterance Generation

After choosing entity and property, we can generate either *question* or *inform* sentences. Template-based natural language generation uses rules (Table 2) to generate question utterances. Questions begin with a question word, are followed by the

Freebase description of the expected answer type $d(t)$, the further followed by Freebase descriptions of entities $d(e)$ and $d(p)$. To fill in auxiliary verbs, determiners, and prepositions, we parse the description $d(p)$ into one of NP, VP, PP, or NP VP. For *inform* system actions, we generate the sentences from triple \langle Bill Gates, organization.founded, Microsoft \rangle to “Bill Gates founded Microsoft” as follows: extract the triple from the text, and disambiguate to KB entities. Then, align to existing triples in KB, fourth. Finally, collect matched phrase-property pairs from aligned triples.

Table 2: Template of questioning. WH represents “Do you know what”.

Rule	Example
WH $d(t)$ has $d(e)$ as NP?	WH election contest has George Bush as winner?
WH $d(t)$ (AUX) VP $d(e)$?	WH radio station serves area New-York?
WH PP $d(e)$?	WH beer from region Argentina?
WH $d(t)$ VP the NP $d(e)$?	WH mass transportation system served the area Berlin?

3.5 Experiment and Result

To compare the weight of two entities, 10 human experts ranked among the 60 entities that were most closely related to the target entity. We asked them to rank the entities as if they were teaching students about the target entities such as “Bill Gates”, “Steve Jobs”, “Seoul”, etc. We considered the human labeled rankings to be the correct answers, and compared them to answers provided by the proposed method and word2vec² (Figure 4); as a similarity statistic we used the average score of Mean reciprocal rank (MRR). We obtained MRR scores 10 times, then got mean and standard deviation by repeating one human labels as the answer and another human labels as the test; this allows quantification of the correlation between human labels. The results show that human-to-human has the highest correlation. Next, the correlation between human and the proposed method is significantly better than between human and word2vec (Figure 4). We found that word2vec has high similarity when entities are of the same type; e.g., Melinda Gates, Steve Ballmer, and Jeff are all “person” in Table 3. However, humans and the proposed system selected entities of different types such as ‘Microsoft’ and “Windows”. Thus, semantic similarity does not necessarily represent the most related entities for explanation about the target entity in the educational perspective. To show property similarity, we plot in the 2D space using t-SNE (Van der Maaten and Hinton, 2008).

² The model of freebase entity embedding is already available in <https://code.google.com/p/word2vec/>

Table 3: Ranked Results of the top 5 entities generated for Bill Gates

Rank	Human	Proposed	Word2Vec
1	Microsoft	Microsoft	Melinda Gates
2	MS Windows	Paul Allen	Steve Ballmer
3	MS-DOS	Harvard Univ.	Bill Melinda Gates Foundation
4	Harvard Univ.	Lakeside School	Feff_Raikes
5	OS/2	CEO	Ray Ozzie

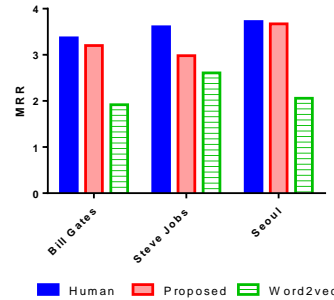


Figure 4: Mean and SD of MRR scores for 10 human labeled rankings

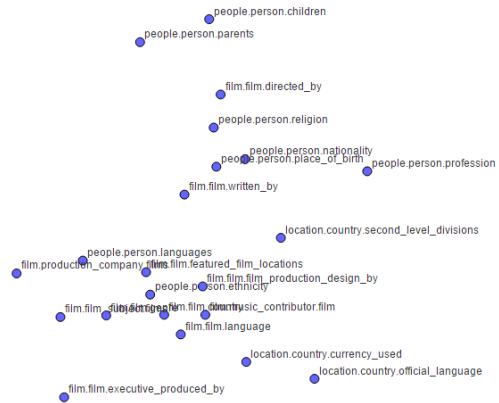


Figure 5: plotting property-embedding vectors

The graph shows that similar properties are closely plotted in 2D space, especially *people.person.children* and *people.person.parents* (Figure 5). This is exactly consistent with our purpose of property-embedding, and our property-embedding model is available³ which includes 779 total properties and 100 dimension.

4 Conclusion

We developed a conversational knowledge-teaching agent using knowledge base for educational purposes. To generate proper system utterance, we obtained the weight between two entities and property similarity. The proposed method significantly improved upon baseline methods. In the future, we will improve our conversational agent for knowledge education more tightly integrated into QA systems and dialog systems.

³ <http://isoft.postech.ac.kr/~kyusonglee/sigdial/p.emb.vec>

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0019523) and was supported by ATC(Advanced Technology Center) Program-“Development of Conversational Q&A Search Framework Based On Linked Data: Project No. 10048448 and was partly supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No. R0101-15-0176, Development of Core Technology for Human-like Self-taught Learning based on a Symbolic Approach)

References

- Jonathan Berant and Percy Liang. 2014. Semantic parsing via paraphrasing. In *Proceedings of ACL*, volume 7, page 92.
- Jenny Brusk, Preben Wik, and Anna Hjalmarsson. 2007. Deal a serious game for call practicing conversational skills in the trade domain. *Proceedings of SLATE 2007*.
- Davide Fossati, Barbara Di Eugenio, Christopher Brown, and Stellan Ohlsson. 2008. Learning linked lists: Experiments with the ilist system. In *Intelligent tutoring systems*, pages 80–89. Springer.
- Arthur C Graesser, Shulan Lu, George Tanner Jackson, Heather Hite Mitchell, Mathew Ventura, Andrew Olney, and Max M Louwerse. 2004. Autotutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers*, 36(2):180–192.
- WLewis Johnson, NingWang, and ShuminWu. 2007. Experience with serious games for learning foreign languages and cultures. In *Proceedings of the SimTecT Conference*.
- Pamela Jordan, Patricia Albacete, Michael J Ford, Sandra Katz, Michael Lipschultz, Diane Litman, Scott Silliman, and Christine Wilson. 2013. Interactive event: The rimac tutor-a simulation of the highly interactive nature of human tutorial dialogue. In *Artificial Intelligence in Education*, pages 928–929. Springer.
- LEE Kyusong, Soo-Ok Kweon, LEE Sungjin, NOH Hyungjong, and Gary Geunbae Lee. 2014. Postech immersive english study (pomy): Dialog-based language learning game. *IEICE TRANSACTIONS on Information and Systems*, 97(7):1830–1841.
- H Chad Lane and Kurt VanLehn. 2005. Teaching the tacit knowledge of programming to novices with natural language tutoring. *Computer Science Education*, 15(3):183–201.
- Sungjin Lee, Hyungjong Noh, Jonghoon Lee, Kyusong Lee, and G Lee. 2010. Postech approaches for dialog-based english conversation tutoring. *Proc. APSIPA ASC*, pages 794–803.
- Sungjin Lee, Hyungjong Noh, Jonghoon Lee, Kyusong Lee, Gary Geunbae Lee, Seongdae Sagong, and Munsang Kim. 2011. On the effectiveness of robotassisted language learning. *ReCALL*, 23(01):25–58.
- Diane J Litman and Scott Silliman. 2004. Itspoke: An intelligent tutoring spoken dialogue system. In *Demonstration papers at HLT-NAACL 2004*, pages 5–8. Association for Computational Linguistics.
- Pablo N Mendes, Max Jakob, Andr es Garc ia-Silva, and Christian Bizer. 2011. Dbpedia spotlight: shedding light on the web of documents. In *Proceedings of the 7th International Conference on Semantic Systems*, pages 1–8. ACM.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Hazel Morton and Mervyn A Jack. 2005. Scenario-based spoken interaction with virtual agents. *Computer Assisted Language Learning*, 18(3):171–191.
- Lasguido Nio, Sakriani Sakti, Graham Neubig, Tomoki Toda, and Satoshi Nakamura. 2014. Improving the robustness of example-based dialog retrieval using recursive neural network paraphrase identification. *SLT 2014*.
- Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(2579-2605):85.
- Kurt VanLehn, PamelaWJordan, Carolyn P Ros e, Dumisizwe Bhembe, Michael B  ottner, Andy Gaydos, Maxim Makatchev, Umarani Pappuswamy, Michael
- Ringenberg, Antonio Roque, et al. 2002. The architecture of why2-atlas: A coach for qualitative physics essay writing. In *Intelligent tutoring systems*, pages 158–167. Springer.
- Kurt Vanlehn, Collin Lynch, Kay Schulze, Joel A Shapiro, Robert Shelby, Linwood Taylor, Don Treacy, Anders Weinstein, and Mary Wintersgill. 2005. The andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education*, 15(3):147–204.