# THE STRUCTURE OF THE DESIGN SPACE

## CO-SITUE Illustrated by a Study in Early Artifact Design

*Niels Ole Bernsen, CCI, Roskilde University and Risø National Laboratory*

*Summary:* The paper presents a framework, called CO-SITUE, for structuring top-down the design space surrounding the design of computer artifacts. CO-SITUE involves the following aspects of design spaces and artifacts: collaboration, organisation, system, interface, task and task domain, user, and user experience. The framework complements the Design Rationale approach to the study of designer reasoning by proposing an overall structure for the problem space in which designer reasoning takes place. To illustrate the use of CO-SITUE in an actual design task, the framework is being applied as a notational frame for analysing the initial artifact specification phase of a spoken natural language dialogue system.

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

## 1. Introduction

The applied science of HCI is currently moving towards analysing entire design processes or phases of such processes. The main reason for this timely endeavour seems to be that, whatever may be the contributions to be had from basic science to the design of advanced computer artifacts, such contributions are likely to be useful only if based on a deep understanding of actual design processes (cf. the papers in J.M. Carroll et al. 1991). To prove itself as an applied science, HCI needs to produce useful *Application Representations* linking the HCI science base with actual design work (Barnard 1991). One approach sharing this general point of view is the *Design Rationale* (DR) approach (e.g., MacLean et al. 1990, MacLean et al. 1991). In DR, design reasoning is described in terms of design *Questions* (Q) for which design *Options* (O) are generated, and explicit design *Criteria* (C) are then applied to the various options in order to allow and record explicit design decisions among the available options. Put crudely, one asks "How should this particular aspect of the artifact be designed ?" In response, a number of solutions to the question are generated. A set of criteria are then developed for use in rationally choosing among those solutions. Finally, the criteria are applied and the preferred solution identified by comparing the trade-offs implied by different solutions. The explicit recording of a design process in such QOC terms would describe the design process as a particular path through the design space surrounding the artifact being designed.

In DR, however, the design space itself is not made visible except through the actual series of QOC's that a particular design process actually passes through. Of course, this series *might* reveal the entire design space relevant to the artifact being designed. The problem is that, without prior knowledge of the structure of the design space we have no systematic way of knowing whether this is the case or not. Similarly, and for the same reason, DR itself provides no systematic way of *generating* the series of QOC's that are relevant to a particular design process. It is certainly not clear at this point to what extent and to which level of detail it might be possible to specify the structure of the design space for interactive computer artifacts in general. However, even the specification of the coarse outlines of this design space might constitute a useful complement to DR and related approaches to the study of design reasoning through providing a principled context within which to locate and possibly generate both individual QOC's and the series of QOC's making up particular design processes.

In what follows, the *CO-SITUE framework* for the applied science of HCI is presented as a characterization of the coarse structure of the design space surrounding the design of interactive computer artifacts (Sect. 2). The framework is then applied to the early artifact specification phase of a collaborative design project carried out by the Speech Technology Centre, Aalborg University, the Centre for Language Technology, Copenhagen University, and CCI. CO-SITUE is presented as a draft notation for capturing design decisions made during initial artifact specification at a coarse level of detail sufficient for capturing major design decisions (Sect. 3). Finally, directions of further work and ways of combining CO-SITUE and DR are discussed (Sect. 4).

## 2. CO-SITUE: The General Idea

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:

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

The basic claim behind CO-SITUE 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 what I shall call a *CO-SITUE complex*. What is being designed is an artifact and an artifact is a CO-SITUE complex.

During the design process, designers have to consider and actually do consider to some extent 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 derived from a consideration of 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, 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. The claim is that designers actually work within the conception of an artifact designated by CO-SITUE, that is, *CO-SITUE describes the overall shape of the design space around an artifact during design at a coarse level of detail*. If that is true, then CO-SITUE does hook into actual design reasoning and hence might provide common ground among designers and HCI scientists alike on which to discuss principles of artifact design.

There are two main points involved in the claim that what is being designed is a (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 designing a system and then looking at user requirements, or first designing a system and then looking at user tasks, or first looking at the user tasks and then designing the system, or first designing a system and then looking at its interface, etc. In design reasoning, multiple constraints are simultaneously 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.

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.* This claim may appear counter-intuitive but is, in fact, straightforward. It has already been noted by several that computer artifacts change the users' tasks (e.g., Norman 1991). It is important to realise, however, that this point about change is more general than that. As seen from within the dynamical design process, all aspects of CO-SITUE are potentially subject to change. The reasons are (1) that computer artifacts *change work conditions:* they change not only tasks but often also collaborative and organisational schemes of work; (2) that artifacts *"change" users* in the sense that the types of target end-users and the requirements on their experience for dealing with the artifact are themselves 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 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 way to change the organisational and collaborative work schemes of users, the way to change their tasks, and the way to select the types of end-users and the experience they need for operating the artifact. The artifact which evolves during the design process is a result of design decisions concerned with all aspects of CO-SITUE. *The design space is a field of possibilities or potential changes, and the scope of those changes include CO-SITUE as a whole.* Thus, for instance, in experimental systems design one may start out 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.

The above claims should not be misconstrued as saying that everything 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 limited working memory to the fact that doctors don't want to transfer responsibility to a medical expert system. 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. 4 below).

## 3. An Example Study of an Early Design Process

The design process to be described here in terms of CO-SITUE is the initial specification phase of a spoken language dialogue system prototype (see Bernsen 1992 for a full account, cf. Larsen et al. 1992).

### 3.1 Design phases

In many cases of computer artifact design it is not possible to clearly separate the initial specification phase from subsequent phases of artifact specification. This is possible to some noteworthy extent, however, when dealing with today's spoken language dialogue systems and the reason is that, due to scientific and technological limitations, one can only specify the artifact to a certain point of detail prior to the knowledge acquisition phase. Many types of result from the knowledge acquisition phase are critically important to further system specification refinement.

In broad terms, the design process was found to have six, partly parallel and iterative, phases:

- initial artifact specification;
- knowledge acquisition;
- detailed architecture specification;
- knowledge representation;
- system implementation;
- system testing.

The term "initial artifact specification" indicates that this is where artifact development begins. The specification refinement process will continue throughout the design process and will interact strongly with the other phases of artifact development. Many design decisions made during initial specification may have to be revised later. The initial specification phase is operationally defined as the phase of artifact specification *up to but not including* the systematic knowledge acquisition phase and the detailed definition of system architecture involving specification of the internal operation of all system modules and the interfaces between the modules. The detailed system architecture specification task was undertaken concurrently with the knowledge acquisition phase (Larsen et al. 1993). It should be emphasized that the distinction made below between design space delimitation and first design decisions within the delimited design space is analytical, not temporal. In particular, *it is very often only when having started to work within an initially vaguely delimited design space that some of the conditions contributing to its firmer delimitation are discovered.* Thus the CO-SITUE analysis provided below represents an analytical (or *post hoc*) approach to initial artifact specification rather than a temporal one. However, CO-SITUE is perfectly compatible with the temporal approach which would normally be adopted during actual design (see Sect. 4 below).

### 3.2 Overall design goal and design constraints

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 time, manpower, available machine power and so on. I will only mention a couple of the most important among these *general feasibility constraints* below. 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 design decisions are to be made. The distinction between realism, usability and naturalness can be viewed as a structured alternative to the nebulous standard notion of "system usability" or "-habitability". Finally, it may be 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 become *interpreted* with respect to the overall design goal. The designed artifact is the outcome of this process of interpretation.

### 3.3 An empty CO-SITUE frame

Here follows an empty version of the CO-SITUE notational frame:

---

**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. DR/QOC arguments:**
**E. 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.

Note that 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.
Note further that point D on DR/QOC arguments will only be discussed in Sect. 4 below. Further conventions will be added as we proceed.

---

### 3.4 CO-SITUE exemplified

Let us now 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.* During initial artifact specification we are dealing with general constraints on the design task. Design reasoning at this stage consists in discovering the constraints, develop their implications and making design decisions on the basis of these. 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 to make appropriate use of the artifact. If such changes are not 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 whose initial design is being 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 rather modest analysis of this kind has been conducted as a field study in a professional travel agency (see Larsen et al. 1993). A successor to the artifact being designed might eventually end up being used by the travel agents of this agency and it is very likely that its introduction might cause changes to their organisational and collaborative schemes of work. If this turns out to be the case, the feasibility of those changes should be investigated much more deeply in order to make sure that the artifact will eventually be realistic and usable. Until this field study has been concluded, we assume that we are dealing with an "individualistic" artifact (a SITUE complex).

*Experience and Interface aspects*

The *E* aspect of CO-SITUE is often comparatively simple. In the case of our spoken language dialogue system prototype, the design decision was to initially aim at novice users. Experienced users would have to accept an artifact designed for novice users even though this might force them to have 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. A more detailed artifact interface can only be developed during the knowledge acquisition design phase. 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 = <u>*spoken telephone dialogue*</u>
T(S) =
T(U) =
U =
E = *novices*

**C. Hypothetical issues:**
**D. DR/QOC arguments:** N/A

**E. Conventions:** See CO-SITUE (0) above

Italics indicate new elements in the design specification (No. n) as compared with the immediately preceeding specification (No. n-1).
Underlining indicates that this aspect of the artifact will remain fixed throughout the design process.
"Null" means that an artifact does not have 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 the design process unless it gets changed to something else because of novel considerations. 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 the question came up about how to handle experienced users. In response to this question, 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. *I,* thus specified, is therefore underlined in the proposed notation. The issue about which design decisions are *really* fixed and which ones are more or less fixed or temporary is not important. The important thing is that the CO-SITUE notational frame makes all important design constraints and decisions explicit. If one so prefers, the difference between underlined and non-underlined design decisions can be regarded merely as a point of emphasis.

If CO-SITUE circumscribes the design space, then the gradual specification as a result of design decisions demonstrates the gradual progress of artifact specification.

## 3.5 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. In the current design task, we immediately ran into 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.

*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 spoken language dialogue systems today are time table enquiries for flights and trains. It was decided in the current design project 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, the realism of the artifact to be designed comes 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 tolerably inferior in performance to humans performing the same tasks. It is possible that these criteria of realism 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 just seen how the realism criteria were *interpreted with respect to the task domain of the artifact to be designed.* The realism criteria as such do not state anything specific about the task domain.

*Usability*

Usability criteria serve to ensure that the system *can do* the tasks done by the human it replaces. If the usability criteria cannot be met then the system will be practically useless. It turns out that many usability criteria can be self-evidently derived 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 give a flavour of the actual contents of the design task, the interpretation of usability with respect to the design goal will be 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 various other assumptions:

- 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 a difficult time getting 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 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 not familiar with the literature on earlier spoken language dialogue systems or with relevant principles of human language processing;

- the system's *semantics* should be appropriate to the chosen vocabulary and grammar. Anyone who could design the linguistic parts of the system would know that;

- 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 connection with the grammer 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 considerations even without invoking cognitive theories of attention span;

- the system should preferably do *speaker-independent* recognition of speech as usability in the chosen task domain is seriously affected by the training process which is otherwise needed for the system to adapt to a new user. In other task domains (e.g., voice process control in the cockpit), speaker-independence may be 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 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 will be limited in performance, users obviously have to know what it 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 seem to exist any procedure for making sure that it is sufficient. Moreover, it is clearly possible for different designers to derive somewhat different sub-sets of the necessary conditions for system usability. Let us take a closer look at how these conditions are derived.

The above usability criteria are *minimal* in the sense that they state conditions necessary for the system to be at all able to replace a human operator in the task domain. They are derived from the overall goal specification of the artifact to be designed 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 usability (cf. the point made about collaborative and organisational aspects above) but it certainly seems to be necessary. In other words, *to determine usability criteria for an artifact to be designed, one has to identify conditions such that, if they are not met by the designed artifact then the artifact cannot do the tasks to be assigned to it.* As long as we do not have any systematic way of identifying such conditions, it will be up to the collective know-how present in the designer team to identify as many important usability conditions as possible. There is no way of guaranteeing that the identified set of usability criteria is complete, and different designers or designer teams may well identify different sub-sets of the usability criteria for a given artifact. This, indeed, seems to happen all the time. Moreover, as indicated, *since important assumptions about the cognitive capabilities and limitations of users may be involved, 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 on 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.

One possibility of ameliorating the situation just described of unprincipled derivation of usability criteria might be to identify some of the basic structural properties of the design space within which system design takes place (see Sect. 4 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;*
*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:**
**D. DR/QOC arguments:** N/A

**E. Conventions:** See above

---

In CO-SITUE (2), the realism and usability criteria and their interpretation have lead to a number of specifications of system performance partly based on assumptions about users. Underlined artifact specifications and user assumptions are mandatory and act as (DR/QOC) criteria in subsequent artifact specifications. However, although mandatory, these criteria 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 in a usable system could still be as little as two (or even one !) word; the notion of close-to-real-time response is 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 that users need to be understood by the system. This of course introduces an extra element of judgment into the CO-SITUE analysis. *However, the judgment is* traceable *from the analysis as long as the S-specification contains an element which, because of assumed self-evidence, is lacking its corresponding element in the U-specification.* In general, the *U*-specification provides justifications for elements elsewhere in the CO-SITUE analysis.

*Naturalness*

However, usability is a basic standard indeed for the design of interactive computer systems. Arguably, a usable system corresponding to the above usability requirements 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. Note that, in the case of “pure” spoken or written natural language interfaces to computers, we actually do have an objective standard of naturalness. This is certainly not the case with respect to most other types of human-computer interfaces. On the other hand, a fully natural dialogue cannot be allowed at this stage 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) *traded for system feasibility* (general as well as scientific and technological). But, just as importantly, *the constraints on naturalness should 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-off has to be principled because user experience has already been fixed to novice experience and because of the assumption that novice users are able to handle only principled limitations on natural 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 lead 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;
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:**
**D. DR/QOC arguments:** N/A

**E. Conventions:** See above

---

What has happened here is that (1) more fixed usability criteria 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 might 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 lead 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 *the question of feasibility has lead to design decisions on user and system tasks.* Generally speaking, in designing an artifact one is also designing the tasks to be performed using it. In other words, it won't 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 and designer preferences, and the scientific and technological requirements and constraints and their trade-offs with the design criteria mentioned already. The CO-SITUE frame came to look like this:

---

**CO-SITUE No.** (4)

**A. General constraints and criteria**

Overall design goal:
- <u>spoken language dialogue system prototype operating via the telephone and capable of replacing a human operator;</u>

General feasibility constraints:
- *7 1/2 m/y's 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;
*Hidden Markov Models;*
*500-600 words vocabulary;*
*components of general system architecture;*
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;
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:**
- *is a vocabulary of 500-600 words sufficient to capture the sublanguage vocabulary needed in the task domain ?*

**D. DR/QOC arguments:** N/A

**E. Conventions:** See above

The design decision to use the Hidden Markov Models approach to speech recognition does not need further justification and I will not go into the components of the general system architecture which were identified during initial artifact specification. 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 lead to a change in system specification (*limited* speaker-independent recognition); (2) a change in the users' task (use *single* sentences) has been decided on; and (3) a hypothesis about the size of the sublanguage vocabulary needed in the task domain has been identified:

(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 felt that limitations in this respect were not of a principled nature, the degree of speaker-independence being a straightforward function of the amount of work spent on training the system's word models.

(2) It was decided that, during each turn of 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 users during their dialogue with the system.

(3) The close-to-real-time- and naturalness design constraints jointly lead 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 is being tested during the knowledge acquisition phase.

### 3.6. Conclusion

CO-SITUE (4) represents both the artifact and the design space surrounding it after completion of the initial artifact specification phase. CO-SITUE (4) is the result of making explicit a number of generic constraints on the artifact to be designed and applying, through a process of interpretation, discovery, trade-off and decision-making, those generic constraints to all aspects of the evolving conception of the artifact. Each general constraint or criterion, and each result of an application of these to the artifact within the design space represents a constraint on the subsequent artifact specification process. Furthermore, the *S, I, T(S)* and *T(U)* aspects of the artifact resulting from initial artifact specification *allow the derivation of an optimal knowledge acquisition model.* 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 (for more detail, see Bernsen 1992).

## 4. Discussion

Perhaps the aspect of the CO-SITUE notational frame which has struck us the most is its compactness. Despite describing a relatively substantial artifact development effort, 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, but it should be emphasized that it has been realised through *excluding* a large amount of design process information. However, the basic justification for the exclusions which have been made 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.

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 characterising the design space surrounding computer artifacts and hence enforces a consideration of each of those aspects during artifact design;

- in this way, CO-SITUE might support the generation of the particular DR/QOC's which drive the design process towards artifact specification;

- maintaining a numbered series of CO-SITUE frames during the design process provides a means of explicitly capturing the most important design decisions made as artifact design evolves as viewed by consensus among the designers. When the design 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 criteria applied;

- in its present form, CO-SITUE in effect provides a series of "snapshots" of the evolving design process including the design space and artifact specification.

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

- CO-SITUE, at least as used above, does not incorporate the design reasoning linking up CO-SITUE (n-1) with CO-SITUE (n). However, as already suggested in the notation used, the DR/QOC-like design reasoning linking up CO-SITUE (n-1) with CO-SITUE (n) can easily be incorporated into CO-SITUE itself;

- CO-SITUE does not incorporate entries for important but so far unresolved design issues or entries for open (in contrast to hypothetical) empirical design issues to be investigated later on in the design process. In brief, each CO-SITUE No. (n) is mainly retrospective rather than being also prospective. However, such lacks may be corrected by adding new entries to the notational frame;

- like its cousin, the DR/QOC-style approach to design reasoning, CO-SITUE offers little in terms of scientific theory. This is unfortunately true of all current approaches to artifact design support. What CO-SITUE offers so far is a general idea of what constitutes artifact design (Sect. 2 above), an analysis of artifacts and the design space surrounding them at the highest level of generality, and a notational frame for dynamically capturing design decisions.

Some issues for further work are the following:

It is of course important to develop and test the notational frame through analysing more ongoing design processes. As suggested above, there are many open issues which may affect the generalisability of the approach. For instance, if it turns out that, for the artifact described, the *C* and *O* aspects are not null, then the *U* aspect in the notational frame may have to be duplicated to take two kinds of users into account: the users that call the system to obtain information and make reservations, and the users in the organisation which has installed the artifact. This revision might also affect the interpretation and even the general criterion of usability as well as leading to a duplication of *T(U)* in the notational frame. One possible outcome of the generalisation process might be a computer tool for capturing design reasoning combining CO-SITUE and DR/QOC.

A second, related line of work is the further development of CO-SITUE itself. If 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, computer artifact 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 of at least two kinds: (1)*taxonomies* of organisations, collaborations, systems, interfaces, interactive tasks and their domains, and users; and (2) *general invariant structures* of organisations, collaborations, systems, interfaces, interactive tasks and their domains, and users. It seems doubtful if further useful articulation of the criteria and constraints of realism, usability, general feasibility, and scientific and technological feasibility is itself feasible. However, (3) the *naturalness criterion* might conceivably become further articulated in a useful manner. Nobody expects 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 a 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 basic science.

## 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.*Working Papers in Cognitive Science* WPCS-92-5. Centre of Cognitive Science, Roskilde University 1992. Contribution to Esprit Basic Research Action AMODEUS.

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

Larsen, L.B., Bernsen, N.O., Brøndsted, T., Dybkjær, H., Dybkjær, L., Music, B., Povlsen, C., and Ravnholt, O.: *Spoken Language Dialogue Systems. A Survey of the State-of-the-Art*

Report 1.1 from the project: Spoken Language Dialogue Systems. STC, Aalborg University, CLT, Copenhagen University and CCI, Risø National Laboratory and Roskilde University. August 1992.

Larsen, L.B., Dybkjær, H. and Dybkjær, L.: System Architecture. Report 1.2 from the project: Spoken Language Dialogue Systems. STC, Aalborg University, CLT, Copenhagen University and CCI, Risø National Laboratory and Roskilde University. January 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.

*Acknowledgements.* I am grateful to Hans Dybkjaer, Laila Dybkjaer and Tove Klausen for helpful comments on this draft. The work was done in part on a grant from the Danish Research Council for the Technical Sciences on "Spoken Language Dialogue Systems" and in part under AMODEUS II, Esprit Basic Research Action 7040.