elsnews 11.2

The Newsletter of the European Network in Human Language Technologies Summer 2002

Lively Discussions Behind Cloister Walls

Impressions from the ISCA Tutorial and Research Workshop on "Multi-modal Dialogue in Mobile Environments", June 17-19, 2002

Andreas Kellner, *Philips Research Labs, Aachen*

The baroque monastery of Kloster Irsee is situated in Bavaria, or more precisely Swabia, about 90km South-East of Munich and is just a short trip away from the famous Neuschwanstein castle. Originally built as a Benedictine cloister in the early 18th century, the monastery complex, which also comprises a wonderful baroque church and a brewery, is now run as a conference hotel, and provided the perfect setting for the ISCA Tutorial and Workshop on *'Multi-Modal Dialogue in Mobile Environments'*. After the successful 1999 Workshop on *Interactive Dialogue in Multi-Modal Systems* (IDS-99), this year's workshop covered various aspects of spoken and multi-modal dialogue systems, with special emphasis on mobile environments, but also highlighted general challenges of multi-modal user interface technology. More than 60 researchers from 15 different countries (including Germany, the Netherlands, Sweden, USA, Japan, and Australia) had come to Bavaria to discuss the latest developments in this field.

The three-day workshop started with a keynote speech by Anthony Jameson from the German Research Institute for Artificial Intelligence (DFKI), who used the example of a speech-controlled cell phone to demonstrate some important usability issues that arise in multi-tasking situations with multi-modal user interfaces. Jameson emphasised that besides technological issues, a deep understanding of user behaviour is crucial for the success of multi-modal interfaces. A live demonstration of a gaze tracker gave an example of the kind of new methodology necessary to support this usability engineering.

Kloster Irsee

From the technical side, robust speech recognition is one of the most important issues for multi-modal dialogues, especially in mobile (and noisy) environments. In the first session of the workshop, various approaches to this challenge were proposed: lip-reading, multi-channel noise reduction, and exploiting structures in language models.

The poster and oral sessions that followed discussed a number of important usability issues for multimodal systems, such as the influence of the system feedback onto the user's reaction and the correct choice of an animated feedback character. Several authors pointed out the necessity of separating the generic dialogue behaviour of a speech system from the task- and domain-specific aspects, and presented different ways of representing external devices and services and their interactive behaviour. >

Inside this Issue:

D'Homme Project Report
David Milward **3**

The Dictionary of the Future – Some Proposed Extensions
Michael Zock **5**

Opinion Column
Kenneth Church **7**

New Editor for *ELSNews*
Geoffrey Sampson **8**

SIGdial Page: Wireless Multimodal – the Next Challenge for Speech Recognition
Roberto Pieraccini **9**

HLTheses: New Research Repository
Maria Vasiliou **10**

FASTY – A Multi-Lingual Approach to Text Prediction
Marco Baroni **11**

New MSc in Lexicography and HLT
Adam Kilgarriff **12**

Information Extraction in the MUMIS Project
H. Saggion, H. Cunningham, Y. Wilks, T. Declerck, P. Witterburg **13**

LREC 2002
A. Atutxa, E. Garcia, K. Sarasola, G. Hernández, and M. Aranzabe **15**

Future Events **16**

Summer 2002

elsnet

ISSN 1350-990X

The day ended with a perfect summer evening in one of the local beer gardens, where participants could test the products of the monastery's brewery.

Unfortunately, the designated keynote speaker for the second day became ill and could not make it to the workshop. So Sadaoki Furui from the Tokyo Institute of Technology, the current president of ISCA, bravely jumped in and gave an inspiring talk entitled *Towards Speech Recognition in Real Environments*. Furui proposed an approach for handling the complexity of conversational speech by integrating multiple information sources in one huge search space.

Looking up to heaven... The workshop participants experience the visual effects of the fresco in the cupola of the monestary's baroque entrance hall

A number of papers and poster presentations on this day made it clear that mobile environments do not necessarily mean car environments. The participants demonstrated various navigation applications for pedestrians and even some "mobile" applications in living-room scenarios. Other contributions focussed on the performance of existing multi-modal technology in real applications under the constraints of current wireless networks.

The theoretical presentations were accompanied by live demonstrations of an in-car navigation system and a multi-modal user interface to an electronic programme guide.

One of the tourist highlights of the workshop was, of course, the guided tour of the monastery's church, where Paul Heisterkamp from DaimlerChrysler not only translated the subtle details of the baroque art of living but also managed to draw some parallels between the Catholic church in those times and the big multinational companies today.

Laila Dybkjaer, Sadaoki Furui, and Roberto Pieraccini discuss the latest advances in dialogue technology

The final day of the workshop started with a tutorial on *Speech Recognition Methods and their Potential for Dialogue Systems in Mobile Environments* by Harald Huening from DaimlerChrysler. Harald gave a comprehensive overview of the state-of-the-art in speech recognition and language modelling technology and the specific challenges that arise for spoken language dialogues in noisy environments. The workshop ended with a set of presentations on user interface issues, where, amongst other contributions, Sharon Oviatt from Oregon Health and Sciences University demonstrated how children adapt to the behaviour of animated characters in a science education task.

In the concluding plenary discussion, the question of "killer-application" for multi-modal dialogue systems was raised once again. Most of the participants agreed that there's probably not just one single application or usage scenario that will push multi-modal user interfaces in the market, but rather it will be the adaptivity and flexibility of multi-modal systems that creates their advantages over uni-modal solutions for specific situations. There are still many unanswered questions and a long road ahead towards the ultimate multimodal dialogue interface. Let us hope that when we come back to the cloisters three years from now, in 2005, for the next Kloster Irsee workshop, we will see how much of the potential of this technology has already made it to the market.

ELSNews is published at the School of Cognitive and Computing Sciences, University of Sussex. It is printed by the University of Sussex Print Unit.

Editor: Jenny Norris.

Editorial Team: Geoffrey Sampson, Steven Krauwer, and Brigitte Burger.

ISSN 1350-990X

© ELSNET 2002

FOR INFORMATION
Contributions to *ELSNews*, and corrections, should be sent to:
lynneca@cogs.susx.ac.uk

Tel: +44 1273 678659
Fax: +44 1273 671320

Material for the next issue is due:
20 September 2002

FOR INFORMATION

Andreas Kellner is Senior Scientist in the Man-Machine Interfaces Research group at Philips Research Laboratories in Aachen, Germany

Email: Andreas.Kellner@philips.com
Web: www.andreas-kellner.de

D'Homme: Dialogues in the Home Machine Environment

David Milward, *Linguamatics Ltd, UK*

Feature

Home devices are becoming ever more sophisticated, and increasing numbers of devices can be controlled across a home network. Automatic curtains, security systems, and integrated, distributed audio/visual systems are already available and can be controlled from a *master controller* such as a home computer or dedicated remote control. However, these interfaces are not a radical improvement on existing home-control devices (e.g., a TV remote control) and may be seen as more complicated to use as they control many devices. More intuitive, natural interfaces are required.

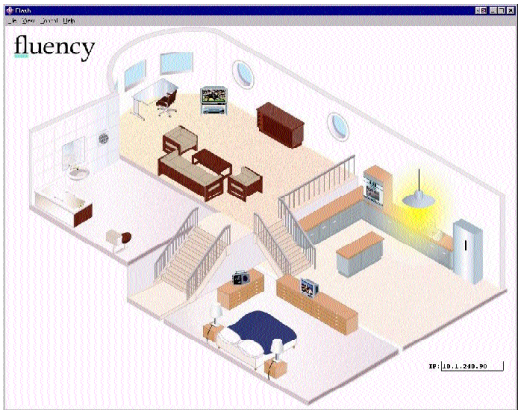
A spoken interface allows users, including the elderly and disabled, to better exploit the benefits of sophisticated devices in a networked and programmable home. With a spoken interface we can query multiple devices, e.g., by asking "Have I left anything on?" as we exit the house. We can control *virtual* devices, e.g., a TV comprising separate speakers, receiver, and screen, as easily as real devices. We can also use language to program devices to interact with each other, e.g., "Turn on the hall light when the front door is opened".

A spoken interface can also be used to control services without specifying which device or combination of devices will achieve this: e.g., "Cool the room to 18 degrees" or "Show the football match from last night". A *night mode* or *vacation mode* might involve sophisticated control of lighting, heating, blinds, and alarms.

Once we have spoken dialogue-based control of devices from within the home, remote control via the telephone is a natural extension. We can phone up the house to ask if we have left the oven on, and then ask the house to turn it off. Spoken dialogue is just as well suited for interactions initiated by the user and by the house: it may be just as important for the devices to contact us (e.g, the burglar alarm, smoke sensor, etc.) as vice versa.

D'Homme was a one-year EU-funded project designed to address the challenges in language understanding and dialogue management for controlling and querying multiple networked devices from inside or outside the home.

Theoretical work in D'Homme included investigation into issues of *plug-and-play*: how do we adapt dialogue management, interpretation, and speech recognition (using statistical or grammar-based language models) as devices are moved around the home, or new devices are added? It also included work on semi-automatic conversion from menu-based systems into dialogue systems (e.g., taking a video recorder's menu-driven interface and converting it for natural spoken interaction).



Practical work included the construction of demonstrators in Spanish, Swedish, and English. The project used simple real devices (on-off or scalar devices, e.g., lights and dimmers), and simulations of sensors and more complex devices such as VCRs. The example above shows a simulated house in a state where the kitchen light is on.

D'Homme also built a portable demo for controlling the lighting in a suite of rooms at the Telia Vision Centre in Stockholm. This was constructed using Lonworks X10 device protocols.

D'Homme stopped short of issues such as microphone placement or speaker identification (it assumed the use of a mobile phone or hand-held remote controllers with integrated microphones, or a clip-on microphone for hands-free use).

The partners in D'Homme were the Universities of Göteborg, Edinburgh, and Sevilla, and the companies SRI, netdecisions (the parent company of Fluency Voice Technology), and Telia.

Theoretical Challenges

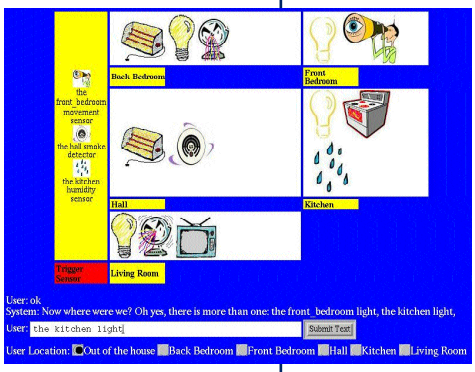
Most current dialogue systems operate in an information-seeking domain (e.g., finding customer requirements for booking a theatre ticket.) The home-control domain is very different and challenges many assumptions about the way in which spoken dialogue systems should be constructed. >

Dealing with plug-and-play issues, where devices may be plugged in and out of a network, provides a challenge to conventional approaches to language modelling and knowledge management in dialogue systems.

User-initiated complex commands, e.g., "Turn on the light and the heater", which may require follow up, do not fit easily into the form-filling paradigm exemplified by the current standard for dialogue systems, VoiceXML. Device-initiated interactions, e.g., "The movement sensor in the garden has triggered", suggest a more event-driven approach.

Plug-and-Play

The development of standards such as UPNP allows devices to be plugged into networks, and to advertise their functionality to other devices. In D'Homme we were interested in seeing if it is possible to provide similar plug-and-play capability at the linguistic level: i.e., to provide new devices with linguistic resources (such as a grammar and lexicon) to allow automatic re-configuration of a home control dialogue system. This cuts across traditional modularisation based on a single lexicon and grammar for the whole task.



For speech recognition and, in particular, for language modelling based on compiling a grammar into the recogniser, D'Homme showed how it is possible to adapt for plug-and-play by using a high level generic grammar, with specialisation to the specific available devices using a feature system.

Speech Recognition for Appliance Control

To provide accurate, speaker-independent recognition requires a grammar or a statistical language model for the domain. D'Homme evaluated both approaches for accuracy at the word level, individual slot values, sentence level, and full semantics. The evaluation was performed on a reasonably substantial corpus of utterances collected for the domain from users of prototype demonstrator systems. The grammar-based system was good at providing a fully correct semantics on utterances by trained users. The statistical system showed more graceful degradation and was particularly good when evaluated according to individual slot values.

Reconfigurable Dialogue and Knowledge Management

For complex devices, such as VCRs and mobile phones, D'Homme showed how existing knowledge from menu-based systems can be reused to create dialogue systems. The resulting dialogues can follow the original menu structure and the user also has the ability to take the initiative and to go immediately to a leaf node in the menu tree, or to supply information before the system asks for it. For example:

System: Welcome to the VCR manager.
User: Add a programme.
System: OK, add a programme.
What channel do you want?
User: Channel one today from eight thirty to nine thirty.

D'Homme also considered how users might set up a dialogue system in their own home, and investigated the use of semantic networks and multi-dimensional inheritance for maximal reuse of information across device types, and to embody necessary inferences (e.g., a dimmer is a kind of light).

Device Initiative

A multi-modal demonstrator (shown left) was designed to illustrate event-driven dialogue management behaviour. The user can click on a sensor to simulate an event such as a smoke detector being triggered. High priority events will interrupt a dialogue, e.g:

User: Turn off the light
System: The hall smoke detector has been triggered

Information from other sensors, e.g., humidity sensors, is only presented to a user at the end of a dialogue exchange.

Alternative Approaches

There are existing commercial products providing speech recognition for control of networked homes. These allow users to associate particular strings of words with particular commands, or sequences of commands, but are not designed for interactive dialogue.

There is also a lot of interest in embedding speech recognition within individual devices, though this does lose the benefit of being able to communicate simultaneously with multiple devices.

Conclusions

Spoken dialogue provides a natural, universal interface for networked home appliances. We are currently working towards deployment in commercial systems.

The home domain presented some very clear research challenges, and we would expect that many of the techniques adopted here will be applicable outside device control.

FOR INFORMATION

David Milward is CTO at Linguamatics Ltd. and was co-ordinator of D'Homme whilst at SRI International.

E-mail: david.milward@linguamatics.com

For more information, project reports, latest demos, and an audio visual presentation, visit the **D'Homme Website:** www.linggu.se/projekt/dhomme

Feature

The Dictionary of the Future – Some Proposed Extensions

Michael Zock, *LIMSI/CNRS, France*

Michael Zock

Dictionaries are a vital component of any natural language processing system, whether natural or artificial. In their modern electronic form, dictionaries have tremendous potential (which is largely under-exploited), provided they are built in a way that allows for their use not only by experts or machines, but also by ordinary language users.

Unfortunately, despite the enormous interest in electronic dictionaries in general and thesaurus-like semantic networks (such as WordNet) in particular, little attention has been paid to the language user. And yet, a lexical database is worthless if the data are not (easily) accessible.

There are many possible ways to make a dictionary useful for people in their daily tasks of processing or learning a language. For example, a dictionary fully interfaced with a word processor would facilitate active reading In such an environment, clicking on a word could reveal its translation, its definition, its usage (in the current context), grammatical information about it, its spoken form, etc. (see Figure 1).

Text to study	Translation
kana/romaji	to do
やまだ : スミスさんは なにを して いますか。 たなか : メールを かいて います。	**Synonym**
やまだ : ブラウンさんは なにを して いますか。 たなか : ほんしゃに でんわ して います。	shitogeru
	Grammatical information
kana/romaji	te-form of the verb suru
Yamada : Sumisu-san wa nani o shite imasu ka? *Tanaka* : Meeru o kaite imasu.	**Sentence pattern**
Yamada : Brown-san wa nani o shite imasu ka? *Tanaka* : Honsha ni denwa shite imasu.	[SUBJECT] wa [SOMETHING] o [VERB *te-form* + imasu]

Figure 1

Additionally, the development of tools that help the user to identify the lemma (by analysis) or the inflected form (by generation) would be very useful here.

Amongst other potential improvements would be functionalities to assist users in:

- Revealing the word that is on their mind (*word access*)
- Memorising words, syntactic structures, and so on (*memorisation and automation*).

The memorisation and automation of words in a syntactic context could be supported by the use of automatically generated exercises in conjunction with flashcards. This could involve the use of goal-driven, template-based sentence generation: the choice of a communicative goal would trigger a specific sentence pattern, which would then be instantiated by the words chosen by the user. This method would facilitate not only the generation of words in context (sentences), but also the generation of exercises containing the words and syntactic structures that the user wants to memorise or automate.

Access to words in electronic dictionaries could be enhanced a great deal. For example, if the dictionaries were built like "mental dictionaries" (associative networks, akin to WordNet, but with many more relations, in particular at the syntagmatic axis), they could assist people in finding not only new ideas (brainstorming) but also the word that's on the tip of their tongue/pen. We have all been in this position – looking for a word (or someone's name) that we know, but not being able to remember (access) it.

Work done by psychologists shows that people who are in the above-mentioned state do know a lot about the word in question (meaning, number of syllables, etc). Also, the word they are producing or thinking of in its place has a lot in common with the target word (initial letter/phoneme, syntactic category, semantic field, etc.). This being so, it should be possible to exploit this fact and build a program that could assist the speaker/writer by revealing the word that is on his/her mind (tongue/pen).

The development of such a program could be guided by speech error data which show that words are stored in two modes : by meaning and by sound. Either (or both) of these can cause problems.

The fact that access may be inhibited by *grapheme* or *morpheme reversal* can be used to find the correct (intended) token. The following steps may be necessary to automate this process:

- Given an incorrect input (e.g., "poteaux" instead of "topos"), make certain permutations and check in the dictionary to see whether there is a candidate
- If there isn't an alternative candidate, then convert the graphemes into phonemes, perform the legal permutations, and check again in the dictionary

>

- If there is more than one candidate, then use semantic criteria or domain knowledge ("bank" as in geography versus finance) to decide on the priority of presentation of alternatives
- If there is still no candidate, then replace some of the sounds by similar phonemes: e.g., /p/ by /b/ or /m/.

Turning to meaning-based access problems, we start from the following assumption:

> *the mental dictionary is a huge semantic network composed of words (nodes) and associations (links), with each being able to activate the other.*

<-A KIND OF-
clinic
sanatorium

<-IS-A-
military hospital
psychiatric hospital

ACTOR
doctor
patient
nurse

Figure 2a: Proposed candidates grouped according to the nature of the link

Finding a word amounts to entering the network and following the links leading from the source node (the first word that comes to your mind) to the target word (the one you're looking for). Suppose you wanted to find the word "nurse" (*target word*), yet the only token coming to your mind were "hospital". In this case the system would display by category (chunks) all the words linked to "hospital" (Figure 2a).

Put differently, the system would build internally a small semantic network with "hospital" in the centre, and as immediate satellites, all the words having a direct link with it (Figure 2b). If the candidate is in any of these lists, the search stops – otherwise it goes on. The user could either take a word occurring in any of the lists or a new token.

Some interesting questions arise in this practical scenario. For example:

- What are the links or associations between words?
- Can we reasonably encode (all or some of) them into a dictionary?
- Where should we look to obtain a list of associations (Mel'cuk's work)?
- Should we allow for adding private information (personal associations)?
- Is it possible to extract this kind of information automatically by parsing an encyclopedia or large amounts of text?, and so on.

In summary, it seems that builders of electronic dictionaries are sitting on an untapped gold mine that is still waiting to be explored and exploited. Yet there is good reason to believe that there is a market for products that integrate more advanced ways of using and accessing lexical information. Let's hope that we will begin to see some of these ideas and methods being incorporated into electronic dictionaries in the not-too-distant future.

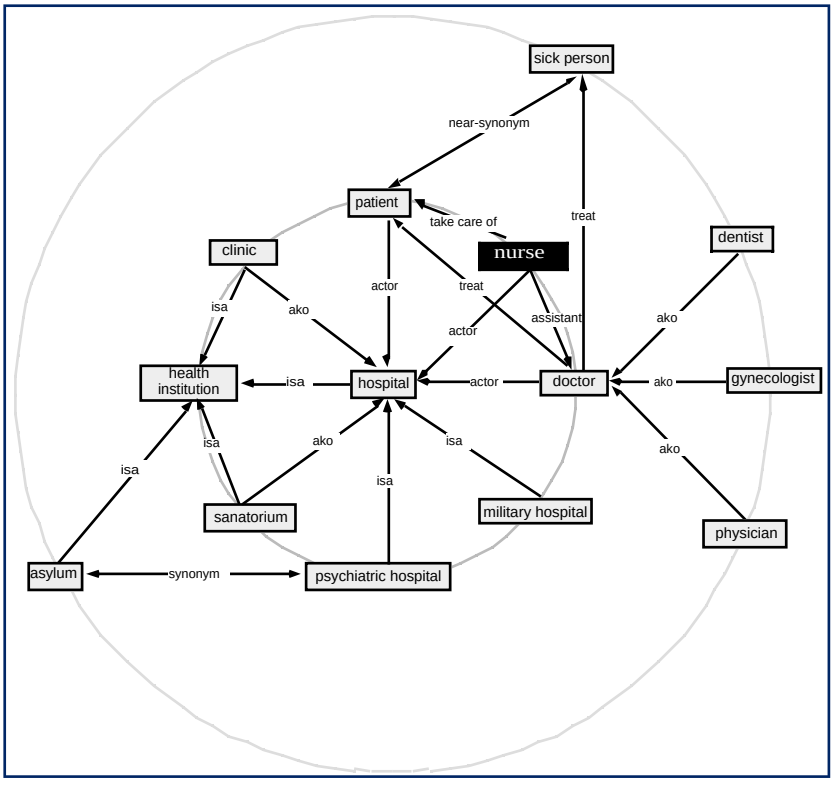


Figure 2b: Search based on propagation in a network (internal representation)

FOR INFORMATION

Michael Zock is permanent researcher at the CNRS, working in the LIR group (Langues Information Représentations) at LIMSI, Orsay, France.

Email: zockm@noos.fr

Web: www.limsi.fr/Individu/zock/

Opinion Column

Strategy

Kenneth Church, *AT&T Labs Research, USA*

The field of Computational Linguistics is doing better and better! It used to be very difficult to prepare a talk for a Computational Linguistics audience because the field was so interdisciplinary that there was almost nothing that you could assume that everyone would already know. The field will really have arrived when a course in speech and language processing is a normal part of every undergraduate and graduate Computer Science, Electronic Engineering, and Linguistics programme – and we're a long way from that. But things are improving … We now have several excellent textbooks such as: Manning and Schütze, *Foundations of Statistical Natural Language Processing*, MIT Press, 1999; and Jurafsky and Martin, *Speech and Language Processing*, Prentice Hall, 2000.

A quick search of the web shows that these textbooks are being used across a wide range of universities, including: Amsterdam; Berkeley; Brown; Charles University (Prague); CMU; Colorado; Columbia; Cornell; Edinburgh; Essex; Göteborg; Helsinki University of Technology; Illinois; IIT Delhi; Penn; Rutgers; Saarbrücken; Sheffield; Stanford; Toronto; Washington; Utah; and more.

There were, of course, many other obstacles that limited the size of the field. It used to be hard to join in on the fun because only a few large industrial laboratories could afford to collect large amounts of data. Thanks to data collection efforts such as the Linguistic Data Consortium (LDC) (www.ldc.upenn.edu/) and ELSNET (www.elsnet.org/ resources.html), and, of course, the Web, data is no longer the problem it used to be. Of course, you can never have too much of a good thing. Balance is fine, but Eric Brill (research.microsoft.com/~brill/Pubs/HLT2001.pdf) and others are finding patterns in billion-word corpora that could not be found in a "mere" million words. In other words, as IBM used to say, more data are better data. Those are still fighting words, but they don't sound nearly as shocking as they once did (now that more people have been able to join in on the fun).

Tools also used to be a problem. I used to teach tutorials at summer schools and conferences, on Unix for Poets (http://www.research.att.com/~kwc/publications.html), arguing that it was easier to do it yourself than beg a computer officer for help. I am really happy to see that these tutorials and tools have improved considerably over the years, and that they are now fully integrated into many of the university courses, as was discussed at an ACL-2002 workshop on Teaching NLP (www.eecs.umich.edu/~radev/TeachingNLP) held on July 7 this year. Some of these tools/tutorials can be downloaded off the web (nltk.sourceforge.net/).

There is always, though, more we could do to promote our field. We could learn a lot from Theoretical Computer Science. Historically, Theory has paid more attention to teaching than we have. They have also worked hard on strategy. There have been many strategy discussions in the theory community over the years: a good recent example is www.research.att.com/~dsj/nsflist.html. The theory community regularly exchange lists of open problems along with difficulty ratings. Students know before they solve a problem whether it is worth a conference paper or a superstar award.

Many organisations in universities, industry, funding agencies, etc., work hard on strategy. There are plenty of examples on the web:

- www.nsf.gov/pubs/2001/nsf0104/strategy.htm
- www.darpa.mil/body/mission.html
- medg.lcs.mit.edu/doyle/publications/sdcr96.pdf
- www.gridforum.org/L_About/about.htm

It is hard to say why strategy is important, but I have noticed, at least within my own institution, that groups that work hard on strategy have grown and prospered over the years. Strategy is never as urgent as the next conference paper deadline, but it is probably more important.

Organisations may or may not follow their own recommendations. The discussion that produces the strategy document is extremely valuable, nevertheless, perhaps more so that anything that happens after the document is finalised. Panels on strategy offer an excellent forum for people to meet and look at the field from a broader perspective. In addition, the theory community has observed that even after the people involved in the original discussion have long since forgotten the outcome, the recommendations continue to live on and influence and broaden the best and most aggressive students for years to come.

I have looked around for discussions of strategy within our field. There are a few examples, though not as many as there should be:

- www.elsnet.org/about.html *and* www.elsnet.org/roadmap.html
- www.ldc.upenn.edu/ldc/about/ldc_intro.html
- www-nlpir.nist.gov/projects/duc/papers/
- LREC workshops (to order proceedings, see www.lrec-conf.org)

The LDC link above was developed a decade ago and was largely responsible for the success of the consortium. If more groups in our field put the same kind of energy into strategy, I believe there would be more success stories like the LDC. >

Opinion (contd.)

A delightful "near miss" is Martin Kay's reflections on ICCL and COLING (under www.coling.org). It establishes direction for the format of COLING conferences in a style that can only be described as "classic Martin". I wish we could somehow convince Martin to write a document in the same delightful style that would establish a direction for the field rather than an atmosphere for a conference.

The big question is what the content of this strategy ought to be. It is natural to start with a survey. There are some excellent surveys out there, such as cslu.cse.ogi.edu/HLTsurvey/ (1996), which is currently being updated. We could then construct a list of core areas by combining that survey with textbooks, course materials, conference proceedings, home pages, etc. In addition, it would be good to construct a list of areas that we might want to reach out into, such as areas covered by closely related conferences on topics like digital libraries, data mining, machine learning, linguistics, information retrieval, artificial intelligence, theory, and so on. A lot of work has been appearing outside our standard venues, perhaps because of the natural tendency to set rather narrow exclusive boundaries. These boundaries, for instance, put Jelinek and Salton on the far side of the line for many years. For the long-term future of the field, we ought to interpret *Computational Linguistics* as generously and as broadly as possible – perhaps as anything concerned with algorithmic processing of the form, interpretation, or use of text, speech, or gesture.

I'm thinking that we could produce something similar to www.research.att.com/~dsj/nsflist.html. In the end, I would like to see a document that is largely focussed on areas that you would expect to see in such a document (topics covered by most textbooks, courses, and conferences) but would also include some "hot" areas where there is a lot of recent excitement, such as biology, data mining, and grid computing. The document should establish a rough consensus about where the boundaries are, and what novices should learn, but do so in a way that is as inclusive as possible (and then some). I could take a stab at such a document here, but rather than do that, I would hope that I could encourage all of you to discuss these issues with your colleagues during the summer conferences. Hopefully, by the end of the summer we will be closer to a consensus. And then we can twist Martin's arm to say it as only he can say it.

FOR INFORMATION

Ken Church is Head of the Department of Data Mining at AT&T Labs Research in New Jersey

Email: kwc@research.att.com

Web: http://www.research.att.com/~kwc

New *ELSNews* Editor

Farewell and Hail

Geoffrey Sampson

With this issue of *ELSNews* we say a fond goodbye to Jenny Norris, Editor since March 2000. With her family Jenny is about to move to Orkney. For the benefit of readers unfamiliar with UK geography, Orkney is a scattering of beautiful North Atlantic islands, several of them inhabited: once Norse, they were annexed by the King of Scotland 500 years ago in lieu of an unpaid dowry. Jenny is changing field into marine biology. She looks forward to Orkney as an ideal place to bring up children, though she says she is going to miss trees (on the islands there are hardly any). We are going to miss Jenny.

Replacing her in the Editor's chair, as soon as formalities can be completed, will be Lynne Cahill (pictured). Lynne has a longstanding association with the University of Sussex, having taken her bachelor's, master's, and doctoral degrees here. Her first degree was in linguistics, and her master's in cognitive studies. Lynne completed her doctorate (on the representation of morphological alternations) under the supervision of Gerald Gazdar, in 1990.

Lynne Cahill

Since then, Lynne has worked at Sussex both as a research fellow and as a lecturer in linguistics. She has recently returned to Sussex after a four year spell at the neighbouring University of Brighton, where she was involved with several different research projects.

Lynne is a keen cricket and football fan (although she doesn't participate!) and also enjoys designing and making clothes, mainly for her two children, Alice (11) and Hannah (5).

The email address for *ELSNews* contributions and correspondence becomes lynneca@cogs.susx.ac.uk.

Welcome to the team, Lynne! We are looking forward to working with you.

Wireless Multimodal – the Next Challenge for Speech Recognition

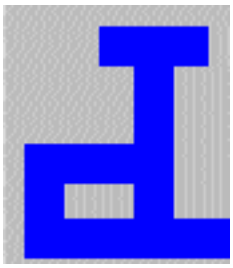
Roberto Pieraccini, *SpeechWorks International, New York*

The increasing automation of call centres, including the replacement of touch-tone interactive voice response (IVR) services with automatic spoken dialogue systems, is an irreversible process that started a few years ago, thanks to companies such as SpeechWorks, Nuance, and Philips. With the increased market demand for wireless communication, and the distinction between personal telephones and hand-held computers becoming fuzzier, new opportunities and challenges are appearing for the speech recognition technology. New wireless communication devices, such as smartphones and personal digital assistants (PDAs), are characterised by increased computing power and high-resolution displays. At the same time, the third generation (3G) wireless networks, which allow the simultaneous flow of data and voice, are gradually being deployed across the globe. However, the ever-decreasing size of wireless portable devices prohibits the use of regular keyboards and makes even limited keypad use difficult. Multimodal speech and visual interfaces are poised to become the interface of choice for personal wireless devices.

Server-based, Distributed, and Embedded Speech Technology

Telephony-based spoken dialogue applications rely today on a server-based architecture, with a simple telephone handset client and speech recognition server communicating voice signals over the telephony network. With a powerful client device and a digital network, it becomes possible to run speech recognition front-end on the device itself, while performing the search on the server: this model has become known as Distributed Speech Recognition (DSR). One advantage of DSR is the reduced bandwidth. Feature vectors can be compressed to between 4.8 and 9.6 kbps compared to the 64 kbps required for conventional telephony speech. A second advantage is that latencies can be lowered with speech signal endpoint detection implemented on the client. Perhaps most significantly, processing power on the client can be used to improve recognition accuracy through improved signal processing, including both noise compensation and a higher sampling rate than the typical 4kHz of the telephone based systems. In view of future deployments of DSR technology, the European Telecommunication Standards Institute (ETSI), has created the Aurora Working Group to standardise the DSR protocol by defining the feature extraction methods.

With enough client-side memory and CPU power, we can even embed speech recognition engines entirely into the client. A compromise might be hybrid systems, where small or common vocabulary speech recognition is performed locally, while large or dynamic vocabularies are handled through DSR.



Multi-modal User Interfaces

The availability of a graphical user interface (GUI) enriches the user experience by providing features that are complementary to the voice user interface (VUI) of speech-only systems. In particular, GUIs facilitate the presentation of visual information that is difficult or impossible to describe verbally (e.g., pictures or maps). Discourse patterns that are typical of speech-only applications, such as confirmation and correction, re-try on timeout, list navigation, etc., may be implemented rather differently with a multi-modal interface. For instance, list navigation for long ad-hoc lists or tables, which is impractical for speech-only systems, is more effective with a visual display. The visual presentation of the current application state and available options (e.g., "What can I say now?") help the user to complete tasks. The Figure shows an example of a wireless multimodal prototype developed at SpeechWorks within the DARPA Communicator project, in collaboration with MapQuest, Compaq, and Lobby7 [1].

One of the challenges of multi-modal user interfaces for wireless devices arises from the possibility of adapting to different situations (*situationalisation*). At different moments the user may be subject to different constraints on the visual and aural channels (e.g., walking whilst carrying things, driving a car, being in a noisy environment, wanting privacy, etc.). Similarly, the same application could potentially be used from devices having different capabilities (eg., WAP, display-less, text-only, audio-less, etc.). In the DARPA project, a simple user interface situationalisation was done for speech-only versus GUI-only or mixed-mode, where the speech-only version of the system has longer prompts in order to compensate for the absence of visual prompting in the forms of highlighted fields or menus. The main challenge of situationalisation is in reducing the cost of designing different UIs for the same application and in devising strategies for adapting the UI depending on the current situation. >

SIGdial Page (ACL Special Interest Group on Discourse and Dialogue)

Summer 2002

elsnet

Architectures, Markup Languages, and Standards

In the face of a growing multi-modal device and network market, including the advent of the 3G network, companies and institutions involved in multi-modal wireless technology are concerned with architectural standardisation issues. Early prototypes have been built, both having unique architectures and around current standards, such as HTML for GUIs and VoiceXML for VUIs. Synchronisation of HTML and VoiceXML documents supports the development of fairly complex multi-modal applications, constrained by the limitations of the two markup language models and the need to mutually refresh them. The SALT (Speech Application Language Tags) extension to HTML provides additional tags that support multimodal applications. There are currently more than 20 companies involved in the SALT forum. In February 2002, the World Wide Web Consortium (W3C) formed a Multimodal Interaction Activity working group, with the charter of creating the specification for a multi-modal standard by February 2004.

Conclusions

There is an exciting growth of activity surrounding the development of GUI, VUI, and multimodal technologies for wireless applications. However, the adoption and penetration of multimodal wireless applications depends not only on the availability of the enabling technologies, but also on other factors such as the reliability and the cost of wireless digital networks and the existence of standards supporting them. DSR and hybrid approaches to speech recognition promise to lower the cost and increase the accuracy of applications. Standards, such as those being determined by the SALT and W3C working groups promise portability, scalability, and the critical mass required to deploy applications.

References

[1] Pieraccini, R., Carpenter, B., Woudenberg, E., Caskey, S., Springer, S., Bloom, J., Phillips, M., "Multi-modal Spoken Dialog with Wireless Devices". In *Proc. of ISCA Tutorial and Research Workshop – Multi-modal Dialog in Mobile Environments*, June 17-19, 2002. Kloster Irsee, Germany.

FOR INFORMATION

Roberto Pieraccini is director of the Natural Dialogue Group at SpeechWorks International, and is a member of SIGDial

Email: roberto@speechworks.com

Aurora Working Group: www.etsi.org/frameset/home.htm?/technicalactiv/DSR/dsr.htm

SIGdial Website: www.sigdial.org

SALT forum: www.saltforum.org

W3C Multimodal Interaction Activity Working Group: www.w3/2002/mmi

Announcement and Call for Submissions

HLTheses: New Research Repository

The Speech and Language Group at The University of Patras Wire Communications Laboratory (Department of Electrical and Computer Engineering) has launched a new programme called HLTheses – a repository of information about current and past PhD research in HLT, which looks set to become an extremely useful and informative reference tool. We hope that ELSNews *readers will take up the invitation to add their own details to the site.*

Opening Announcement

HLTheses is a website devoted to current research in Human Language Technology (HLT). It is an ELSNET initiative, co-sponsored by ISCA and EACL, aiming to aid HLT researchers and promote their research by providing information on relevant PhDs worldwide. HLTheses contains extended PhD abstracts (completed and on-going), CVs, and contact information of PhD authors and researchers. The intention is for the site to reference all PhD theses related to HLT worldwide, thus serving as a state-of-the-art guide and a place where researchers can locate others who work in similar areas.

Currently, only a small number of abstracts are available for viewing on HLTheses, but the number is expected to increase shortly.

Call for PhD Abstracts

HLTheses invites all researchers holding or pursuing a PhD in any area of Human Language Technology to submit extended abstracts of their PhD theses (and brief CVs) in order to be listed in the HLTheses repository. The PhDs need not be completed: abstracts for on-going PhD are equally eligible.

With plans for a large promotional campaign, PhDs presented in HLTheses are expected to create a great impact. Additionally, the intention is to produce annual printed volumes containing details of the theses completed every year, sorted by research area.

FOR INFORMATION

Visit the site at http://HLTheses.elsnet.org.

For detailed submission information, visit: http://hltheses.elsnet.org/information/subminfo/howsubmthes.htm

For more information contact Maria Vasiliou
Email: mvasiliou@wcl.ee.upatras.gr

FASTY – A Multi-lingual Approach to Text Prediction

Feature

Marco Baroni, *Austrian Research Institute for Artificial Intelligence (OeFAI), Vienna*

Written communication and information exchange is a vital factor in human society. Impairments, which lead to a reduction of typing speed, therefore, severely influence quality of life and cut off a person from equal participation in the information society.

This problem is being addressed by FASTY – a European project funded by the Fifth IST R&D Framework Programme. The project started in January 2001, and will run until December 2003. The consortium is led by Fortec – Vienna University of Technology (Austria). The other partners are the Austrian Research Institute for Artificial Intelligence (Austria), FTB (Forschungsinstitut Technologie-Behindertenhilfe, Germany), the Department of Linguistics at Uppsala University (Sweden), Multitel ASBL (Belgium), IGEL GmbH (Germany), Elisabethinum Axams (Austria), IkuT – Ingenieurbüro für Kunst und Technik (Germany), and Facultés Universitaires Notre-Dame de la Paix (Belgium).

Since languages display a high degree of redundancy, low-speed typists can be supported by Predictive Typing (PT) systems. Such systems attempt to predict subsequent portions of text by analysing the text already entered by the writer. Character-by-character text entry is replaced by making a single selection as soon as the desired word or sequence is offered by the system in the selection menu.

State-of-the-art programs for PT claim Keystroke Saving Rates (KSR) of up to 75%. This does not mean, however, that the text generation rate increases by a factor of four. Using PT consumes time, because the user needs to read the selection menu and make a decision. Only substantial KSRs will lead to an increase of communication speed. To double the text generation rate of a typical mouth-stick user, the program must offer a KSR of about 65%.

Such high rates are currently only achieved for English, a language almost unique in having a very limited set of inflectional endings. This property makes it ideally suited to the currently most popular PT technology, which uses a statistical approach based on the probability of word n-grams. By adapting programs designed for English to other languages (especially highly inflected ones), the KSR drops significantly (usually below 30%). Therefore, most motor/speech impaired persons will experience no gain in text generation rate from existing programs.

FASTY aims at providing impaired speakers of languages other than English with PT systems that perform as well as those that nowadays are available for this language only. FASTY is currently being implemented for Dutch, French, German, and Swedish, but it is based on a generic modular architecture, with a clear separation between processes and language-specific resources. This should make adaptation to other languages relatively easy.

The target languages of the FASTY project are highly inflecting. Depending on the syntactic context, words take different forms. As already mentioned, this makes standard n-gram language modelling techniques less effective. Thus, additional methods that are able to cope with syntactic constraints are needed. Furthermore, in most of FASTY's target languages (i.e., Dutch, German, Swedish), productively formed compounds are written as single orthographic strings (in contrast to English, where compound terms are groups of words separated by a blank character or, at least, a hyphen). This causes serious problems in terms of lexical coverage and data sparseness to systems that do not perform some type of compound processing.

The FASTY language component includes the following modules: word- and part-of-speech-based n-gram prediction; grammar-based prediction; compound prediction; morphological lexicon; user lexicon; collocation-based prediction. The modules are driven by a controller engine that manages the input requirements of each component, establishes the required input data from the context, and combines the outputs in a meaningful way.

Preliminary experiments indicate that the n-gram-based models, despite the problems mentioned above, still provide reasonable predictive power, and they constitute the core of the FASTY language component.

The grammar-based module performs a partial parse of the current input, and it ranks the predictions provided by the other modules on the basis of the grammatical information provided by the parse. Moreover, in contexts where all predictions by the core component are syntactically ill-formed, the grammar-based module generates well-formed predictions using the morphological lexicon.

The compound prediction module allows the user to type (nominal) compounds in multiple steps. The user can choose to complete the word constituting the first part of the compound (if it is in the prediction list), and then re-enter the prediction loop for the current word, now getting predictions for the second part of the compound (and this process can be repeated as many times as necessary to obtain longer compounds). >

Compound prediction is based on a set of compound-specific statistical models trained on a corpus where compounds have been split into modifiers and heads. In preliminary research to be reported at ECAI02 and COLING02, we have found that our split compound model leads to an improvement in compound word KSR of more than 15%, over a baseline model without compound analysis. However, integrating compound- and whole-word prediction so that compound completions do not get in the way of simple-word predictions turns out to be a rather difficult task, and we are still experimenting with alternative integration strategies.

The morphological lexicon provides the necessary morphosyntactic information to the grammar-based module. It also functions as a last resort prediction source, if all the other modules run out of completions before the user finds the word s/he intends to type.

The user lexicon is an additional, dynamic resource intended to support the style and vocabulary preferred by a particular user. It contains words and n-grams collected from the texts written by the user so far, and thus may contain words and phrases that may not be present in the general dictionary but are of importance for the user (e.g., names of people the user often addresses, specific terminology s/he is using, etc.) The user dictionary is automatically augmented during text entry, allowing the prediction of words new to the system after the first time they are used. At the end of the session the user can choose whether to save or discard new terms.

The final version of FASTY will also include a collocation-based module to provide predictions based on the degree of textual association between words (or word-classes). This module has not yet been implemented.

The language resources for all the FASTY languages are nearly ready, and we expect to complete the implementation of a realistic prototype of the whole system (excluding collocation-based prediction) very soon.

At this stage, the biggest issues are those pertaining to the integration of the various components of the system.

FOR INFORMATION

Marco Baroni is a researcher at the OeFAI, currently working at the FASTY project, and focussing in particular on compound prediction

Email: marco@ai.univie.ac.at

More about FASTY: www.fortec.tuwien.ac.at/fasty

Announcement of New Course

Like a Horse and Carriage: New MSc in Lexicography and HLT

As Human Language Technologies (HLTs) move into the marketplace, so the need for large-scale resources is everywhere acknowledged. This is now a commonplace. How we create such resources, and of sufficient quality, however, is not. While machine-readable dictionaries have been widely explored over the last fifteen years, and offer a great deal, they have limitations: the format of the data is frequently not as consistent as the computationalist requires; other critical data is simply absent; and the description of the language is often flawed.

How can this be remedied? One part of the answer is to train people in both lexicography and NLP, so they will know both what is required and how it can be accurately found and represented.

At the University of Brighton, leading lexicographers Sue Atkins and Michael Rundell have joined forces with the Information Technology Research Institute (ITRI)'s HLT laboratory to set up a Masters course linking HLT and dictionary-making. It is the first MSc of its kind.

The course, running for the first time in 2002-03, is designed for recent graduates and for more experienced lexicographers or technologists who wish to complement their skills with those from the "other side".

Course modules are also offered on an intensive one-week basis, for professional updating and development. Alongside the core modules on lexicography (writing lexical entries which accurately reflect corpus data) and lexical computing (HLT as applied to the lexicon and lexicon development) there are modules on: corpus design and use; automating lexicographic tasks; bilingual lexicography and the use of parallel corpora in translation; dictionary project management; and corpora and language teaching.

ITRI has already collaborated with Michael Rundell in producing the Macmillan English Dictionary for Advanced Learners, the first English dictionary to be compiled with the aid of "word-sketching" software, which summarises parsed corpus material for the lexicographer.

Adam Kilgarriff, Course Leader, says: "We're very excited to have this new course to offer. It's unique in the world and, as lexicography and HLT have more and more need for each other, it's a fast-moving and fast-growing area, full of new possibilities. Our doubly-expert graduates are going to be in great demand."

The course is currently open for recruitment: see www.itri.brighton.ac.uk/courses/MScLex for details.

Information Extraction from Football Reports in the MUMIS Project

Horacio Saggion, Hamish Cunningham, *and* ***Yorick Wilks,*** *University of Sheffield;* ***Thierry Declerck,*** *DFKI;* ***Peter Wittenburg,*** *Max-Planck-Institure for Psycholinguistics*

Feature

The Multimedia Indexing and Searching Environment (MUMIS) is an on-going EU-funded project within the Information Society Programme (IST) of the European Union, Human Language Technology (HLT) section. The MUMIS objective is to develop technology to produce formal annotations about essential events in multimedia programme material. The project consortium consists of the University of Twente/CTIT (The Netherlands), the University of Sheffield (United Kingdom), the University of Nijmegen (The Netherlands), Max-Planck-Institut für Psycholinguistik (Germany), ESTEAM AB (Sweden), DFKI (Germany), and VDA (The Netherlands).

Project Overview

The vast amount of multimedia information available and the need to access its essential content accurately to satisfy users' demands encourages the development of techniques for multimedia indexing and searching. It is well known that there are no effective methods for automatic indexing and retrieving of image and video fragments on the basis of analysis of their visual features. MUMIS carries out indexing by applying Information Extraction (IE) to multimedia and multi-lingual information sources, merging information from many sources to improve indexing quality, and combining database queries with direct access to multimedia fragments on the multimedia programme. The key insight is that combining IE from speech transcripts (good temporal data but noisy text) with news reports (bad temporal data, clean text) can lead to a high quality conceptual index.

Various software components operate off-line to extract information from multi-source linguistic data in Dutch, English, and German and to produce a composite index of the events on the multimedia programme. The domain chosen for tuning the software components and for testing is football – in particular the Euro 2000 Championships.

The on-line part of MUMIS consists of a state-of-the-art user interface allowing the user to query the multimedia database (e.g., "The fouls committed by Beckham"). The user is first presented with selected video key-frames as thumbnails that can be played, obtaining the corresponding video and audio fragments.

A corpus of collected textual data in the three languages was used to build a multi-lingual lexicon and shared ontology of the football domain. Based on this shared model, three different off-line IE components, one per language, have been developed. These are used to extract the key events and participants from football reports and to produce XML output. A merging component or cross-document co-reference mechanism has been developed to merge the information produced by the three IE systems. Audio material is being analysed in order to obtain transcriptions of the football commentaries (spontaneous speech). The database is being populated through key-frames extraction from MPEG streams around a set of pre-defined time marks – obtained from the IE component.

In this article we focus specifically on the technology being developed at the University of Sheffield.

English Information Extraction System

At Sheffield we are working on IE from English sources, while other project participants are dealing with sources in Dutch and German. An analysis of the domain has led MUMIS to propose 31 types of event for a football match: kick-off; substitution; goal; foul; red card; yellow card; etc.). The elements to be extracted that are associated with these events are: players; teams; times; scores; and locations on the pitch.

The Sheffield IE system is conceptualised as a Java front-end system based on finite state transduction followed by a Prolog back-end system for inference over a classification hierarchy, implemented in SICStus Prolog.

Java Components of the System

The finite state machinery is based on ANNIE, a free IE system available as part of GATE, a General Architecture for Text Engineering (http://gate.ac.uk/) [*see* ELSNews *11.1 – Ed.*]. GATE provides language and processing resources that can be used to deploy natural language processing applications. ANNIE comes with tokeniser, sentence splitter, gazetteer lookup, semantic tagging, part-of-speech tagging, and orthographic name matcher. A rule-based lemmatiser has also been implemented and integrated into the system.

>

Semantic Tagging in the System

Gazetteer lookup and semantic tagging have been adapted for the football domain. We are using JAPE – Java Annotation Pattern Engine – a pattern-matching engine implemented in Java to identify and annotate regular expressions over annotations produced by any language processing components in the system. JAPE grammars are sets of rules that act on annotations assigned in earlier phases, in order to produce annotated entities. Semantic tagging is essential to identify typical expressions and domain jargon that can be semantically interpreted without relying on full parsing. For example, an expression like "Fiore (Totti, 82)'' indicates substitution, and can be identified with gazetteer lookup and a regular grammar. Regular grammars and semantic tagging components have been developed through careful corpus analysis.

Prolog Components of the System

The Prolog back-end system has been adapted from previous IE projects at Sheffield (e.g., LaSIE). We are using bottom-up chart parsing enriched with semantic rules that construct a naive semantic of each sentence in first order logical form. The parser uses two grammars: the first is a domain-dependent grammar used to produce logical forms for the entities of the football domain (e.g., players, teams, etc.); the second is a context-free phrasal grammar of English, enriched with features and values.

The discourse interpreter is based on a World Model (WM) representing the ontological (or hierarchical) knowledge about a particular domain. We have adopted the XI Knowledge Representation Language – a formalism that allows the user to code and operate with symbolic knowledge. XI is compiled into Prolog, making it possible to mix procedural knowledge with the constructs of the basic declarative formalism. The ontology of the football domain has been encoded into XI and properties about domain entities and events have been modelled. The interpreter works by mapping the information produced by the parsing and semantic interpretation into an evolving Discourse Model (DM) of the input text. The WM contains rules allowing the deduction of new knowledge from the "explicit'' information found in the text: for example, the fact that a player plays for a particular team, or the fact that the participants of a "substitution'' event must be players playing for the same team, are encoded in the WM.

Event and entity co-reference is the basic mechanism that allows the system to fill in properties for each event in the domain. Co-reference is a unification-based process that is based on:

- the notion of distance between nodes in the ontological hierarchy
- a measure of similarity between entities and events computed using the values of associated properties
- rules that constrain co-reference.

The co-reference algorithm is a general one that requires careful coding of the semantic properties of events and entities of the domain. Solving co-reference is essential for the IE task.

Consider the example below:

Poborsky fouls the Barcelona winger. Collina books him.

In this piece of text there are two possible antecedents for the pronoun *him*: *Poborsky* and *the Barcelona winger*; but assuming that the text is coherent, the pronoun *him* cannot co-refer with *the Barcelona winger* because a *book* event presupposes that someone committed a foul, in this case, the player *Poborsky*.

The system is completed with a template extraction and writing algorithm used to read the event instances deduced by the system and to produce XML output.

Evaluation

Evaluation of the system is being carried out using GATE's automated precision and recall evaluation tool, AnnotationDiff. The semantic tagging component was measured at 91% precision, 76% recall – comparing favourably with results obtained using a default named-entity recognition system that also uses gazetteer and grammar rules.

The Future

We have adapted language and processing resources for the MUMIS project. Our IE system integrates finite state machinery implemented in Java with full syntactic analysis and discourse interpretation implemented in Prolog. The IE system is operative and we are in the process of producing XML annotations for the full corpus.

FOR INFORMATION

Horacio Saggion, Hamish Cunningham, and **Yorick Wilks** are all based at the University of Sheffield's Department of Computer Science

Thierry Dederck is a senior consultant at DFKI's LT lab and is currently leading the MUMIS project

Peter Wittenburg is Head of the Technical Department of the Max Planck Institute for Psycholinguistics

Contact person for queries: Horacio Saggion
Email: h.saggion@dcs.shef.acuk

MUMIS Web site:
parlevink.cs.utwente.nl/projects/mugmis.html

LREC 2002: Multimodality, Standards, and Treebanks

A. Atutxa, E. Garcia, K. Sarasola, G. Hernández, *and* ***M. Aranzabe,*** *University of the Basque Country*

We left San Sebastian, anxious to get to sunny Las Palmas de Gran Canaria. Our bags were filled with sunglasses, aftersun lotion, swimsuits… Everyone was expecting sunshine and our aim was to combine hearing and talking about language resources with finding and testing the beach resources. Unfortunately, the latter was to prove impossible, since we saw the sun for only a couple of minutes a day. It wasn't cold or raining (except for Monday afternoon) but where was the sun? Anyway, we didn't find tropical weather, but we did discover a successful congress, with many interesting presentations and participants.

This year's LREC was extraordinarily productive, with 365 papers and 18 satellite workshops (which was good for science but not for the health – with more than 2175 pages in the six volumes of proceedings, our backs were suffering!)

Kepa Sarasola, Gregoria Hernández, Maxux Aranzabe, and Elena Garcia (back in San Sebastian) with their mountains of Proceedings

It was virtually impossible to attend all the sessions, but we would like to pinpoint two topics of particular interest. Firstly, *Multimodality*, with 25 papers in the main conference and two workshops, is a new emergent subject that hasn't been present at previous conferences. Secondly, *Treebanks*, *Syntactic Annotation*, and *Parsing* (with an uncountable number of posters, presentations, and references).

It is very interesting to see how Multimodal systems are able to integrate knowledge from several areas. Sometimes, working on our small specific tasks, it is difficult to see how our work could interact with other people's. If you'd been feeling like this, LREC 2002 was a good place to be – very interesting and thought-provoking.

The workshops on Syntactic Annotation and Treebanks and the workshop entitled "Beyond PARSEVAL: Towards Improved Evaluation Measures for Parsing Systems" were of special interest as well. With respect to the treebanks, to see the increasing effort made in order to build such expensive and, on the other hand, fundamental tools was really encouraging. And not only treebanks, but all kinds of annotated syntactic corpora. One always thinks that such expensive resources can only be developed for English, but no! It is refreshing to see that this is no longer true. We saw treebanks for Portuguese, Korean, Czech, Russian, and German in the conference and the same multilingual idea came up too, in the "Beyond PARSEVAL" workshop. Here, the bases established in the previous LREC conferences on the evaluation of parsing systems were revised. Different groups expressed their concern about finding suitable evaluation methods for parsing. The goal is to find an evaluation measure that can be applied to all kinds of languages and all kinds of systems. We saw the emergence of something similar to SENSEVAL, but for parsing: in principle, this would be open to any language that could provide a parsing system and an annotated corpus to evaluate.

There was also discussion of collaboration and standardisation tasks. We were introduced to OLAC, the Open Language Archives Community *[see* ELSNews *9.4 – Ed.]*, that aims to create a worldwide virtual library of language resources – see www.language-archives.org (we know this isn't a new community, but we hadn't heard of it before) – and the ISO/TC 37/SC 4 committee led by Nancy Ide and Laurent Romary. This committee aims to prepare international standards and guidelines for language resource management. They asked for collaboration, and it seemed that Nancy got lots of replies, as she was very happy in the Gala Dinner (and was the first to dance!)

Among some LREC participants we detected a new disease – EoI flu – a direct result of an invitation from the European Commission to submit Expressions of Interest (EoI) as an opportunity for Europe's research community to help in preparing the first calls of FP6. In this context, LREC became an excellent environment to meet other researchers. HLT *per se* won't be a specific research subject, and everybody tried joining broader groups including speech, written language, and any other aspect that will give the key to any of the following:

- Applied IST research addressing major societal and economic issues
- Knowledge and interface technologies
- IST future and emerging technologies.

Now that we are all getting ideas for next autumn, what a difficult autumn we are going to have! FP6 flu is coming soon! How many meetings! How many projects to define!

On the Sunday we were leaving, we discovered that the whole of Gran Canaria had been really sunny for the whole week – everywhere except Las Palmas, that is. This seems to be a repeating pattern. Many of the other delegates had managed to visit the rest of the island. What a pity! If we had only known, we wouldn't look as pale as we still do.

ELSNET

Office
Steven Krauwer,
Co-ordinator
Brigitte Burger,
Assistant Co-ordinator
Monique Hanrath,
Secretary
Utrecht University (NL)

Task Groups
Training & Mobility
Gerrit Bloothooft,
Utrecht University (NL)
Koenraad de Smedt,
University of Bergen (NO)

Linguistic & Speech Resources
Antonio Zampolli,
Istituto di Linguistica Computazionale (I) and
Ulrich Heid, Stuttgart University (D)

Research
Niels Ole Bernsen, NIS
Odense University (DK)
and Joseph Mariani,
LIMSI/CNRS (F)

Executive Board
Steven Krauwer,
Utrecht University (NL)
Niels Ole Bernsen, NIS,
Odense University (DK)
Jean-Pierre Chanod,
XEROX (F)
Jose Maria Cavanillas,
SEMA (E)
Björn Granström,
Royal Institute of Technology (S)
Nikos Fakotakis,
University of Patras (EL)
Ulrich Heid,
Stuttgart University (D)
Denis Johnston, BT
Adastral Park (UK)
Joseph Mariani,
LIMSI/CNRS (F)
José M. Pardo,
Polytechnic University of Madrid (E)
Tony Rose, Reuters (UK)
Geoffrey Sampson,
University of Sussex (UK)
Antonio Zampolli,
University of Pisa (I)

Future Events

July 22-23 *International Workshop on Computational Approaches to Collocations*: Vienna, Austria.
Email: colloc02@oefai.at URL: www.ai.univie.ac.at/colloc02/index.html.

July 23 *Workshop on Semantic Authoring, Annotation, & Knowledge Markup (SAAKM 2002)* (in conjunction with ECAI2002): Lyon, France.
Email: handschuh@aifb.uni-karlsruhe.de URL: saakm2002.aifb.uni-karlsruhe.de

July 27-31 *The Fifth Teaching and Language Corpora Conference (TALC 2002)*: Bertinoro, Italy.
Email: talc5@sslmit.unibo.it URL: www.sslmit.unibo.it/talc

Aug 24 – Sept 1 *The 19th International Conference on Computational Linguistics (COLING-2002)*: Taipei, Taiwan.
URL: http://www.coling2002.sinica.edu.tw/

Aug 31 – Sept 1 There a number of workshops being held in conjunction with COLING–2002. Details from COLING-2002 web site (above).

Aug 31 – Sept 1 *The Sixth Conference on Natural Language Learning (CoNLL-2002)* (in conjunction with COLING–2002): Taipei, Taiwan.
Email: Antal.vdnBosch@kub.nl URL: www.aclweb.org/signll/cfp.html

Sept 2-5 *The Seventh International Workshop on Speech and the Computer (SPECOM-2002)*: St-Petersburg, Russia.
Email: specom@iias.spb.su URL: http://www.spiiras.nw.ru/speech

Sept 4-6 *The Sixth Workshop on Semantics and Pragmatics of Dialogue (EDILOG-2002)*: Edinburgh, UK.
Email: edilog@ed.ac.uk URL: http://www.ltg.ed.ac.uk/edilog/

This is only a selection of events – see http://www.elsnet.org/cgi-bin/elsnet/events.pl for details of more events.

Note that the list of ELSNET member nodes, which usually appears in the space, has been held over to make room for the report on the recent LREC on page 15.

What is ELSNET?

ELSNET is the European Network of Excellence in Human Language Technologies. ELSNET is sponsored by the Human Language Technologies programme of the European Commission; its main objective is to foster the human language technologies on a broad front, creating a platform which bridges the gap between the natural language and speech communities, and the gap between academia and industry.

ELSNET operates in an international context across discipline boundaries, and deals with all aspects of human communication research which have a link with language and speech. Members include public and private research institutions and commercial companies involved in language and speech technology.

ELSNET aims to encourage and support fruitful collaboration between Europe's key players in research, development, integration, and deployment across the field of language and speech technology and neighbouring areas.

ELSNET seeks to develop an environment which allows optimal exploitation of the available human and intellectual resources in order to advance the field. To this end, the Network has established an infrastructure for the sharing of knowledge, resources, problems, and solutions across the language and speech communities, and serving both academic and industry. It has developed various structures (committees, special interest groups), events (summer schools, workshops), and services (website, e-mail lists, *ELSNews*, information dissemination, knowledge brokerage).

Electronic Mailing List

elsnet-list is ELSNET's electronic mailing list. Email sent to elsnet-list@let.uu.nl is received by all member site contact persons, as well as other interested parties. This mailing list may be used to announce activities, post job openings, or discuss issues which are relevant to ELSNET. To request additions/deletions/changes of address in the mailing list, please send mail to elsnet@let.uu.nl.

Subscriptions

Subscriptions to *ELSNews* are currently free of charge. To subscribe, visit **http://www.elsnet.org** and follow the links to *ELSNews* and "subscription".

FOR INFORMATION
ELSNET
Utrecht Institute of Linguistics OTS, Utrecht University,
Trans 10, 3512 JK, Utrecht, The Netherlands
Tel: +31 30 253 6039
Fax: +31 30 253 6000
Email: elsnet@elsnet.org
Web: http://www.elsnet.org