STRUCTURING THE DESIGN SPACE

Niels Ole Bernsen, CCI, Roskilde University and Risoe National Laboratory

Summary: The paper presents ongoing work on a usability engineering framework, called CO-SITUE, for structuring top-down the design space surrounding the design of computer artifacts. CO-SITUE involves, in addition to general constraints and criteria on the design process, the following aspects of design spaces and artifacts: collaboration, organisation, system, interface, task and task domain, user, and user experience. To illustrate the use of CO-SITUE in an actual design task, the framework has been applied as a notational frame for analysing the initial artifact specification phase of a spoken natural language dialogue system. The framework is viewed as complementing approaches to the analysis of design reasoning such as the Design Rationale approach to design space analysis using Questions, Options and Criteria.

Keywords: Design space representation, artifact design, designer reasoning, design rationale.

1. Introduction

Much has been said and written on the complexity of the problem facing designers of advanced information technologies and applied scientists of human-computer interaction (HCI) alike. The overall conclusion is that there is a huge gap between the complex, real life diversity of users, systems, work contexts, tasks, etc., on the one hand, and anything which has so far been proposed to provide principled, predictive support for the design of usable artifacts on the other. The dilemma for HCI is the following: either such support actually does address in detail part of the complex, real life diversity mentioned, but then its predictive scope turns out to be quite narrow and cannot easily, if at all, be generalised for broader coverage; or the support is maintained at a general level of description in order to maintain broader coverage, in which case its applicability for predictive purposes to concrete design processes is shallow.

This gap may well be unbridgeable. Prescriptive design guidelines will never bridge it and have lately fallen into disrepute exactly because bridging attempts based on them tend to result in an unmanageable complexity of rules and heuristics. The theories and systems available for predictively supporting the detailed design of usable artifacts each have a very limited domain of application and do not lend themselves easily to extensive generalisation (e.g., Card, Moran and Newell 1983, Kieras and Polson 1985, Payne and Green 1986). The basic science of cognitive psychology has proved of less use to HCI than many used to believe a few years ago. There are two main reasons for this. Firstly, the necessary scientific foundations are not in place. For most usability engineering problems, the relevant support from scientific theory does not currently exist. Secondly, even if the scientific foundations were there, the lessons from the relationship between physics and engineering teach us that applied science is much more complex than simple deduction from basic principles to some real life (design, engineering, or otherwise) problem at hand. Realistic problems exhibit a complexity and variability which far surpasses the conceptualisations of basic theory.

The physics/engineering example suggests, however, that more principled support for usable artifact design in information technologies might be achievable than is currently the case. To investigate ways and means of providing such support, many HCI scientists are turning towards deeper analyses of full-scale design processes (cf. the papers in J.M. Carroll et al. 1991). The understanding gained is intended to be fed back into requirements for new developments of the science base of HCI, new ways of representing the knowledge in the science base for the purpose of predictively supporting design processes and, ultimately, operational support tools based on science (Barnard 1991). This situation forms the overall context of the Esprit Basic Research project AMODEUS II. The applied science of HCI is no longer viewed as being a matter of directly importing the science base into the design process. In view of the complexity of realistic problems one has, at least for the time being, to be more modest. Predictive support for usability during early design may be a matter of combining designer craftsmanship with (a) advanced rapid prototyping methods involving empirical user testing, (b) focused application of science-based methods where available, be it through consultancy, walk-through support, development of computer-based tools or in other ways, (c) use of relatively low-level and robust conceptual frameworks which can be turned into computer-based tools, and (d) communication with designers through example design studies in which usability issues are analysed in depth, the idea being that 'intuitive generalisation' performed by designers themselves will do some of the work that scientific theory is currently unable to accomplish.

The present contribution addresses points (c) and (d) above in presenting first steps towards the development of a robust, top-down framework for characterising the design space within which software information systems designers work, and starting at the coarsest level of description. The paper is a path-finder document for an idea rather than the presentation of a mature theory. In section 2, an iteration is made over the issue of what is being designed and the CO-SITUE framework is presented as a coarse-grained analysis, for the purpose of supporting usability engineering, of a central part of the design space within which artifact design takes place. Section 3 presents the full set of components of the CO-SITUE framework and an empty CO-SITUE frame. Sections 4 and 5 describe the application of the CO-SITUE frame notation to the first phase of the design of a particular information service system, i.e., a spoken language dialogue system. Section 6 concludes the analysis and relates the work to another approach currently under development in AMODEUS II, i.e. the Design Rationale framework for design space analysis using Questions, Options and Criteria.

2. CO-SITUE and the Design Space

The question "What is being designed?" may appear trivial but in fact is not. One may want to say that, obviously, what is being designed are systems or, if one prefers, systems and user interfaces intended to facilitate human performance on specified tasks. This is false. We have learned that normally if not always, the introduction of information technologies as intermediaries between users and their tasks and work domains change the tasks themselves. Systems and interfaces are not simply being tailored to pre-specified tasks. So what *is* being designed? Let us call what is being designed an artifact. What is an artifact? It turns out that artifacts are much more complex entities than suggested above and this fact seems to have implications for the understanding of design practice and for the nature of the design support approaches which are feasible in HCI.

In addressing the question of what is an artifact, let us look at CO-SITUE first. CO-SITUE stands for the following aspects of the artifact which is being designed:

C = Collaborative aspects.
O = Organisational aspects.
S = System aspects.
I = Interface (or more generally: system Image) aspects.
T = Task aspects including task domain aspects.

U = User aspects.
E = User experience aspects.

User performance on the designed artifact will be a function of all these aspects, i.e.:

UPERF = f(C,O,S,I,T,U,E)

Put an artifact designed for one organisational setting into another, different organisation and the artifact may not be used any more. Or change the user population from occasional users into full-time professionals and they may end up being highly frustrated when using the artifact. User performance is the decisive measure of the usability and efficiency of the artifact as a tool for human work. The basic claim behind CO-SITUE, therefore, is that in designing a 'system' what is actually being designed is not merely a system but something much more complex. What is being designed is a particular CO-SITUE complex. Or, what is being designed is an artifact and artifacts are CO-SITUE complexes.

It is important to note from the outset that there are other generic constraints on design tasks than those designated by CO-SITUE. The overall design space framework to be presented in this paper has three main components, i.e., CO-SITUE, the general design goal and a number of general constraints and criteria on the design process. The latter two components are presented in Sect. 3 below. For the moment, let us consider the question whether artifacts are 'really' CO-SITUE complexes. The justification for claiming that this is at least approximately true is the following. Information systems artifacts have an *interface* at which *users* accomplish their *tasks*. In the case of interactive information systems, task accomplishment is interactively shared between user and system and this will be marked by the introduction of a distinction between *system's task* and *user's task* below. Intended users' *experience* is an important parameter in artifact design. *System* performance is a major parameter in early artifact design. Many information systems artifacts are being designed to fit into *collaborative* schemes of work and many artifacts, even though they are not intended for collaborative use, are designed to fit into specific *organisational* frameworks. Each of these CO-SITUE aspects obviously have their own intrinsic complexity. The claim here is that any analysis of design space structure, however coarse-grained, will at least have to include the CO-SITUE aspects. This is also true of design space structure analysis from the particular point of view with which we are concerned, namely, usability engineering. The aim of usability engineering is the optimisation of user performance and the artifact being designed is analysed with this aim in mind. This implies that, e.g., interface technicalities or system technicalities are only relevant to the analysis to the extent that they make a difference to user performance. Otherwise, they are left out of consideration.

During the design process, designers have to consider and actually do consider to some extent and at various levels of detail all the types of aspect relevant to artifact design summarised in CO-SITUE. The artifact is designed the way it is in order to satisfy multiple criteria and constraints an important part of which are derived from a consideration of the CO-SITUE aspects of the artifact. The claim here is not that designers do consider all possible CO-SITUE aspects of the artifact, or all relevant aspects, at appropriate levels of detail, or that they consider the aspects which they do consider in any systematic fashion to make sure that no relevant aspect has been overlooked and left unanalysed. In fact, designers currently have no way of making sure that this happens. Instead, they work with personalised stopping rules and evaluation functions (Goel and Pirolli 1992). The claim is rather that designers actually work within the conception of an artifact designated by CO-SITUE; that is, *CO-SITUE describes an important part of the overall shape of the design space around an artifact during design.* If that is at least approximately true, then CO-SITUE does link up with actual design reasoning and hence might provide common ground among designers and HCI scientists on which to pursue principled and predictive approaches to usable artifact design.

There are two main points involved in the claim that what is being designed is a particular (CO-SITUE) *complex*. Firstly, during the design process the various aspects of CO-SITUE constantly interact. This means that neither normatively nor in actual design practice is there such a thing as first specifying a system and then looking at user requirements, or vice versa, or first specifying a system and then looking into user tasks, or vice versa, or first specifying a system and then specifying its interface. In design reasoning, multiple constraints derived from very different aspects of the artifact to be designed are simultaneously and continuously brought to bear within the design space and user performance on the resulting artifact is a function of the sum of the design decisions made. The resulting physical artifact is an embodiment of a specific CO-SITUE complex. It represents a particular set of commitments within the design space. Which one, cannot be determined from the physical artifact itself through 'reverse engineering' because of the complexity involved. The CO-SITUE frame notation has been designed to capture the specific set of commitments which constitute the emerging artifact during design and are relevant to usability engineering (see below).

Secondly, there is an important sense in which designers design, not only systems and their interfaces but also collaborative and organisational schemes, tasks, users and the degree of experience of users. Usable artifact design is not a one-way traffic of optimising the way constraints from those domains influence the usability of the resulting artifact. It has already been noted by several that computer artifacts change users' tasks (e.g., Norman 1991). However, the point about change is more general than that. As viewed *from within* the dynamical and creative design process, all or most aspects of CO-SITUE are potentially subject to change as a result of design decisions. The reasons are that (1) computer artifacts change work conditions: they change not only tasks but very often also collaborative and organisational schemes of work; (2) artifacts 'change' users in the sense that the types of target end-users and the requirements on their knowledge and experience for dealing with the artifact are themselves to some extent variable design options; and (3) that, rather more obviously, during the design process the actual system to be designed and its interface are themselves variable design options.

In other words, when designing an artifact one constantly has to identify and select between options concerning not only the way the system is to be built, or the way the interface is to be built, but also concerning the possible ways to change the organisational and collaborative work schemes of users, the possible ways to change their tasks, and the possible ways to select the types of end-users and the knowledge and experience they will need for operating the artifact. The artifact which evolves during the design process is a result of design decisions which are normally concerned with all aspects of CO-SITUE. The design space is a field of possibilities or potential changes, and the scope of those changes is CO-SITUE as a whole. Thus, for instance, in an experimental design process where the goal is to generate ideas for concrete applications of a new technology, one may start by designing an artifact with which naive users can easily accomplish tasks a, b and c, and end up with an artifact with which only trained users can accomplish tasks b, c, d and e using an interface very different from the one which was originally intended and requiring a rather different collaborative and organisational context. It is true that most design processes are more constrained than that but this does not affect the general point.

One common, current idea for producing relatively low-level HCI science support of design processes is that of developing taxonomies for tasks, systems, interfaces or users. Such taxonomies might be useful, it is argued, since locating an artifact-to-be-designed within them might facilitate property inheritance from the set of properties characterising a specific type of system, interface, user or task within a particular taxonomy. However, if what is being designed during artifact design are CO-SITUE complexes, it follows that the research goal of constructing *independent* taxonomies for tasks, systems, users, interfaces or possibly also for collaborative or organisational schemes of work, represents a misguided approach to design-applicable HCI. In the design context, tasks, interfaces, users and systems can only be described and analysed with comprehensive reference to one another. It is, therefore, not surprising that no such taxonomies have yet been devised by HCI researchers nor proved successful to design practice. Indeed, the prediction is that such taxonomies are not only misguided, but impossible to construct unless they start combining features of users, tasks, interfaces, etc. (compare Brooks 1991).

The above claims should not be misconstrued as stating that every aspect included in CO-SITUE is a legitimate variable during the design process. CO-SITUE itself is not, nor are, for instance, general aspects of users' cognitive architecture. These are *constants* or *invariants* rather than variables which designers can only ignore at their peril. Bad design often results from overlooking such invariants, from users' limited working memory to the fact that doctors do not want to transfer responsibility to medical expert systems. The existence of such invariants opens up the possibility that further articulation of invariant CO-SITUE aspects of the design space might provide increased support for artifact design (see Sect. 6 below).

If CO-SITUE provides an approximate, coarse-grained characterisation of a central part of the design space surrounding an artifact during design, the design process itself becomes one of making explicit the constraints relevant to the particular artifact being designed and trading them off against one another in handling the design options that become apparent. This way, CO-SITUE gives rise to the definition of interacting design commitments at increasing levels of detail which can be represented in a frame notation supported, whenever necessary, by relevant design documents. The following three sections describe how this process worked during the initial specification phase of a spoken language dialogue system prototype. The prototype is the first of two experimental prototypes of increasing sophistication designed for eventual product development by a Danish research consortium including the Speech Technology Centre, Aalborg University, the Centre for Language Technology, Copenhagen University and CCI. The system will allow users to book flights, change flight reservations and obtain information on Danish domestic flights using spontaneous spoken natural language over the telephone. The design of the first prototype has aimed at ensuring that a usable system can indeed be built given rather hard technological constraints. The design of the second prototype will aim at improving the naturalness of the system. A full account of the initial specification phase is given in Bernsen (1993). Project results are described in Larsen et al. (1992), Larsen et al. (1993), Dybkjaer and Dybkjaer (1993) and Dybkjaer, Bernsen and Dybkjaer (1993). The description below focuses on usability aspects and is analytic (or *post hoc*) rather than temporal and historical. It shows that the artifact is indeed a CO-SITUE complex and how the initial design specification phase can be represented as a limited set of interacting commitments at increasing levels of detail. CO-SITUE has yet to be applied *ab initio* to the detailed temporal structuring of design processes. The analysis reported in this paper was actually used as the basis for the subsequent knowledge acquisition phase of the design project and in defining the knowledge acquisition model (Bernsen 1993, Dybkjaer and Dybkjaer 1993).

3. Overall Design Goal and Design Constraints

As said earlier, the overall design space framework to be presented has, in addition to CO-SITUE, two more components. The design task has an overall design goal and a number of generic design constraints. The *overall goal* of the design task is to develop a state-of-the-art spoken language dialogue system prototype operating via the telephone and capable of replacing a human operator. The *generic design constraints* are the following: The design task is constrained, firstly, by limitations in cost, time, manpower, available machine power and so on. There may also be relevant legal, social and political constraints. These are *general feasibility constraints*. Secondly, the design task is constrained by *scientific and technological feasibility*. Thirdly, the artifact to be designed is subject to a number of important conditions concerning its *realism, usability and naturalness.* These conditions help shaping (or constraining) the initial design space within which subsequent design decisions are to be made. The distinction between realism, usability and naturalness can be viewed as a somewhat more structured alternative to the ill-defined notion of 'system usability' or 'system habitability'. Finally, it is valuable for the sake of making the design process transparent to include a fourth type of design constraint, namely, *designer preferences*. What happens during the design process, then, is that the generic design constraints are applied, i.e. interpreted and traded off against one another in a process of interpretation, discovery, justification, trade-off and decision-making, with respect to the overall design goal of designing a particular CO-SITUE complex. The implemented physical artifact is the eventual outcome of this process. Since CO-SITUE aims to support usability engineering rather than design reasoning in general, the constraints from general feasibility, scientific and technological feasibility, and designer preferences will in general only enter into the CO-SITUE representation when they are judged relevant to usability issues during design.

Here follows an empty version of the CO-SITUE notational frame. The distinction between general constraints and criteria (A) and their application within the design space (B) amounts to a distinction between general constraints and criteria, on the one hand, and interpretations of the general constraints and criteria and trade-offs between them with respect to the particular artifact under development (or with respect to the overall design goal), on the other. The frame includes the distinction already mentioned between system's task (T(S)) and user's task (T(U)).

CO-SITUE No. (0)

A. General constraints and criteria
Overall design goal:
General feasibility constraints:
Scientific and technological feasibility constraints:
Designer preferences:
Realism criteria:
Usability criteria:
Naturalness criteria:

B. Application of constraints and criteria to the artifact within the design space
C =
O =
S =
I =
T(S) =
T(U) =
U =
E =

C. Hypothetical issues:
D. Conventions:
C = Collaborative aspects.
O = Organisational aspects.
S = System aspects.
I = Interface (or more generally: system image) aspects.
T(S) = System Task aspects including task domain aspects.
T(U) = User Task aspects including task domain aspects.
U = User aspects.
E = User experience aspects.
CO-SITUE No. ( ) indicates the number of the current CO-SITUE specification.

---

4. CO-SITUE Exemplified

Let us apply the CO-SITUE framework analytically to the initial specification phase for a spoken language dialogue system. The notational frame itself is indifferent to the order in which information is entered into it. The reason is that, in general, no particular order can be normatively imposed on the way in which frames have to be filled during initial artifact specification. In the present case we are dealing with general, problem-structuring constraints on the design task. Design reasoning at this stage consists in discovering the constraints, developing their implications and making design decisions on this ground. Each decision represents the binding of a variable leading to a more specific conception of the artifact being developed. Let us start with a number of constraints which seem rather straightforward, such as the overall design goal and the following:

Collaborative and Organisational aspects

Many artifacts are to be introduced into collaborative and organisational work contexts in which it is important to consider the possible changes needed in collaborative and organisational schemes of work in order that appropriate and efficient use be made of the artifact. If such changes are not properly analysed and evaluated, the artifact may eventually not be used at all or may not be used optimally by its intended users. Some artifacts, such as the one described here, may not initially appear to have any clear collaborative or organisational aspects to them. If analysis of the work domain shows this to be the case, then there is no need to take collaborative and organisational aspects into consideration during the design process. As for the current design task, a first analysis of this kind was conducted as a field study in a professional travel agency. A successor to the artifact being designed may eventually end up being used by the agency workers and it has already been established that its introduction would cause changes to their organisational and collaborative schemes of work. These changes must be taken into account eventually. However, it will be assumed in what follows that we are dealing with an 'individualistic' artifact (a SITUE complex).

Experience and Interface aspects

The *E* aspect of CO-SITUE can be comparatively simple. In the case of our spoken language dialogue system prototype, the design decision was to initially aim at novice users in order to create a fully self-explanatory system. Unless ways were found later to accommodate them, experienced users would therefore have to accept an artifact designed for novice users even though this might force them to listen to (for them) redundant information from the system during each occasion of use.

At the most general level of description, the *I* aspect of CO-SITUE is simple and fixed throughout the design process since we are dealing with spoken dialogue over the telephone. Thus we have:

CO-SITUE No. (1)

A. General constraints and criteria

Overall design goal:
- *Spoken language dialogue system prototype operating via the telephone and capable of replacing a human operator;*

General feasibility constraints:
Scientific and technological feasibility constraints:
Designer preferences:
Realism criteria:
Usability criteria:
Naturalness criteria:

B. Application of constraints and criteria to the artifact within the design space:

C = *Null*
O = *Null*
S =
I = *Spoken telephone dialogue*
T(S) =
T(U) =
U =
E = *Novices*

C. Hypothetical issues:
- *How to accommodate experienced users?*

D. Conventions:
Italics indicate new elements in the design specification (No. n) as compared with the immediately preceeding specification (No. n-1).
"Null" means that the artifact does not embody a certain aspect of CO-SITUE.

The *C, O, I* and *E* components of *CO-SITUE* have now been fixed to the extent possible at this stage. During the remainder of the design process, *E* fixed as *novice* continues to exert its influence on design decisions unless it gets changed to something else because of novel considerations which might help resolve the hypothetical issue of how to accommodate experienced users. If changes are made to *E*, the changed *E* will then exert its constraining influence on the design process. As a matter of fact such a change happened later during the early knowledge acquisition phase when a promising option was discovered concerning how to make the artifact adapt on-line to experienced users. This illustrates how, during the design process, E is itself being designed and re-designed. CO-SITUE aspect *I* having been fixed as spoken telephone dialogue, on the other hand, will remain fixed throughout the design process. This also illustrates how design commitments, or the assignment of values to variables in the design space, act as constraints on subsequent design reasoning and design decision. The issue about which design decisions are *really* fixed and which ones are more or less fixed or temporary is not important. The significant point is that the notational frame makes explicit all important design constraints and decisions relevant to usability engineering.

5. Criteria and Preferences

During initial design specification we want to identify the type of artifact to be designed, specify important aspects of its performance, identify design tools and basic technologies for incorporation into the artifact, and develop the broad overall system architecture. We also want to prepare subsequent phases of the design process. As regards the focal issue of usability, the current design task requires consideration of the linked criteria of *realism, usability and naturalness*. These are general criteria to be used throughout the design process in evaluating different design options prior to the making of design decisions. The following description demonstrates a number of cases in which the criteria of realism, usability and naturalness support the making of early design decisions, sometimes in conjunction with other kinds of constraint on the design process. The net result is a gradual shaping of the design space from the top down.

Realism

The artifact to be designed should be *realistic,* that is, it should demonstrate the use of the generic technology (i.e., spoken language dialogue systems) in a task domain in which such systems might actually be superior, equivalent or at least acceptably inferior to other known ways of performing similar tasks. Among the preferred task domains for experimental spoken language dialogue systems today are time table enquiries for flights and trains. It was decided to select one such domain, i.e., that of information on flight travels and reservation of tickets to be obtained over the telephone. This task domain is realistic in the sense that, in all or most countries today, the telephone is already being used extensively for the tasks indicated. In principle, the relevant tasks might be mechanised instead by having users perform long series of telephone keystrokes or through the use of screen, keyboard and mouse. However, the first option implies rapidly decreasing usability as the task-relevant dialogues grow in complexity, and the second option is not (yet) available to most potential users.

In summary, then, the realism of the artifact to be designed derives from the facts that it (a) addresses real and known user needs, (b) is preferable to other technological solutions currently available and (c) is assumed to become, from the users' point of view, tolerably inferior in performance to humans performing the same tasks. It is possible that these realism criteria are valid for all artifacts designed to replace humans on some task or set of tasks. Criterion (c) is expanded in the usability criteria below. Note that criterion (a) makes an assumption about user needs. Note also that if the realism criteria cannot be met then the system will be practically useless. This means that realism criteria are normally basic to the decision whether to design a system at all. Finally, we have seen how the realism criteria were interpreted with respect to the task domain of the artifact. The general realism criteria themselves do not state anything specific about the task domain.

Usability

Roughly speaking, usability criteria serve to ensure that the artifact *can do* the tasks done by the human it replaces or, speaking more generally, that the artifact can serve the tasks it is being designed to serve. This means that if the usability criteria cannot be met then the system will be practically useless. Strictly speaking, usability in this sense sometimes is a matter of degree. However, it seems useful to think of usability criteria as conditions which *unconditionally* have to be met by the artifact, thereby contrasting them with naturalness criteria (see below). This way of conceiving of usability criteria emphasizes the importance of being able to identify all usability aspects of the artifact during early design. It turns out that many specific usability criteria can be self-evidently derived from or interpreted with respect to the design goal. Usability criteria normally are basic to the decision whether to design a system at all, unless we are dealing with pure research systems.

In view of the importance of usability to HCI, and to illustrate the actual contents of the design task, the interpretation of usability with respect to the design goal is presented rather comprehensively in what follows. A usable spoken language dialogue system should meet at least the following requirements, most of which are rather self-evident implications of the overall design goal in conjunction with assumptions about users and tasks:

- the system should *understand* all or most of the users' utterances in their appropriate task context as evidenced from its responses to users. In cases where the system fails to understand an utterance, it should be able to repair the dialogue through appropriate responses to the user. This feature of the system is often called "robustness". It is self-evident that if a system is not sufficiently robust in this sense it will not be usable;

- the recognised*vocabulary* should be large enough to encompass all or most terms relevant to completing the dialogues necessary for the chosen tasks. If the vocabulary is too limited, then users will have an unacceptably difficult time trying to get the system to do what they want;

- the system's *grammar* should be natural to users, i.e., the system should recognise and understand the varieties of syntactic forms that users find it natural to use during their dialogue with the system. It is a fact of cognition that it is practically impossible for users to remain within the bounds of an unnaturally restricted syntax during natural dialogue. This point may be less self-evident than others on this list and hence might be overlooked by designers who are not familiar with the literature on earlier spoken language dialogue systems or with relevant principles of human language processing;

- the system's handling of *discourse phenomena* should be natural to users during their dialogue with the system. The principle behind this condition is the same as in the case of grammar above. The point about lacking self-evidence made in that context holds true here as well;

- the system should *respond to user input* in something not too remote from real time. This is evident from practical usability considerations even without invoking cognitive theories of attention span;

- the system should do*speaker-independent* recognition of speech as usability in the chosen task domain would be damaged by requiring the training process which is otherwise needed for spoken language understanding systems to adapt to new users. In other related task domains (e.g., voice process control in the flight cockpit), speaker-independence is less important;

- the system should preferably do *continuous speech recognition* as constraints on users' sentence pronounciation seriously affect the usability of the system in the task domain. Again, in other related task domains this requirement may be less important;

- the system should clearly communicate to users *which tasks* they can accomplish with the system in the chosen task domain and the system should possess the *task domain information* necessary for users to accomplish those tasks. Since the system's performance will be limited as compared to that of human operators, users obviously have to be told what the system can and cannot do. And when the system announces it capability of doing something, it should be able to do it.

Although this list is (mostly) self-evident and states conditions necessary to usability, there does not currently seem to exist any procedure for making sure that it is sufficient. For this reason, moreover, it is possible for different designer teams, depending on their collective know-how, to derive somewhat different sub-sets of the necessary conditions for system usability. This, indeed, seems to happen all the time. This is clearly an area in which principled support for usability engineering would be beneficial. Note that the current design process is in some sense a limiting case since the *C* and *O* aspects of CO-SITUE have been set to null and the *I* aspect is extremely simple. In the general case, usability may be affected by all aspects of CO-SITUE.

Let us take a closer look at how the above commitments on usability are derived. The usability criterion is *minimal* in the sense that it states conditions necessary for the system to be at all able to replace a human operator in the task domain. The design commitments derived from it are derived from the overall design goal specification in conjunction with task domain information and information on user needs and users' cognitive capabilities and limitations. An important cue as to how this derivation takes place is the following. In general, when a system is usable, it *can do* the tasks done by the human operator it replaces. This may not be sufficient to ensure ultimate usability (cf. the point made about collaborative and organisational aspects above) but it certainly seems to be necessary. In other words, *to determine specific usability criteria for an artifact during early design, one has to identify conditions such that, if they are not met by the designed artifact then the artifact cannot do the tasks assigned to it given the users assigned to it*, or, in the general case, given CO-SITUE as a whole. CO-SITUE may thus provide some support for the derivation of usability commitments in the design space.

The 'meta-criterion' just stated for the identification of specific usability criteria for a given artifact may provide some support during early design. But this support is not sufficient even with respect to user and task-related aspects of usability. Since, as indicated, important assumptions about the cognitive capabilities and limitations of users may be relevant, the expertise needed for deriving a reasonably complete set of usability criteria during early design is not necessarily present in the designer team and strong penalties may have to be paid later if some crucial usability criterion has been overlooked. Some way of testing, as early as possible in the design process, the cognitive assumptions about users underlying the design of a particular artifact is therefore mandatory. The same applies to other aspects of the artifact denoted by CO-SITUE. CO-SITUE is intended to provide support for the derivation of usability commitments and identification of usability issues which may be in need of empirical testing or other forms of investigation, and a further articulation of the basic structural properties of the design space within which system design takes place may yield additional support (see Sect. 6 below).

CO-SITUE No. (2)

A. General constraints and criteria

Overall design goal:

- spoken language dialogue system prototype operating via the telephone and capable of replacing a human operator;

General feasibility constraints:
Scientific and technological feasibility constraints:
Designer preferences:

Realism criteria:
- *The artifact should meet real and/or known user needs;*
- *The artifact should be preferable to current technological alternatives;*
- *The artifact should be tolerably inferior to the human it replaces, i.e., it should be at least usable (to be expanded in the usability criteria);*

Usability criteria:
- *Make sure that the artifact* can do *the tasks done by the human it replaces;*

Naturalness criteria:

B. Application of constraints and criteria to the artifact within the design space:

C = Null
O = Null
S = *Understand user utterances in task domain;*
Ability to repair if understanding fails;
Large enough task-related vocabulary;
Natural grammar;
Appropriate semantics;
Natural discourse handling;
Close-to-real-time response;
Speaker-independent recognition of continuous speech;
Complete task domain information;
Clear and comprehensible communication of what the system can and cannot do;
I = Spoken telephone dialogue
T(S) =
T(U) =
U = *Need for specific types of flight information;*
Need for natural grammar and discourse handling;
Communication failure in case of delayed response;
Need to use continuous speech in task domain;
Risk of communication failure in case of lacking task domain information;
Risk of communication failure in case of lacking knowledge about what the system can and cannot do;
E = Novices

C. Hypothetical issues:
- How to accommodate experienced users?

In CO-SITUE (2), the realism and usability criteria and their interpretation have led to a number of design decisions on system performance partly based on assumptions about users. It is clear that the commitments made so far do not completely specify a number of artifact properties relating to them. For instance, the task domain has not yet been fully specified; the task-related vocabulary recognised by a usable system might still consist of as little as two (or even one !) word; the notion of close-to-real-time response is somewhat vague and so is the notion of speaker-independent recognition. Note also that some of the most self-evident user needs are absent from the list. There is no need to complicate the CO-SITUE description of the evolving design space with trivial claims such as that users need to be understood by the system. This of course introduces an extra element of judgment into the CO-SITUE analysis. However, judgments are traceable from the analysis as long as the *S*-specification contains elements which, because of assumed self-evidence, are lacking their corresponding elements in the *U*-specification. In general, the *U*-specification provides justifications for elements elsewhere in the CO-SITUE analysis.

Naturalness

Usability is a basic standard indeed for the design of interactive computer systems. Arguably, a usable system corresponding to the above commitments on usability could be developed by having a dialogue with users which is totally controlled by the system so that the user will only have to respond by "yes" or "no" throughout. However, such a dialogue would not be *natural*, i.e., it would not even remotely correspond to the way in which spoken language dialogues are normally conducted in the task domain. In other words, a natural dialogue requires that users can use language (i.e., vocabulary, grammar, semantics and discourse) to some considerable extent freely during their dialogue with the system.

It is interesting to note that in the case of 'pure' spoken or written natural language interfaces to computers, we actually do have an objective standard of naturalness. We know what natural spoken language communication is like and can easily recognise unnatural limitations imposed on it. This is not the case with respect to most other types of human-computer interfaces. With respect to the current design effort, a fully natural dialogue cannot be allowed given the task domain and the state of the relevant science and technology; so there have to be some constraints on naturalness. In other words, *naturalness can be* - and in fact has to be, in this case just as in most or all other cases of contemporary systems design - *traded for system feasibility* (general as well as scientific and technological). However, just as importantly, *the constraints on naturalness should themselves be principled and natural* in the sense that they can be easily assimilated and practiced by users, possibly based on some initial advice communicated to them by the system itself. The trade-offs have to be principled and natural because user experience has already been fixed to novice experience and because of the assumption that novice users are able to handle only principled and natural limitations on system performance.

Constraints on system naturalness may affect at least the tasks the users can perform with the system, the task domain covered by the system, the mode of user-system interaction and the types of users who can operate the system. The resulting design decisions led to significant reductions in the interactive tasks of user and system:

CO-SITUE No. (3)

A. General constraints and criteria

Overall design goal:
- Spoken language dialogue system prototype operating via the telephone and capable of replacing a human operator;

General feasibility constraints:
Scientific and technological feasibility constraints:
Designer preferences:

Realism criteria:
- The artifact should meet real and/or known user needs;
- The artifact should be preferable to current technological alternatives;

- The artifact should be tolerably inferior to the human it replaces, i.e., it should be at least usable (to be expanded in the usability criteria);

Usability criteria:
- Make sure that the artifact *can do* the tasks done by the human it replaces;

Naturalness criteria:
- *Maximize the naturalness of user-interaction with the system within the scope of feasibility;*
- *Unless a naturalness criterion cannot be met for feasibility reasons, it should be incorporated into the artifact being designed;*
- *Constraints on system naturalness resulting from trade-offs with system feasibility have to be made in a principled fashion based on knowledge of users in order to be practicable by users;*
- *Constraints on system naturalness have to be clearly communicated to users;*

B. Application of constraints and criteria to the artifact within the design space:

C = Null
O = Null
S = Understand user utterances in task domain;
Ability to repair if understanding fails;
Large enough task-related vocabulary;
Natural grammar;
Appropriate semantics;
Natural discourse handling;
Close-to-real-time response;
Speaker-independent recognition of continuous speech;
Complete task domain information;
Clear and comprehensible communication of what the system can and cannot do;
Intelligible, practicable and principled limitations on natural system performance;
I = Spoken telephone dialogue
T(S) = *Make system limitations clear to users from the outset;*
Provide information on and allow booking of flights between two specific cities;
T(U) = *Obtain information on and perform booking of flights between two specific cities;*
U = Need for specific types of flight information;
Need for natural grammar and discourse handling;
Communication failure in case of delayed response;
Need to use continuous speech in task domain;
Risk of communication failure in case of lacking task domain information;
Risk of communication failure in case of lacking knowledge about what the system can and cannot do;
Need for principled limitations on natural system performance;
E = Novices

C. Hypothetical issues:
- How to accommodate experienced users?

What has happened here is that (1) more fixed usability commitments have been introduced. (2) Distinction has been made between the tasks to be performed by the artifact and the tasks to be performed by the user during user-artifact interaction. This distinction comes very natural in the case of natural language dialogue systems but may prove useful to the analysis of other types of user-artifact interaction as well. (3) A further assumption about users has been introduced. (4) Trade-offs between naturalness and feasibility have led to initial specifications of system and user tasks. An important point illustrating the claim that the design process is a process of designing a CO-SITUE *complex* is that questions of feasibility have led to design decisions on user and system tasks. In other words, it will not do in general to assume that we have well-defined tasks to start with and then simply design a system and an appropriate interface in order to support or perform those tasks.

Feasibility constraints

Let us finally add to the map of the evolving design space some important general feasibility constraints, designer preferences and scientific and technological requirements and constraints, and their trade-offs with the design criteria mentioned already. From a usability engineering point of view, such requirements and constraints should be incorporated into the CO-SITUE notation only if and when they interact with issues of realism, usability and naturalness. There is no need, therefore, to include details on, e.g., the speech recogniser or the semantic frames used to represent the meaning of natural language expressions. The CO-SITUE frame came to look like this:

CO-SITUE No. (4)

A. General constraints and criteria

Overall design goal:
- Spoken language dialogue system prototype operating via the telephone and capable of replacing a human operator;

General feasibility constraints:
- *9 man/years available for the first version of the prototype;*
- *Limited machine power available;*

Scientific and technological feasibility constraints:
- *Limited capability of current speech and natural language processing;*
- *Open research questions;*

Designer preferences:
- *Use of the Dialogue Design Tool (DDL);*

Realism criteria:
- The artifact should meet real and/or known user needs;
- The artifact should be preferable to current technological alternatives;
- The artifact should be tolerably inferior to the human it replaces, i.e., it should be at least usable (to be expanded in the usability criteria);

Usability criteria:
- Make sure that the artifact *can do* the tasks done by the human it replaces;

Naturalness criteria:
- Maximize the naturalness of user-interaction with the system;
- Unless a naturalness criterion cannot be met for feasibility reasons, it should be incorporated into the artifact being designed;
- Constraints on system naturalness resulting from trade-offs with system feasibility have to be made in a principled fashion based on knowledge of users in order to be practicable by users;
- Constraints on system naturalness have to be clearly communicated to users;

B. Application of constraints and criteria to the artifact within the design space:

C = Null
O = Null
S = Understand user utterances in task domain;
Ability to repair if understanding fails;
Large enough task-related vocabulary;
Natural grammar;
Appropriate semantics;

Natural discourse handling;
Close-to-real-time response;
Limited speaker-independent recognition of continuous speech;
Complete task domain information;
Clear and comprehensible communication of what the system can and cannot do;
Intelligible, practicable and principled limitations on natural system performance;
500-600 words vocabulary;

I = Spoken telephone dialogue

T(U) = Obtain information on and perform booking of flights between two specific cities;
Use single sentences (or max. 10 words);

T(S) = Make system limitations clear to users from the outset;
Provide information on and allow booking of flights between two specific cities;

U = Need for specific types of flight information;
Need for natural grammar and discourse handling;
Communication failure in case of delayed response;
Need to use continuous speech in task domain;
Risk of communication failure in case of lacking task domain information;
Risk of communication failure in case of lacking knowledge about what the system can and cannot do;
Need for principled limitations on natural system performance;

E = Novices

C. Hypothetical issues:

- How to accommodate experienced users?
- *Is a vocabulary of 500-600 words sufficient to capture the sublanguage vocabulary needed in the task domain ?*

In the current context, the interesting new points in CO-SITUE (4) are the following: (1) Trade-offs between the naturalness criteria and feasibility have led to a change in system performance specification (*limited* speaker-independent recognition). (2) A change in the users' task (use*single* sentences) has been decided. And (3) a hypothesis has been made about the size of the sublanguage vocabulary needed in the task domain.

(1) Limited speaker-independent recognition was adopted for general feasibility reasons. Given the amount of time available for completing the prototype, it was agreed that *limited* speaker-independent word recognition might be acceptable in the first prototype. This would save an amount of effort in training Markov word models, which might preferably be invested in other aspects of artifact development. Although limited speaker-independence adds to the fragility of system performance, it was decided that limitations in this respect were not of a principled nature, the degree of speaker-independence being largely a function of the amount of work spent on training the system's word models.

(2) It was decided that, during each turn in the user-system dialogue in which users address the system, they should use brief utterances, and preferably single sentences in doing so. This will increase the likelihood of utterance recognition and understanding by the system. At the same time, this constraint on the user-system dialogue seems to be *principled* and thus meets the naturalness criterion. The assumption is that the brevity constraint can be easily understood, assimilated and respected by the intended users during their dialogue with the system.

(3) The close-to-real-time and naturalness design constraints jointly led to the design decision that a vocabulary of 500-600 recognised words should be sufficient for the selected user tasks. Fewer words recognised would constitute an unnaturally restricted sublanguage for the task domain whereas a larger vocabulary would endanger the close-to-real-time design constraint. This design decision clearly rests on a *hypothesis* about the nature of the task domain

sublanguage and therefore has been made explicit in CO-SITUE as a crucial hypothetical issue. The hypothesis was tested and confirmed during the knowledge acquisition phase. This testing conformed to the idea that artifacts are CO-SITUE complexes, since the two parameters of user task and system task were manipulated until the resulting user sublanguage vocabulary was smaller than required.

6. Concluding Discussion

CO-SITUE (4) represents both the artifact and the design space surrounding it after completion of initial artifact specification. The representation is the result of making explicit a number of generic constraints and criteria on the artifact to be designed and applying, through a process of interpretation, discovery, justification, trade-off and decision-making, those generic constraints and criteria to all aspects of the evolving conception of the artifact. Each general constraint or criterion, and each design commitment resulting from applying these to the artifact within the design space represent additional constraints on the subsequent artifact specification process.

There is no reason to expect that a series of CO-SITUE frames will be filled in any systematic fashion. Design reasoning, *qua* dealing with open-ended and ill-structured problems, may use a limited commitment-mode control strategy (Goel and Pirolli 1992) under which different aspects of the design problem are being addressed concurrently and interchangeably and no one problem needs to be completely solved before other problems are being addressed. Instead, designers are taking advantage of, often linked, multiple problem-solving contexts in their generation, evaluation, re-consideration and modification of design solutions. As a consequence, successive CO-SITUE frames are generally expected to be incrementally filled in a seemingly random, across-the-board fashion.

In the design process described, subsequent artifact specification was done through two parallel approaches. One was a more detailed specification of system architecture, the other was empirical knowledge acquisition through *T(S)* simulation using the Wizard of Oz methodology. Interestingly, the *S, I, T(S)* and *T(U)* aspects of the artifact resulting from initial artifact specification allowed a direct derivation of an optimal knowledge acquisition model. Knowledge acquisition for the artifact should be done through a simulation of an artifact which reflects the design commitments represented in CO-SITUE (4). If empirical knowledge acquisition is carried out according to this model, there is an optimal likelihood that the results obtained will be valid for the implemented artifact (Bernsen 1993).

A noteworthy aspect of the CO-SITUE notation is its compactness. Despite describing a relatively substantial artifact development effort with a focus on usability issues, most of the important design decisions (excluding only the overall system architecture) and the generic constraints and criteria on which they are based can be represented rather compactly. This no doubt is a virtue of the notation although it should be emphasized that it has been realised through excluding a large amount of design process information. What matters, however, is that what has been included is viewed by consensus in the designer team as constituting the most important criteria, constraints and design decisions adopted so far which relate to usability engineering issues. It can safely be claimed that use of the CO-SITUE notation has been helpful in keeping track of designer consensus during our development of the first prototype dialogue system and that the existence of the compact design space representation in CO-SITUE (1)-(4) will make a difference to our work on the following, second prototype.

The virtues of CO-SITUE in its current form can be summarised as follows:

- CO-SITUE offers a framework for making explicit the most general aspects, criteria and constraints characterising the design space surrounding computer artifacts and hence enforces a consideration, during artifact design, of each of those from the point of view of usability engineering;

- maintaining a numbered series of CO-SITUE frames during the design process provides a means of explicitly and succinctly capturing the most important design decisions made as artifact design evolves as viewed by designer consensus. When the designer team is large and/or organisationally or geographically distributed, CO-SITUE can be useful in keeping track of designer consensus and the reasoning behind it, including the ultimate justification of the specific design commitments made;

- CO-SITUE in effect provides a series of 'snapshots' of the evolving design process including design space and artifact specification.

Some drawbacks of CO-SITUE in its present form are:

- the CO-SITUE notation does not incorporate the design reasoning linking up CO-SITUE (n-1) with CO-SITUE (n). Instead, design reasoning has been represented above in the form of brief discussions linking each consecutive pair of CO-SITUE frames. Ways should be sought of providing more structure to the representation of design reasoning (see below);

- CO-SITUE in itself offers little in terms of scientific theory. This is unfortunately true of all current and realistic approaches to artifact design support. What CO-SITUE offers so far is an analysis of artifacts and the design space surrounding them at the highest level of generality, and a notation for analytic and dynamic capture of design decisions combined with informal representations of the reasoning behind them.

Some open issues are:

- how does CO-SITUE work when its comes to representing the details of tasks and interfaces of the artifacts being designed?

- how much extra overhead on the design process does use of the CO-SITUE notation imply?

More open issues will be pointed out below. It is the aim in the context of AMODEUS II to address these drawbacks and open issues of CO-SITUE as it stands. The following two sub-sections describe the relationship between CO-SITUE and the Design Rationale approach to design space analysis using Questions, Options and Criteria (DR/QOC for short) and the issue of a 'level 2' articulation of CO-SITUE, respectively.

6.1 CO-SITUE and Design Rationale

DR/QOC represents an approach under development in AMODEUS II to providing additional structure to design reasoning. The idea is that Design Rationale can be captured through design space analysis in terms of Questions, Options and Criteria (QOCs), which helps provide explicit analytic structure to the design space surrounding an artifact (MacLean et al. 1990, MacLean et al. 1991). CO-SITUE and DR/QOC are complementary approaches and the current aim is to combine them into a framework for usability engineering which can be turned into a computer-based tool. The main similarities, differences and complementarities between CO-SITUE and DR/QOC in their present forms are the following:

Similarities: Both DR/QOC and CO-SITUE represent forms of design space analysis which, generally speaking, is a natural style of reasoning to designers. Both approaches are therefore potentially useful to the structuring of design discussions. Both are analytical accounts of design processes rather than historical documentations of them. Both representations provide useful documentation for supporting effective communication within design projects, with future designer teams responsible for extending or maintaining the artifact, with experts in science-based approaches to artifact evaluation, and possibly also with end-users. The representational formats involved are simple enough to be understood by a variety of people. On the other hand, both types of design space analysis face the challenge of having to offer more in cost-benefit terms than familiar design representations such as generating scenarios, sketching possible solutions, listing required attributes, using rapid prototyping and producing prototypes. Finally, the fact that design space analysis is a natural style of reasoning to designers of course does not imply that their reasoning maps directly into the CO-SITUE and DR/QOC notations and their underlying conceptual structures. More research is needed on this issue which is crucial to the eventual adoption by designers of CO-SITUE and/or DR/QOC.

Differences and complementarities: The main strengths of DR/QOC as compared with CO-SITUE are that DR/QOC-style design space analysis (1) explicitly aims at supporting, during the design process, the development of an organised space of alternative and innovative designs or possible solutions, as well as (2) facilitating the analysis of weaknesses of proposed solutions. Such DR/QOC representations might be directly incorporated into the CO-SITUE frame notation. However, it seems preferable to develop DR/QOC representations in the form of separate links between the individual CO-SITUE frames in a series representing a particular design process. The design process would then be analytically represented for usability engineering purposes as the series: CO-SITUE (1) -> DR/QOC (1) -> CO-SITUE (2) -> DR/QOC (2) -> . . . -> CO-SITUE (n) -> DR/QOC (n). The representation DR/QOC (n) would provide the design reasoning which justifies the design decisions made between CO-SITUE (n-1) and CO-SITUE (n). Furthermore, DR/QOC (n) would not necessarily have to represent a comprehensive analysis of every design decision made between CO-SITUE (n-1) and CO-SITUE (n) but might focus on the most important decisions for which it would offer a detailed analysis, whereas other decisions might be discussed in the informal style used in the present paper.

CO-SITUE, on the other hand, (1) attempts to make explicit the overall semantic structure of design spaces, i.e. to systematically conceptualise the aspects of the artifact being designed which need to be taken into account from a usability engineering perspective. DR/QOC does not represent overall design space structure. Linked to that, and in contrast to most DR/QOC work done so far, (2) CO-SITUE represents a top-down approach to design processes in that it aims at capturing those constraints and criteria, and their gradual application within the design space, which are most important from a usability engineering point of view. Outside of this unfolding context, it can be claimed, DR/QOC representations make little sense as an aid to design. The unfolding design context involving general criteria and constraints and the design commitments based on these are potentially relevant to any subsequent design decision to be made during the design process. This context provides the design space structure which often carries the specific style of the artifact and is used in justifying the more specific design arguments which are applied to the solution of specific design problems. It follows that the general criteria, constraints and design commitments dealt with by CO-SITUE necessarily have to be represented also in a DR/QOC-style approach.

Linked to these two points, again, is the fact that (3) CO-SITUE explicitly captures design constraints which do not directly express usability issues but which, nevertheless, affect usability issues by enforcing trade-offs with the criteria of usability and naturalness. It may be claimed that capturing such constraints is necessary to the analytic capturing of design reasoning even though they sometimes trade off *against* the naturalness of a particular design. Indeed, the apparent opposition between purely historical design process accounts and purely justificational accounts is fictitious. Any historical account will contain justifications, and any justificational account, however analytically expressed, will contain history. Since design problems are ill-structured and do not have unique and correct solutions, what matters is to provide as much analytic structure to them as possible in order to support the justificational strength of their solutions. Finally (4), it is an important question for further work whether a CO-SITUE representation of the design process can facilitate identification of those design issues for which a detailed DR/QOC analysis could provide the most benefit. This is an as yet unsolved problem within the DR/QOC approach, the problem being that a 'complete' DR/QOC design space analysis of medium- and large-scale design processes would probably, for most types of design processes, require too much effort to be realistic in cost-benefit terms.

The similarities, differences and complementarities between the DR/QOC and CO-SITUE approaches noted above would seem to justify the attempt to combine them into an explicit and analytic representational framework for design processes.

6.2 Level 2 Articulation of CO-SITUE

If CO-SITUE as presented above is considered a general level articulation, or a *Level 1* articulation, of the structure of design spaces, the question arises whether a *Level 2* articulation of design space structure is possible. If approximately valid, CO-SITUE captures the most general aspects of design spaces and artifacts. These aspects are thus invariants involved in any computer artifact design process. As remarked above, such design processes have to reckon with many more invariants characterising artifacts and design spaces. The problem is to articulate these invariants in a systematic fashion which both maintains a sufficiently high level of generality and is directly relevant to practical design reasoning. Such invariants are potentially of at least two kinds:

(1) *A taxonomy of CO-SITUE complexes* describing design-relevant combinations of aspects of organisations, collaborations, systems, interfaces, interactive tasks and their domains, and users. As argued above (Sect. 2), independent individual taxonomies of these aspects are probably neither relevant to artifact design nor possible in the first place. It is an empirical question to what extent a taxonomy of CO-SITUE complexes is feasible. Their articulation would probably have to proceed via:

(2) *The identification of general invariant structures* of organisations, collaborations, systems, interfaces, interactive tasks and their domains, users and features of the novice-to-expert development process. This would constitute a Level 2 articulation of CO-SITUE and forms the subject of ongoing work. In the particular context of AMODEUS II, the Cognitive Subsystems framework developed by Barnard is an obvious candidate for articulation of the *U* aspect of CO-SITUE (see Phil Barnard's paper in this volume).

It is an open question whether much further useful articulation of the general criteria and constraints of realism, usability, general feasibility, and scientific and technological feasibility is possible or worthwhile. However, it seems likely that the naturalness criterion might profitably become further articulated. There clearly is work to be done on how to define naturalness with respect to multi-modal interfaces, for instance.

We cannot expect systematic HCI theory to ever overcome the gap between general constraints and criteria, and general CO-SITUE-like aspects of design spaces and artifacts, on the one hand, and the immense complexity of detail met with when interpreting constraints, criteria and aspects with respect to particular artifacts or design goals, on the other. However, some further systematic, top-down articulation of constraints, criteria and CO-SITUE aspects might prove useful in practical design work as one possible way of making the design process more rational, or more systematic, or more transparent, or more amenable to receiving input from an appropriately developed basic science. The CO-SITUE framework is a candidate approach along these lines but it is clearly too early to make strong claims about its capabilities to deliver.

References

Barnard, P.: Bridging between basic theories and the artifacts of human-computer interaction. In J.M. Carroll (Ed.): *Designing Interaction. Psychology at the Human-Computer Interface.* Cambridge: Cambridge University Press 1991.

Bernsen, N.O.: Design of a Spoken Language Dialogue System. A Study of the Initial Specification Phase. *Esprit Basic Research Action AMODEUS Working Paper* RP3-WP3,1993.

Brooks, R.: Comparative Task Analysis: An Alternative Direction for Human-Computer Interaction Science. In J.M. Carroll (Ed.): *Designing Interaction. Psychology at the Human-Computer Interface.* Cambridge: Cambridge University Press 1991.

Card, S.K., Moran, T.P., and Newell, A.: *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates 1983.

Carroll, J.M. (Ed.): *Designing Interaction. Psychology at the Human-Computer Interface.* Cambridge: Cambridge University Press 1991.

Dybkjaer, L., Bernsen, N.O. and Dybkjaer, H.: Knowledge acquisition for a constrained speech system using WoZ. In *Proceedings of the EACL* 1993 (in press).

Dybkjaer, H. and Dybkjaer, L.: *Wizard of Oz Experiments in the Development of the Dialogue Model for P1*. Report 3, Spoken Language Dialogue Systems. STC Aalborg University, CLT Copenhagen University and CCI Roskilde University, February 1993.

Goel, V. and Pirolli, P.: The structure of design problem spaces. *Cognitive Science* 16, 395-429, 1992.

Kieras, D.E., & Polson, P.G. (1985). An approach to formal analysis of user complexity. *International Journal of Man-Machine Studies,* 22, 365-394.

Larsen, L.B. (Ed.), Bernsen, N.O., Broendsted, T., Dybkjaer, H., Dybkjaer, L., Music, B., Povlsen, C., and Ravnholt, O.: *Spoken Language Dialogue Systems. A Survey of the State-of-the-Art.* Report 1, Spoken Language Dialogue Systems. STC Aalborg University, CLT Copenhagen University and CCI Roskilde University, June 1992.

Larsen, L. B. (Ed.), Broendsted, T., Dybkjaer, H., Dybkjaer, L., and Music, B.: *Overall Specification and Architecture of P1*. Report 2, Spoken Language Dialogue Systems, STC Aalborg University, CCI Roskilde University, CST University of Copenhagen, February 1993.

MacLean, A., Bellotti, V. and Young, R.: What Rationale is There in Design? In Diaper, D., Gilmore, D., Cockton, G. and Shackel, B. (Eds.): *Proceedings of INTERACT '90: Third IFIP Conference on Human-Computer Interaction*. Amsterdam: Elsevier North-Holland 207-212.

MacLean, A., Young, R., Bellotti, V. and Moran, T.P.: Questions, Options, and Criteria: Elements of Design Space Analysis. *Human-Computer Interaction* Vol. 6, 1991, pp. 201-50.

Norman, D.A.: Cognitive artifacts. In J.M. Carroll (Ed.): *Designing Interaction. Psychology at the Human-Computer Interface.* Cambridge: Cambridge University Press 1991.

Payne, S., & Green, T. (1986). Task Action Grammars: A model of the mental representation of task languages. *Human-Computer Interaction*, 2, 93-133.

Acknowledgements: Many thanks are due to Phil Barnard, Hans Dybkjaer, Laila Dybkjaer, Allan MacLean and Tove Klausen for helpful comments on earlier drafts. The work was done as part of the Esprit Basic Research project 7040 AMODEUS II whose support is gratefully acknowledged.