

IDEA: An Interactive Dialogue Translation Demo System Using Furhat Robots

Jinhua Du, Darragh Blake, Longyue Wang, Clare Conran, Declan Mckibben,
and Andy Way

ADAPT Centre, Dublin City University, Dublin, Ireland
{firstname.secondname}@adaptcentre.ie

Abstract. We showcase **IDEA**, an *I*nteractive *D*ialogue *E*trAnslation system using Furhat robots, whose novel contributions are: (i) it is a web service-based application combining translation service, speech recognition service and speech synthesis service; (ii) it is a task-oriented hybrid machine translation system combining statistical and neural machine learning methods for domain-specific named entity (NE) recognition and translation; and (iii) it provides user-friendly interactive interface using Furhat robot with speech input, output, head movement and facial emotions. IDEA is a case-study demo which can efficiently and accurately assist customers and agents in different languages to reach an agreement in a dialogue for the hotel booking.

Keywords: Dialogue translation · Furhat robot · Entity translation

1 Introduction

Applications of machine translation (MT) in many human–human communication scenarios still present many challenges, e.g accurate task-specific entity recognition and translation, ungrammatical word orders in spoken languages etc. Human–human dialogue translation is a more demanding translation task than the general-purpose translation tasks in terms of the recognition and translation of key information in the dialogue, such as the *person* (*who*), *location* (*where*), *time/date* (*when*), and *event* (*what*) etc. An example of the low accuracy of entity recognition and translation in dialogues between customers and hotel agents from public translation systems is shown below:

<i>Source:</i>	我想定一个{30号}的(大床房)。[五点]可以到酒店。
<i>Reference:</i>	I would like to <u>book</u> a (king room) on {the 30th this month}. I will check in at [five o'clock].
<i>Google:</i>	I want to <u>set</u> a {30th} (bed room). [Five o'clock] can go to the hotel.
<i>Baidu:</i>	I'd like to <u>make</u> a (big bed room) for (No. 30). You can get to the hotel at [five].

In this example, we use {}, () and [] to highlight the time/date and room type entities and their corresponding translations; use to highlight the verb of

the event. We can find from this commonly-used dialogue sentence that check-in date, room type and the verb were not translated accurately either by *Google* or *Baidu*. Wrong translations of these key information will impede an effective and efficient communication between the customer and agent.

Accordingly, we carried out a study on task-oriented dialogue machine translation (DMT) with semantics-enhanced NE recognition and translation. As a case study, we developed an interactive DMT demo system for the hotel booking task: **IDEA**, which can assist customers in one language to communicate with hotel agents in another language to reach an agreement in a dialogue. A text input/output-based demo system of IDEA describing the working mechanism can be found at <https://www.youtube.com/watch?v=5KK60gMPDpw&t=5s>.¹

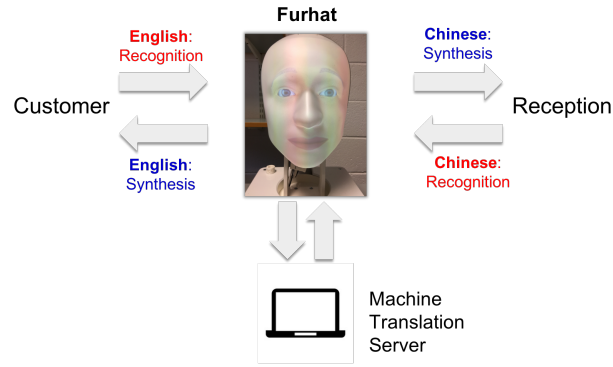


Fig. 1. System architecture of IDEA

2 System Description

2.1 System Architecture and Workflow

The architecture of IDEA is illustrated in Figure 1. In the hotel booking scenario, Furhat robots provide the speech recognition and speech synthesis services between customers and hotel agents. The DMT Server provides a Web Service for the translation of messages between the customer and the agent through Furhat robots. A key component in our demo system is a Semantic Module which combines statistical and neural machine learning methods for the understanding, extraction and candidate translation of key entities/information in the ongoing dialogue, such as “customer name”, “arrival time”, “room type” and so on.

¹ News regarding the Furhat-based DMT demo system can be found at: <https://twitter.com/adaptcentre/status/932957814301044737>.

The text input/output-based interface of IDEA is shown in Figure 2. In the hotel booking scenario, customers and agents speak different languages.² Customers can access the hotel website to request a conversation with an agent. Then the agent accepts the customer’s request to start the conversion. Messages between the customer and agent will be automatically translated to the customer’s or agent’s language, and the semantic information (key entities) is automatically recognised, extracted and translated to achieve the intention of the hotel booking.

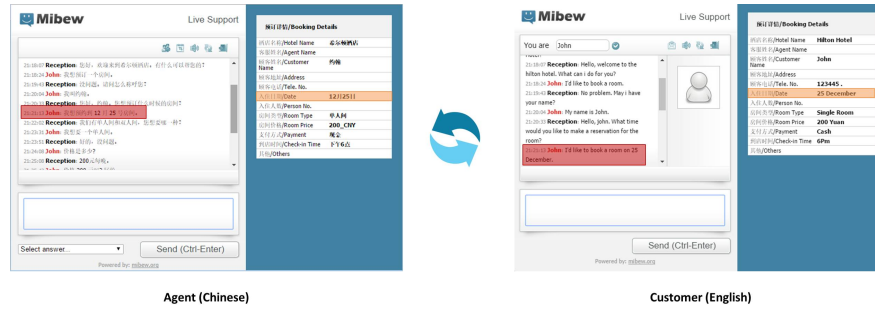


Fig. 2. Interface of IDEA

Figure 3 shows the detailed workflow of IDEA which demonstrates how each module in the system works. In general, customers and agents alternately input texts or speak like a human–human question–answering scenario. We propose a task-oriented semantics-enhanced statistical and neural hybrid method to recognise entities by inferring their specific types based on information such as contexts, speakers etc. Then, the recognised entities will be represented as logical expressions or semantic templates using the grounded semantics module. Thus, the entities can be correctly translated in the contexts. Finally, candidate translations of semantically represented entities will be marked up and fed into a unified bi-directional translation process. Refer to [1] for technical details.

3 Evaluation and Application

We evaluate IDEA on different aspects and results show that: (1) the hybrid NE recognition and translation in the dialogue improves over 20% compared to general-purpose systems regarding the hotel booking scenario; (2) the improved NE translation improves the overall translation performance by over absolute 15% BLEU points on English–Chinese hotel booking scenario; (3) the success rate of booking is improved by over 30% compared to public translation systems.³

² In order to be more understandable, the rest of the paper will assume that customers speak English and agents speak Chinese.

³ If all pre-defined key information such as *room type*, *check-in date*, *room price* etc. can be recognised and translated correctly, then we regard it as a successful dialogue.

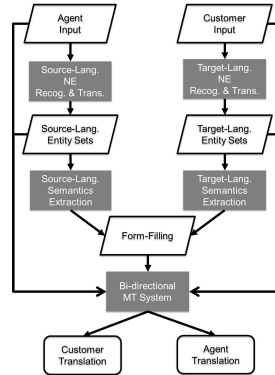


Fig. 3. Workflow of IDEA

The advantages of our task-oriented DMT system include: (1) focusing on the recognition and translation of key information in the dialogue to alleviate misunderstanding; (2) an extra task-oriented semantic module reduces the reliance on large-scale data, and easy to deploy on portable devices; (3) the head movements and facial expressions motivated by speaker’s mood and manner of speaking etc. can increase the engagement; (4) the mode of using Furhat robots for interaction (head movements, facial expressions etc.) can be alternatively implemented as an avatar in mobile applications; (5) the techniques proposed and developed in the hotel booking scenario can be quickly adapted and applied to other task-oriented dialogue translations, such as customer services in finance, telecom, retail etc.

4 Conclusion and Future Work

In this paper, we demonstrated IDEA, a task-oriented dialogue translation system using Furhat robots for a hotel booking scenario. Evaluations on different aspects show that our DMT system can significantly improve translation performance and success rate of the task, and can be easily to extended to different task-oriented translation services.

Acknowledgement. We would like to thank the reviewers for their valuable and constructive comments. This research is supported by the ADAPT Centre for Digital Content Technology, funded under the SFI Research Centres Programme (Grant 13/RC/2106), and by SFI Industry Fellowship Programme 2016 (Grant 16/IFB/4490).

References

1. Wang, L., Du, J., Li, L., Tu, Z., Way, A., Liu Q.: Semantics-Enhanced Task-Oriented Dialogue Translation: A Case Study on Hotel Booking. In: The Companion Volume of the IJCNLP 2017 Proceedings: System Demonstrations, pp. 33–36, Taipei, Taiwan, November 27–December 1, 2017.