DISC – Overview of Results

Niels Ole Bernsen and Laila Dybkjær
Natural Interactive Systems Laboratory (NIS), Odense, Denmark
nob@nis.sdu.dk, ld@nis.sdu.dk

Spoken language dialogue systems (SLDSs) are becoming mainstream technologies which provide a wealth of different services in an increasing number of languages. In general, SLDSs are complex systems which incorporate speech recognition, speech generation, natural language understanding and generation, dialogue management and database components. The building of successful applications not only demands expertise in those technologies but also in systems integration, human factors, design and development support tools and, increasingly, call centre technology. Difficult research problems remain to be solved in most of the areas mentioned above, and best practice guidance is notoriously absent in the field of dialogue engineering.

The DISC project - spoken language DIalogue Systems and Components - best practice in development and evaluation - was launched in June 1997 with the aim of developing a first Best Practice Guide on how to develop and evaluate SLDSs and their components (cf. Elsnews 6/4 1997, 6-7). Ending in December 1998, DISC was immediately continued in DISC-2 which ran until the end of 1999. The first phase of DISC was dedicated to the development of current practice reviews of the DISC SLDS aspects (see below), a detailed best practice dialogue engineering development and evaluation methodology, and a range of design support concepts and software tools. The second phase, i.e. DISC-2, focused on testing the validity and usability of the draft Best Practice Guide, the concepts and the tools, and on the integration, packaging and dissemination of the final DISC Best Practice Guide. Throughout DISC, the Advisory Panel which by the end of 1999 counted more than 50 researchers and industrial developers from across the world, has been an invaluable source of critique and comments on progress.

The DISC current practice reviews charted current SLDSs development and evaluation practice, producing about 50 in-depth analyses of existing SLDSs and components together with the following approach to dialogue engineering best practice. An SLDS is viewed as having six major *aspects:* speech recognition, speech generation, natural language understanding and generation, dialogue management, human factors, and systems integration.

Each aspect can be analysed in terms of a *'grid'*. A grid contains an aspect-specific description of the state-of-the-art technical problem space facing the developer, including technical properties, interrelationships among properties, and advice on which properties to include in particular applications. Within the grid problem space – or outside it, since new ideas appear all the time - the developer must make the decisions most appropriate for the application to be developed. In DISC, the grid problem space is structured in terms of the *issues* facing the developer, the *options* the developer must choose from per issue, and the *pros and cons* with respect to each option.

Orthogonal to the "static" grid description, each aspect may be analysed in terms of a development *life-cycle* which decomposes the development process into iterative phases and issues to be addressed in each phase. Integral to the life-cycle is the continuous *evaluation* of progress and results. As DISC progressed, evaluation of SLDSs aspects gained prominence due to the many unsolved research issues in this field. In response, DISC has developed a generic *evaluation template* which can be used to characterise each evaluation criterion for

use in evaluating aspect-specific properties of SLDSs. Furthermore, the grid analyses have been used to systematically generate a set of evaluation criteria per aspect.

In addition to the above, best practice guidance must incorporate guidance on available platforms, methods and supporting tools per aspect. These have been surveyed in a series of DISC reports. Moreover, DISC has itself produced a series of development support tools and guidelines. These include:

- guidelines and testing protocols for the development of speech recognition components for SLDSs;

- a software tool for evaluating speech synthesis components in SLDSs;

- guidelines for the acquisition of lexical data for SLDSs;

- CODIAL, a software tool in support of cooperative system dialogue design;

- SMALTO, a software tool in support of speech functionality (pertaining to what the speech modality is (not) good for) decisions in early design.

The core result of DISC is the web-based DISC Best Practice Guide (www.disc2.dk) which resulted from turning everything mentioned above into a comprehensive website. In addition to the ingredients described above, the DISC Best Practice Guide includes a comprehensive *glossary* of dialogue engineering terminology, *references* to the literature including all DISC publications, and brief *checklists* per aspect.

Due to generous support from ELSNET, the DISC website will continue to be updated in the future. The DISC consortium, therefore, welcomes all comments from those who have tried out any part of the DISC Best Practice Guide. You may email the authors of the present article who will make sure that your comments reach the right people. However, the DISC website offers several other possibilities of interacting with the DISC consortium which includes NIS (Odense, Denmark, coordinating partner), CNRS-LIMSI (Paris, France), IMS (Stuttgart, Germany), KTH (Stockholm, Sweden), Vocalis Ltd (Cambridge, UK), Daimler-Chrysler AG (Ulm, Germany) and ELSNET (Utrecht, The Netherlands).