Language and Speech Technologies in Natural Interactive Systems

Niels Ole Bernsen, Odense University

*Which technologies are being developed or made available by ESPRIT networks other than ELSNET? And how could they be integrated with, or interfaced to, natural language and speech processing? Niels Ole Bernsen introduces **i3net**, the **European Network for Intelligent Information Interfaces,** and describes its relevance for the ELSNET community.*

Let me first explain why it is a little early to speak on behalf of **i3net** in any detailed or programmatic fashion. i3net was launched in September 1996, to serve members of the European community at large working in intelligent information interfaces. However, its full 'hard core' membership has only recently been added to the network. The reason for this is that i3net will, from June 1997 onwards, also support thirteen research projects. These so-called **i3 projects** were only selected in March 1997; their partners are currently busy preparing their project programmes, and everybody is still waiting for the formal Commission Decision that will allow the projects to start. Moreover, it is only when this work has begun that we will be in a position to invite others to join i3net.

Still, it is generally clear where the emphases of i3net and i3 projects lie. First, there is an emphasis on innovative human-centred interfaces for interacting with information for the broad population. This implies a focus on computer systems and interfaces that are easy and intuitive to use, as well as a focus on providing support for ordinary, nonprofessional activities. Secondly, i3 has a definite emphasis on social computing, or computing systems for communities of users, rather than on the single user-single computer paradigm. And thirdly, most i3 projects aim at developing running research prototypes with significant involvement of communities of users, thereby ensuring prototype realism and usability.

Let us take a brief look at the i3 projects to see how each of them aim to explore those general directions. Based on the evaluation report, three groupings of i3 projects have emerged. The first consists of five projects (**eRENA**, **eSCAPE**, **POPULATE**, **AMUSEMENT** and **PERSONA**) which address issues of networked virtual environments; these include interactive TV, large-community participatory entertainment, the linking of different virtual environments, 3D avatar booths and interactive multimodal personalised information seeking. The second grouping has three projects (**MAGIC LOUNGE**, **CORMIS** and **HIPS**) which address mixed or augmented realities — including virtual meeting places for geographically dispersed communities; speech-based Web search; and wearable/portable assistants for (for example) conferences and local tourism. The third group, which consists of five projects (**CAMPIELLO**, **MAYPOLE**, **LIME**, **CO-NEXUS** and **PRESENCE**), looks at intelligent information interfaces for local communities, such as electronic whiteboards; large screen interactive TV; tourists meeting the local community; interfaces for the sharing of local

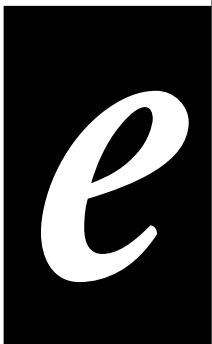*Ole Bernsen at the most recent ELSNET EB meeting*

activities; shared local community memory systems; and self-learning agents representing individual user interests.

This brief description already shows that all these projects are interrelated at many different levels; and this was further demonstrated at the joint i3 kickoff meeting in Brussels in late March 1997. i3net is currently mapping out these interrelationships in order to support them through the creation of cross-project **Special Interest Groups**.

On the face of it, i3 may not appear to have much in common with ELSNET's long-term goal, namely

> *to build integrated multilingual speech and NL systems with unrestricted coverage of both spoken and written language. (ELSNET-2 Network Programme, June 1996)*

And it is true that the focus of i3 lies elsewhere. However, the list above mentions both speech-based Web search and tourists. We know that tourists often do not speak the local language of the country they're visiting terribly well. So tourists would be interested in multilingual computer support for their activities, whether they want to have information of the standard tourist varieties or whether they want to get in touch with the local community. Users of the self-learning agents and the personalised information seeker might have similar needs for multilingual support when retrieving information in foreign languages, and they would also be likely to welcome innovations such as

summaries of large amounts of textual information.

So when we look more closely, a much broader picture emerges, which could amount to a vision of the future. And ELSNET, i3net and many more actors worldwide may already be contributing to this vision, irrespective what they believe they are doing.

What does this 'vision' look like? Computing systems will no longer be considered mere tools which their users somehow get to execute. Instead, a computing system will be a combination of a tool and a communication partner. In the old days, or in the traditional scenario, humans used to solve problems together using memos, planning documents, blueprints, physical models, timetables, telephone calls, faxes, reports and so on. A typical scenario would involve two or more persons doing problem-solving together, in one place; in that process, the people involved would draw upon their combined knowledge and skills, using all sorts of external information sources to support their work, and communicating in natural human ways, through speech, pointing, gesturing, facial expression, emotion, sketching, demonstrating, and so on. In other words, the traditional scenario is characterised by: natural human-human commu–

nication, in one place, with limited access to external knowledge.

In this vision of the future, human problem solving will involve people being together in virtual space; they would use the system as an almost all-knowing tool they can communicate with, and which may well be present on the wireless, screen, or whatever, as a virtual human. So the future scenario is characterised by natural 'human-human-machine' communication, ubiquity, and unlimited knowledge access. Activities of people in these future virtual spaces will of course not be limited to joint problem solving; they will also include entertainment or simple conversation; and enhanced realities and "real" local communities offer additional challenges. But these things merely add more facets to the vision.

I do not believe that this vision is just one possible future scenario, or even an arbitrary daydream: it appears to be the *necessary* end-point of the technological developments we are witnessing right now. Arguably, no other end-point makes sense.

The implication of all this is that i3net needs ELSNET's technologies, and increasingly so; that ELSNET will increasingly need those of i3net; and that both will increasingly be looking for technologies from computer vision, machine learning, advanced graphics,

acoustics and haptics, augmented reality and elsewhere. Moreover, all actors will be involved with multimodality, human factors, user-centeredness, human communication and behaviour, and user adaptation. The number of application domains for which these technologies will be useful is unlimited.

The vision of natural interactive systems I have sketched here gives rise to many questions, of course. One of the most interesting (and pressing) ones is: what can the networks do to accelerate progress towards this goal, and to increase the efficiency with which it can be achieved?

Call for Propopals
ELSNET Bullet Courses

ELSNET Bullet courses are short intensive courses which aim to enhance knowledge transfer between academia and industry, and are geared primarily towards industrial participants. To encourage activities in many different countries, and to make it possible for many different Member Nodes to be involved, ELSNET would like to co-organise future bullet courses with Member Nodes.

Member Nodes are therefore invited to submit proposals for topics and local organisation of future ELSNET Bullet Courses. Topics should be industry-oriented (i.e. concern technology and products that can be used immediately), and the targeted audience should be well-defined, in order to allow successful publicity. Although courses will in principle be international, they may also be offered in the local language if the target audience commanding this language is big enough.

ELSNET Bullet Courses are in principle self-financing. The first course (on Spoken Dialogue Systems), which took place from 16-18 April 1997, has shown that the formula is practically and financially viable. Both potential financial risks and profits will be shared by ELSNET and the local organiser. However, the local organiser will be able to make use of ELSNET's infrastructure, experience and administrative support. In return, the whole ELSNET community will benefit from the specific locally available content and/or market expertise brought in by the co-organising Member Node.

ELSNET plans to organise its second bullet course in the spring of 1998, and to turn it into a bi-annual event from then on.

For more information and an information package, please contact **Yvonne van Holsteijn**, ELSNET Assistant Coordinator (elsnet@let.ruu.nl).