The pragmatic combination of different cross-lingual resources for multilingual information services

Hans Uszkoreit[1], **Feiyu Xu**[1], **Jörg Steffen**[1], **Ilhan Aslan**[2]

DFKI GmbH, LT-Lab[1], IUI-Lab[2]
Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany
{uszkoreit, feiyu, steffen, aslan}@dfki.de

Abstract

We will describe new cross-lingual strategies for the development multilingual information services on mobile devices. The novelty of our approach is the intelligent modeling of cross-lingual application domains and the combination of textual translation with speech generation. The final system helps users to speak foreign languages and communicate with the local people in relevant situations, such as restaurant, taxi and emergencies. The advantage of our information services is that they are robust enough for the use in real-world situations. They are developed for the Beijing Olympic Games 2008, where most foreigners will have to rely on translation assistance. Their deployment is foreseen as part of the planned ubiquitous mobile information system of the Olympic Games.

1. Introduction

The application is a bundle of web-based mobile multilingual information services for the visitors of large international events. Since most of the information is not produced in all of the offered languages, cross-lingual technologies need to be employed for overcoming the existing language barriers (Uszkoreit and Xu (2004)). The application scenario is the 2008 Olympic Games in Beijing, a highly demanding setting for a cross-lingual application because of the heterogeneous nature of the information and the involved languages.

If machine translation were reliable and robust enough to support the mapping of all provided information to the languages of the users, one general high-quality MT system would be sufficient to turn monolingual information services into the desired multilingual system. The current state of the art in machine translation, however, does not provide the degree of reliability required for the intended services. When foreign visitors need help in Beijing, they often are in situations in which they do not have the time and patience to cope with faulty or unintelligible information. Therefore our services try to circumvent the fully automatic translation through a general purpose MT system whenever possible.

The three main approaches employed in COMPASS for crossing the language boundaries are multilingual generation, phrasebook translation and free machine translation. For each approach, we exploit existing resources whenever possible. Among such resources are monolingual and bilingual dictionaries, terminologies, tourist phrasebooks, and existing MT systems. Some additional resources that had to be developed were designed in such a way that they could be reused in future applications. An example is ontology for the service features and location objects and for some application specific domains that are designed in RDF schema.

Our task is to embed the multilingual techniques in an information service platform in such a way that they are really useful for the real world situations during Olympic Games 2008. Therefore, each specific translation service is in fact a complex system itself, composing of textual translation, location-based services, speech generation and multimodal interaction.

Intelligent multimodal user interface technologies ensure comfortable access via the small screens of PDAs and smart phones and the limited input options for handheld computers (Aslan et al. (2005a), Wahlster (2003)). Users can choose their input methods for search depending on the situation: via speech or writing and they can decide whether they want to read or hear the found information. The COMPASS system "knows", where a user is located and in which situation he is and exploits this information for providing the best responses. The scientific and technical innovation of the COMPASS2008 system is to combine translation technology, speech input and output, personalization, situation-aware services and an open semantic structure oriented architecture, so that users of COMPASS are supplied anytime and anywhere with the best information in their own language.

In this paper, our focus is the trans-lingual and cross-lingual aspects of the multilingual information services and not the cross-lingual retrieval and question answering techniques already explored in many systems (Braschler and Schäuble (1999), Carbonell et al (1997), MUST (2002), Xu (2003)).

2. Cross-lingual and Translation Techniques

2.1. Multilingual Template-based Generation

For all types of information that can be obtained as structured data or easily converted into records of a conventional relational database, multilingual template-based generation is utilized for producing the requested information in the respective target languages. This method is highly efficient and reliable. A certain degree of stylistic variation is included in order to avoid monotony by frequent repetition (Busemann and Horacek (1998)).

In COMPASS, the input for the template-based generation can be delivered by various components: location-based search based on GPS information, food search, restaurant search or user input. For example, a user query such as "Take me to a Chinese restaurant close to here" will trigger the positioning service and location search for restaurants within the smallest radius. Then a

record will be filled with Chinese restaurant names and their addresses. The multilingual template generation will translate it into Chinese. The corresponding person, in this case, most probably, the taxi driver will obtain the request, which is spoken in Chinese by the automatic speech generation.

2.2. Phrase-book Translation

Whenever utterances in a relevant limited domain need to be translated for communication partners who do not master the source language, we try to use a phrasebook-like specialized translation scheme. In some cases, the user can select the source phrase through a menu, in other cases an input utterance is mapped into the semantically equivalent or closest source phrase. The phrase-book approach has the advantage that the translations are correct. It works because in most situations the choice of possible input utterances is limited by the context. In order to exploit this dependence, we have developed dialogue models for the relevant situation types. Our dialogue models resemble finite-state automata. At each state the user can select a number of utterances in one or several steps. Once an utterance is selected, the user can translate the utterance. The translation can be displayed, spoken, or in some cases, also transmitted to services via phone or SMS. The user can also back off or jump to another situation.

If the situation and device permit the input of typed utterances, these utterances will first be matched against the expected input utterances of the state in the dialogue model. When a natural language utterance is mapped to a source phrase in the phrase book and then translated, the source phrase is displayed for the user as feedback in order to permit the detection of possible misunder-standings.

In some dialogue states, input may contain names of locations, times, dates or currency expressions. In this case, the displayed input option is a sentence template with some open slots that can be filled through special pull-down menus or through typed or penciled input.

2.3. Free Text Translation

Free translation by a rule-based state-of-the-art machine translation system is applied when users want to translate texts at their own risk using the so-called *transearly* service (see 3.1). We have assessed a few translation systems based on availability and on a simple evaluation procedure. Since there do not exist many systems with Chinese as a source or target language, the number of considered systems was small for translations into or out of Chinese. The best system for translations involving Chinese[1] turned out to be a commercial system from Beijing, provided by our Chinese project partners. For other language pairs, we tested the web-based Systran MT Service and the LOGOS Open Source machine translation system. Although the translation quality is in many cases still limited, it is sufficient for covering basic needs of a tourist in Bejing. Especially the HuaJian MT System works well in typical tourism situations, maybe because one of the first applications of the system was a handheld tourist translation assistant.

[1] we have applied the BLEU score evaluation method and subjective confirmation of the BLEU results.

However, even the free text translation service will first check if the input can be mapped into a phrasebook source phrase for increased reliability. The choice of the translation system for the deployed services will depend on licensing policies and the business model for the service provision.

3. COMPASS Translation Center

The COMPASS Translation Center provides a list of translation services that are relevant for the stay of foreigners. It combines the free textual translation mentioned in 2.3 and application-specific translation services:

- **transearly**,, online translation service
- **smart dining**, assistance with restaurant and menu selection and ordering
- **taxi talk**, assistance in dialogues with taxi drivers
- **resc you**, emergeny assistance

Figure 1: COMPASS Service Client

Smart dining and *taxi talk* combine multilingual generation with phrasebook translation. *Resc you* combines phrasebook translation with a free general text translation for selected situations. *Transearly* combines phrasebook translation with free text translation. We argue, why the respective demands and constraints of each of the four services are met by the selected combinations of approaches. All these services can be called by user requests and one service can call another service, for example, the smart dining service can access the transearly service too.

3.1. Transearly

As described in 2.3, our open-domain machine translation service is an integrated web-based online service using all of available open-domain machine translation resources. In the specific Olympic Games scenario, the language pair Chinese and English is very important. The embedded translation system is also able to translate a Chinese sentence into its Pinyin representation. Pinyin is a system of romanization (phonemic notation and transcription to Roman script) for standard Mandarin,

which can be used as pronunciation guide for foreigners. Furthermore, we have integrated the Chinese speech synthesis that can speak for the users.

Furthermore, the *transearly* service is also useful for computational linguistic lectures because of big coverage of language pairs. Our current system supports more than 15 language pairs.

Figure 2: Free textual translation on mobile device

3.2. Smart Dining

China is famous for its unique food culture exhibiting a high degree of culinary diversity and regional variety. The smart dining service helps foreigners to get to know and enjoy this part of the Chinese culture even if they do not understand the language. Smart Dining is an information assistant and at the same time translation service, specialized for restaurants and food. The service is available for small handheld devices (Aslan et al, 2005b). Foreigners can find their preferred restaurants and dishes in their own languages. Different search options will be provided. Users can find restaurants in their neighborhood and select restaurants and dishes according to their tastes. The system knows the restaurants' menus and the ingredients of listed dishes, so that it can help to find the most preferred combinations and avoid discomfort and noxious effects caused by aversions and allergies. Using the integrated translation and speech synthesis, users can let their PDA or smartphone display or pronounce the names of dishes and beverages in Chinese. Thus food ordering will be easy and successful. The food descriptions are accompanied by pictures and background information. This facilitates decision making and communication with the restaurant staff. The attractive user interface and the wealth of instructive information provide valuable assistance as well as some pre-dinner edutainment.

We have developed fine-grained ontology for restaurants, food and beverage. Our ontology is multilingual, supporting three languages. The ontology is used at same time as our multilingual lexicon for translation, in addition to data structuring task.

Concerning food ontology, we are faced with the culture difference between Chinese dinning tradition and western dinning tradition. A simple example is that a Chinese dinner is divided into cold dish, warm dish, soup and dessert, not starter, main course and dessert. In order to help foreigners to find the Chinese dishes easily, we have classified the Chinese dish into the western course additionally. In the following, we give an example of a dish in Chinese and English:

```
<food>
  <food_id id="1"/>
     <name_zh>北京烤鸭</name_zh>
        <name_py>bei jing kao ya</name_py>
        <small_img>food_1_small.jpg</small_img>
        <audio>food_1.mp3</audio>
         <food_zh>
                <name>北京烤鸭</name>
                <food_taste_zh value="咸"/>
                <food_kind_zh value="鸭"/>
                <food_origin_zh value="北京菜"/>
                <food_course_zh value="主菜"/>
                <ingredients>...</ingredients>
                <preparation>..</preparation>
                <eating_method> ...</eating_method>
                <comments>肉质鲜嫩...</comments>
        </food_zh>
        <food_en>
                <name>Beijing Roast Duck – Peking
Duck</name>
                <food_taste_en value="salty" />
                <food_kind_en value="duck"/>
                <food_origin_en value="Beijing"/>
                <food_course_en value="main_course"/>
                <ingredients>...</ingredients>
                <preparation>...</preparation>
        </food_en>
</food>
```

Figure 3: Navigation of restaurant information

In the current system, we have collected more than 300 dishes and 200 restaurants in Beijing. Dish, beverage and restaurant and location information is interlinked.

3.3. Taxi Talk

To be mistakenly dropped by a taxi driver at an unknown place in China, is a true horror scenario for many foreign tourists. Not only during the Olympic Games the communication with a taxi driver belongs to the impossible challenges, if you do not master the Chinese language. The COMPASS2008 taxi-dialog-assistant is the mediator between a taxi driver and the foreigners. The system helps the visitors, to find their destination in their own language. It highlights and pronounces the location name in Chinese, so that the driver can be informed explicitly. Furthermore, the system

will ask the driver for the distance and the prices before the trip. The driver can interact with the mobile device and select the appropriate answers from a selection of choices. The assistant translates the most common requests, such as “please stop here” or “please turn on the air condition”. Of course, the system cannot support a deep intellectual conversation with the Chinese taxi driver. However, our assistant will provide sufficient communication support ranging from ordering a taxi till paying the correct fare. The effective and comfortable assistance is achieved through an intelligent combination of translation, speech generation and an interactive mobile user interface.

We have modeled four dialog states: order taxi, before, during and after transportation. For each state, we have modeled dialogs in three languages in a very detailed way. The template-based generation translates the user query into Chinese. The current *Taxi Talk* supports three languages Chinese, English and German. It is easily to be extended to other languages. In figure 4, you see the phrases relevant for the state “before transportation”:

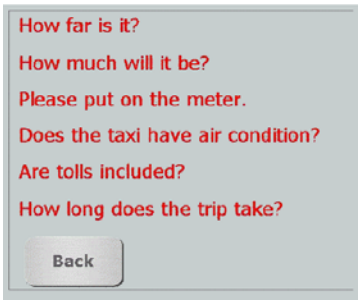

Figure 4: Phrases before taxi transportation

4. Conclusion and Future Work

The described combination of services stands for a wider class of realistic near term applications that incorporate language technology in a conservative but scalable way. Although state of the art machine translation is far from being accurate and robust, machine translation, backed up by other cross-lingual technologies, can be applied in a reliable service.

We are currently extending the application by special services for emergency situations such as medical problems, accidents, loss of valuables or documents, crime and cases of serious conflicts. In these services, combined in the component RescYou, the phrase based translation approach guided by strict dialogue models is followed even more strongly in order to provide the needed reliability.

On the other hand, we will investigate in our upcoming field tests of the entire application in Beijing how useful the free translation service is to the test users and how we can extend and improve this part of the application. The results of the field tests will be interpreted as guidance for the last phase of the project and appropriately published.

To sum up, the COMPASS information services fill a true need for language technology in situations where cross-lingual assistance is badly needed. They combine different types of language technology such as

- simple language resources (ontology resources, bilingual dictionaries and phrase book sample translations)
- more sophisticated language technologies such as CLIR, NEE and speech generation
- existing machine translation systems

with mobile multimodal information services in a novel but well thought through and carefully constrained way. Evaluation will be performed in a field test in Beijing under realistic conditions with test users from several countries.

5. References

Aslan, I., Xu, F., Uszkoreit, H., Krüger, A., Steffen, J. (2005a). COMPASS2008: Multimodal, Multilingual and Crosslingual Interaction for Mobile Tourist Guide Applications. *In Proceedings of Intelligent Technologies for Interactive Entertainment: First International Conference, INTETAIN 200*5, Madonna di Campiglio, Italy, November 30 – December 2, 2005, pp. 3 - 12

Aslan, I., Xu, F., Steffen, J., Uszkoreit. H., Krüger, A. (2005b). COMPASS2008: The Smart Dining Service. *In Proceedings of INTETAIN 2005*, Madonna di Campiglio, Italy, November 30 – December 2, 2005, pp. 299-302.

Braschler, C. P.and Schäuble, P. (1999). Cross-Language Information Retrieval (CLIR) Track Overview. *In Proceedings of the eighth Text REtrieval Conference (TREC-8)*, held in Gaithersburg, Maryland, November 17 – 19, 1999.

Busemann, S., Horacek, H. (1998). A Flexible Shallow Approach to Text Generation. *In: Eduard Hovy: 9th International Natural Language Generation Workshop (INLG '98)*, August 5-7 pp. 238-247, Niagara-on-the-Lake, Ontario, Canada, 1998

Carbonell, J., Yang, Y., Frederking, R., Brown, R. D., Geng, Y., Lee, D. (1997). Translingual Information Retrieval: A comparative evaluation. *In Proceedings of the Fifteeth International Joint Conference on Artificial Intelligence*, August 1997.

Uszkoreit, H., Xu, F. (2004). Modern Multilingual and Crosslingual Information Access Technologies. *In Proceedings of Multilingual Information Service System for the Beijing 2008 Olympics Forum*. The 7th China Beijing International High-Tech Expo (CHITEC), May 21 2004, Beijing, China.

Must (2002). Must — Multimodal and Multilingual Services for Small Mobile Terminals. *Heidelberg, EURESCOM Brochure Series*, May 2002. (http://www.eurescom.de/public/projects/P1100-series/P1104/default.asp)

Wahlster, W. (2003). SmartKom: Symmetric Multimodality in an Adaptive and Reusable Dialogue Shell. *In: Krahl, R., Günther, D. (eds.): Proceedings of the Human Computer Interaction Status Conference 2003*, June 2003, Berlin: DLR, pp. 47 – 62.

Xu, F. (2003). Multilingual WWW — Modern Multilingual and Cross-lingual Information Access Technologies. *In Knowledge-Based Information Retrieval and Filtering from the Web. Witold Abramowicz (Ed.)*, Kluwer Academic Publishers, pp. 165 – 184, 2003.