Using Task Context Variables for Selecting the Best Timing for Interrupting Users

Sonja Gievska[(1)] *& John Sibert*[(2)]

[(1)] The George Washington University
Washington D.C., 20052, USA
sonjag@gwu.edu

[(2)] The George Washington University
Washington D.C., 20052, USA
sibert@gwu.edu

Abstract

This paper presents a framework that helps in selecting the most appropriate timing for interruption as a way to mediate human interruptions by the computer. The conceptual framework is based on the new Interruption Taxonomy and uses Bayesian Belief Networks as a decision-support aid. A proof-of-concept model was constructed for the experimental setting used in the exploratory study that was also part of this research. The steps in constructing the model that was built into the first version of the interruption mediator will be presented to show, in detail, how one might use the proposed framework for mediating interruptions.

1. Introduction

Recent trends in software development directed toward intelligence, distribution, and mobility have brought sophisticated software artifacts that often come with some unwanted side effects; frequent interruptions, for instance. The results of the experiments on the disruptive consequences of interruptions have shown that people make more mistakes, have difficulties remembering, hesitate and delay in making decisions, and in general are less effective when exposed to interruptions. As interruptions naturally occur during any communication including human-computer interaction (HCI), even a small reduction in their harmful effects can have significant benefits.

Steady progress has been made toward identifying and understanding what factors make some interruptions more disruptive than others. Task complexity [1], [2], coordination method used to handle interruptions [10], interruption point at which interruptions arrive [2], similarity between the ongoing and the interruption task [4], interruption modality [7], etc. have proven to affect task performance and user's emotional state in context of interruption.

The theory and traditional user-interface design guidelines do not address the interruption problems entirely. Recommendation and empirical results still do not generalize to wide majority of application domains and systems. There is a lack of a general framework to guide interface designers in developing more tacit and graceful interaction that can leverage the strengths and support the weaknesses of humans in presence of interruptions.

This paper describes the design specifics of implementing the first research version of the interruption mediator based on the proposed framework. A new taxonomy that identifies and organizes the relevant context factors for selecting the most appropriate timing for interrupting people is developed. A key feature of the model is the employment of a suitable decision-theoretic support that is needed for making decision when to interrupt the user. The effectiveness of the model was experimentally measured in terms of improved performance, and decrease of the disruptive effects of interruption on user's socio-emotional state, such as: feelings of stress, distraction, annoyance, frustration, etc. The reader should refer to [5] and [6] for a detailed presentation and discussion of the experimental results.

2. Interruption Taxonomy

As a basis for the framework a new Interruption Taxonomy is outlined to categorize a variety of traceable information needed to exhaustively describe the context space. The Interruption Taxonomy includes a set of abstractions that helps unify the issues previously considered by other researchers in a variety of different disciplines, proposes new ones, and suggests avenues for further exploration. By organizing the context information needed for mediating interruptions in a coherent framework, it attempts to improve the methodology of the design process. Interruption-related information is categorized according to context: Task Context, User Context, and Environment Context. A graphical representation of the taxonomy three-dimensional space is presented in Figure 1.

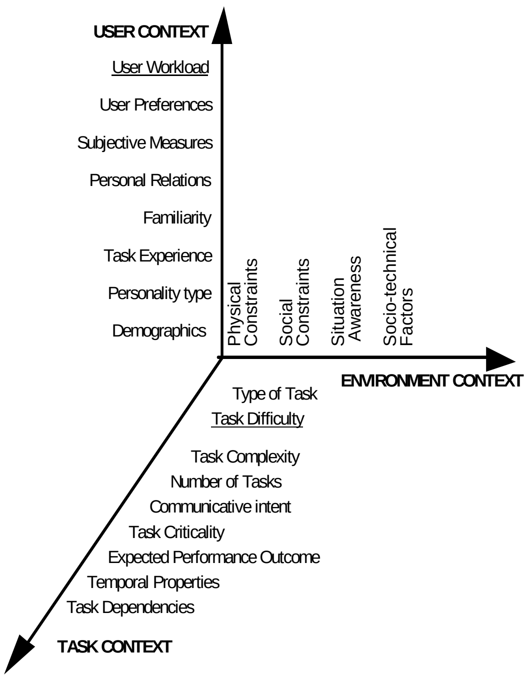

Figure 1: Interruption Taxonomy.

2.1. Task context

The Task Context dimension includes a number of attributes to capture the pragmatics of domain tasks. Some of these categories represent invariant properties of tasks that can be obtained from the domain-specific background knowledge about the application space, and the specifics of the particular interaction. Others should be drawn from the information gathered from a variety of sources (e.g., perceptive devices, interaction event tracing). It is clear that the inclusion of domain-specific user activities besides those related to human computer interaction are needed to provide a broader context view and to accommodate interaction with other devices in mediated spaces.

Task-related context knowledge is crucial in disambiguating the meaning and the relevance of the interruption task in regard to the current user endeavor. In general, the system knowledge of intelligent systems concerning users usually includes user's goals, plans, capabilities, attitudes, and knowledge [3]. The portion of the system knowledge concerning user's goals and plans, which are ultimately related to one another, is partially represented by the categories in the Task Context space.

2.2. User context

User-related taxonomic categories are included to support the representation and reasoning about a particular situation as the user views it. Finding solutions to the problems associated with interruptions that are effective with respect to some objective criteria (i.e., task characteristics) is necessary, but not sufficient [3]. To a large extent, the appropriateness of system behavior is also a function of the state of the individual's own comfort level. A user perspective and preferences may constrain the space of solutions and possible ways of handling interruption. The user-related categories in this taxonomy lie somewhere on a specialization scale from generic to individual. Generic categories target user groups, while individual user characteristics contain information specific to a single user (e.g., preferences, likes, dislikes).

2.3. Environment context

The Environment Context dimension extends the system context knowledge with information that captures the physicality and the dynamics of the working space where the human-computer interaction takes place. The inclusion of the environment-related categories attributes substantially greater sensitivity to the system, namely, the ability to adapt to a social setting, physical and organizational constraints, or the particularity of the current situation. Four types of environmental conditions are included in the taxonomy, but the addition of others is also possible.

Physical conditions - The effectiveness of the selected presentation modalities depends on the physical conditions in the surrounding space. Physical limitations could constrain possible modes and presentation techniques.

Social constraints – Detecting the presence of other individuals, and integrating the explicit knowledge of different types of well-known social settings will allow the system to recognize them, apply the appropriate social rules and constraints, and adapt its behavior accordingly. By integrating various social rules, the Interruption Taxonomy has the potential to support both social and task-oriented coordination of interruption.

Situation patterns - Recognizing situations associated with risk, forced choices, excessive workload and accountability is crucial since they are more likely to affect human behavior. Interruptions are considered natural accompaniments of crisis and predictors of system vulnerability. In presenting information to a user in an emergency situation, the challenge is to make the information accessible in a way that will improve her understanding of the situation.

Socio-technical factors - Human performance and closely related human workload are affected by certain organizational characteristics of the working environment and the system under consideration.

Two categories, User Workload and Task Difficulty, are theoretical constructs. This framework proposes mapping these categories to other context variables that belong to all three taxonomy dimensions. Much of the constructs of task difficulty and user workload (i.e., mental workload) are based on theoretical concepts and empirical findings in the relevant literature.

3. Interruption model

We have adopted an approach based on Bayesian Belief Networks (BBNs) to represent the causal relationship between different pieces of information and to integrate rules for how to use, maintain, and reason with interruption-related knowledge. The Bayesian network constructed for selecting the most appropriate timing of the interruption is shown in Figure 2. As shown, the decision on the most appropriate timing of interruption (i.e., Interruption Timing) depends on inferring the state of several hypothetical (non-observable) variables: *Interruption Relevance* (A), *Sensitivity to Interruption* (B), *Individual Differences* (C), *Environmental Conditions* (D) and *Urgency of interruption* (E). The circled areas in Figure 2 represent the parts of the graph that relate each of these variables with the relevant taxonomic categories. Most of the nodes in the network drawn as oval boxes correspond to the taxonomy categories.

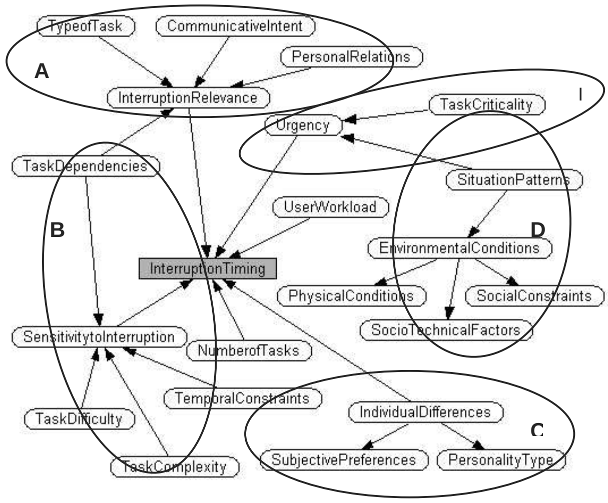

Figure 2: High-level dependencies between taxonomic factors for inferring the timing of interruption.

As stated previously, the theoretical constructs, User Workload and Task Difficulty are inferred using the interruption-related knowledge represented by the taxonomy. The BBN for inferring the difficulty of a task is presented in Figure 3. The set of categories selected from all three taxonomy dimensions could be broken down into three groups: (1) factors that are used to portray the "objective" difficulty of a task based on what is known about that task in general, (2) factors reflecting the particularities of a given situation, (3) characteristics to account for the individual (i.e., "subjective") perspectives on how difficult a task is, and (3) environmental influences. The justification for the selection of the characteristics is based on the theory and empirical evidence reported in relevant literature.

The BBN-model for inferring the user workload was constructed in a similar manner. It should be noted that Figures 2 and 3 depict variables and relations only on the highest level. In practice, parts of the network became quite complex by adding more exhaustive context representation.

To apply the proposed framework for mediating interruptions in a particular application domain requires: (1) a selection of the context variables from the taxonomy relevant in that particular domain; (2) an identification of the sources of information needed for sensing the states of variables; and (3) a construction and training of the Bayesian network which includes the relationships amongst the variables and the specification of the conditional probabilities implied by the relationships.

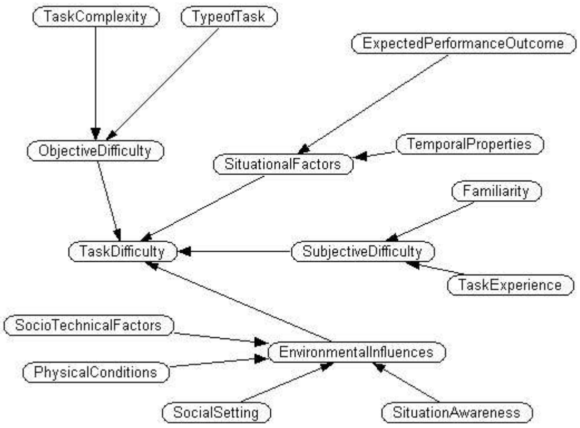

Figure 3: BBN-based model for inferring the difficulty of a particular task.

4. Implementing the Interruption Mediator

The first phase in implementing the interruption mediator was to create an appropriate context representation as a basis for mediating interruptions. The construction of the interruption model was tightly-coupled with the Interruption Taxonomy. In order to encourage greater clarity, it was decided to focus on the Task Context dimension.

The UI designers should select the set of variables, taxonomic categories, and their relationships after investigating and deciding on their relevance for mediating interruption in the system under investigation. A pilot study with fifteen participants was conducted to help construct the interruption model for this particular experimental environment. Domain-specific knowledge, subjective and objective measures gathered during the pilot study, were used to investigate the relevance of each taxonomic factor and to select a suitable way to categorize it.

4.1. Experimental tasks

As a test-bed application, we used a two-task experimental system developed at the US Naval Research Laboratory (NRL) in Washington D.C that has already been used for interruption-related studies. Two experimental tasks bear high resembles to military-like computer games and simulations. The primary (interrupted) task is a resource-allocation task named Three-Strike (TS). The objective is to attack and destroy three destinations utilizing available resources, ten heavy and ten light tanks and a certain amount of fuel and munitions (Figure 4a). On their missions, users are encountering resistance from differing locales and different kinds of obstacles based on a stochastic model of the TS task.

The interrupting task is Tactical Assessment (TA) task presented to a user at random points while she is performing the primary task. In this task, the user plays the role of a fighter aircraft pilot looking at a radar-screen-like display where three types of objects appear (Fugure 4b). The objective is to indicate whether the approaching object is hostile or neutral based on a specific set of rules. The decision of the “pilot” is assisted by an intelligent-automated component that colors the objects as red (hostile), blue (neutral) or yellow (when the assessment can not be made). The user is to confirm the hostile/neutral indications or give the appropriate classifications of the yellow objects based on a set of rules.

4.2. Interruption-related knowledge

One of the hallmarks of an intelligent, attentive system is that it attempts to help a user achieve her underlying goals, even when they are not explicitly stated. The principle way for a system to know a user’s underlying goal(s) is by recognizing observable user actions for achieving a domain-relevant goal. Thus, sensing user activities (not necessarily human-computer interaction), and identifying their characteristics and interdependencies is a way to reason about the user’s current goals and plans. Interruptions can arrive in a number of different scenarios, and it is of crucial importance to know the exact task context in which it occurs so that the mediator can select the most appropriate timing to interrupt the user.

Domain-specific knowledge was used as a basis for coupling and chunking the primary task structure, while subjective and performance measures gathered during the pilot study were used to validate the proposed organization. First, the structure of the primary task was divided into a set of higher-level subtasks using goal-based analysis. Five subtasks have been identified as potential cases for mediating interruption. A number of corresponding lower-level activities and interaction events were associated with each higher-level subtask. Lower-level activities were broken down to the lowest level of interaction events. All possible alternatives of interaction events needed to execute a particular lower-level activity were identified. This step was especially important so that the systems which tracked user behavior in terms of interaction events could recognize different subtasks and distinguish between various context situations.

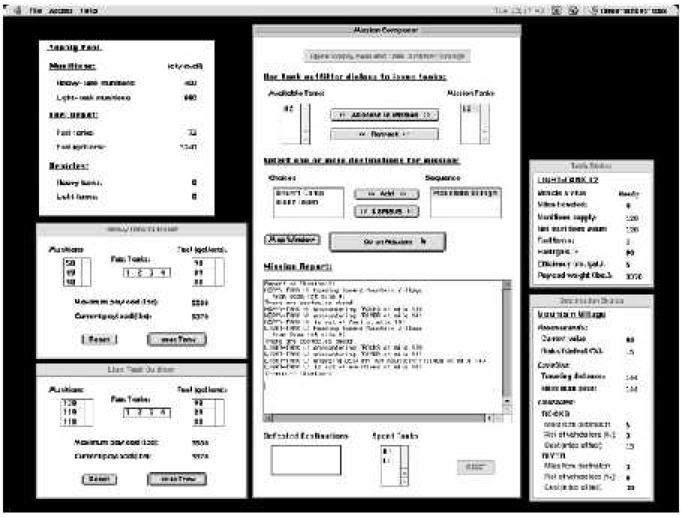

Figure 4. a) Primary Task Interface

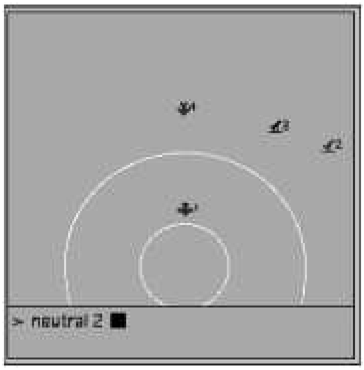

Figure 4. b) The Interface of the Interrupting Task

Once the tasks, subtask and interaction events were identified a suitable classification was needed to categorize them. Task categorization was level-specific and it was done in the following manner: higher-level subtasks were categorized in terms of task stages, such as planning, evaluation, execution as proposed in [9]. Domain-specific lower-level subtasks were actions associated with each task stage. General HCI categories were used to classify the interaction events on the lowest level (e.g., scroll, select, type, read, etc.).

Interruption requires coordination of human behavior, and since coordination is "a process of managing dependencies among activities" [8], an important factor for coordinating (mediating) interruptions is the identification of interdependencies between the primary and the interruption task. Even though the model of this particular dual-task setting was specified with no dependencies between the interrupted and the interrupting task, it was very important to describe the dependencies between primary subtasks to help in inferring the interruption sensitivity of each particular point.

The relationships between higher-level subtasks seemed to only partially describe a particular interruption context. Examination of lower-level relationship was needed especially for describing transitions between subtasks and cases when a subtask was executed partially or in a few iterative steps. Several functional relationships, such as *causal*, *producer-consumer*, *cooperation* were defined to describe higher levels of associations (represented by unidirectional or bidirectional block arrows in the figure). Lower-level relationships included categories such as *transfer*, *communication*, or *shared resources* [8]. Sometimes more than one applied. Pilot study participants were asked to rate the perceived strength of relationship between same-level and different-level subtasks that were used when assigning the causal probabilities between the nodes of the constructed BBN.

4.3. Constructing the BBN-based model

The selection and categorization of the taxonomic factors was followed by the phase of constructing the BBN-based model. The context factors were interrelated, and built into a decision network that decides when the most appropriate time to interrupt the user is. The conceptual interruption model was simplified to focus on the most essential factors needed to describe the task context, while user- and environment-related factors were left unexplored. Therefore, the parts of the network (Figure 2) that relate to *Individual Differences* and *Environmental Con*ditions were not operationalized for this research version of the interruption mediator. We could say that this version of the mediator bases its decisions mainly on how sensitive a certain context to interruption is, which is represented by the variable *Sensitivity to Interruptions*, its associated links and related taxonomic factors (area B in Fig. 2).

4.3. Data collection and evaluation

Instead of passive observation of human-computer interaction, the pilot study was conducted in a manner of invisible experiment by observing participant's behavior during normal interaction with the system. Inspection of the pilot study data indicated that the sensitivity of particular interruption points can be predicted based on the task-related context representation included in the Interruption Taxonomy.

The results of all participants were aggregated to distinguish effects based on the characteristics of particular situations (task context), rather than of individual participant. This allowed a larger training set to be used for BBN initialization and training than if the implementation was restricted to data from a single user.

For this particular implementation, the timing of the interruption was mapped and coded to a certain interruption point within subtask (e.g., before or after a particular interaction event). The system keeps records of the times and contexts of all relevant interaction events. Each time the specified situation arises (i.e., sequence of interaction events), the system infers the user's goals based on the context of her interaction before and at the time of the interruption. The system adjusts in such a way that a user is not interrupted during interruption points sensitive to interruptions (user performance degrades), deferring the interruption task for the next appropriate moment.

The Bayesian network was constructed and trained off-line with the pilot study data as a training set. The training set was used to adjust the corresponding causal probabilities of the BBN nodes. One particular state of the decision network (Figure 5) shows the observable variables and their assigned probabilities for the following situation; the user was interrupted when she opens the destination status dialog box after she reviewed the destination map. We have examined how much the states of the variables linked to Sensitivity influence the beliefs about that node. The results have confirmed that the difficulty of the task and the strength of the relationship were most likely to produce the greatest change in the belief of Sensitivity.

Figure 5 shows the links from Sensitivity and Interruption Timing to the Utility node (U – hexagon box), capturing the idea that a user will perform better when the interruptions coming at "highly-sensitive" points are deferred (utility = 89.8). On the contrary, if we consider situation in which the user performs "easy" subtask (complexity = *two*), loosely-coupled with the proceeding task (strength of relationship = *one*), the expected utility corresponding to each decision choice will change (not shown in this figure) and the best decision will be to present the interruption immediately (utility for Immediate = 92). The decisions for all relevant situations incorporated in this version of the interruption mediator were obtained with similar network reasoning. In general, the research version of the interruption mediator uses the outcomes of the trained Bayesian network, which helped to differentiate between sensitive situations that require deferring interruptions, and those for which immediate interruptions would be appropriate.

4.4. Lessons learned

The Interruption Taxonomy can help UI designers focus attention on relevant context factors important for mediating interruptions. It offers a wide variety of tangible factors to describe the context of interest with enough detail and depth to inferr the sensitivity of a particular interruption points. Our experience has shown that:

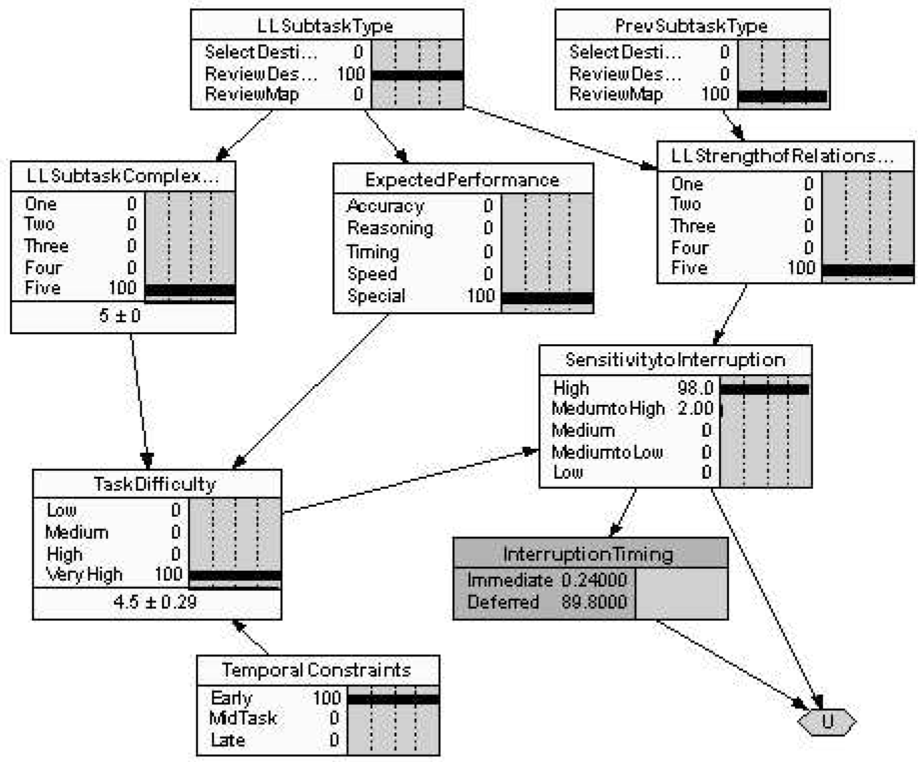

Figure 5. An instance of the decision network for inferring the interruption timing.

- Ø The analyses of individual users' records can help in the identification of relevant interruption contexts and situations. They also helped us in interpreting ambiguous results (e.g., "lengthy" pauses associated with reading that should be separated from the effects of interruption per se).
- Ø Task analysis may be needed to span across different levels of task abstraction to provide fine-grained task description. More than one classification may be selected from the set of existing and widespread-used taxonomies to categorize tasks and subtasks at different levels.
- Ø Applying recommendations and suggestions given by other researchers in the field is a not a straightforward task. The trends observed during the pilot study, and later confirmed by the formal experiment contrast to some previous results reported in related research works. For instance, the severe degree of disruption during complex subtasks was expected because of the cognitive load associated with them. What was perhaps more informative for our design was that interruptions at the beginning of certain subtasks subsequent to complex subtasks caused lengthy resumption times as well. The likeliest explanation could be that the user was still maintaining the goal state as a basis for subsequent related actions that are tightly-coupled with the finished subtask (e.g., outfitting a tank depends on users' strategy and plans established during the "planning" stage). The finding is in line with, but also somewhat contradictory to the suggestions that the most opportune moment for interruptions could be the moments associated with a start of a task [1], [2]. The contradictory results may be due to differences in our experimental designs, especially the types of tasks that were used. Gaining a user's attention

after one task is completed, but before a new task is initiated could apply to certain situations, such as performing independent non-related tasks [1], [2]. Whenever cognitively-taxing processes (e.g., reflecting, evaluating, planning) take place in-between related task or actions, care should be taken not to interfere with potentially fragile cognitive state of the user.

5. Future considerations

The first version of the interruption mediator coordinates interruptions in response to situations that can be defined using the task context characteristics. Mediator's interventions are based on explicit model of the task and situation rather than on a general or customizable user model. The results of the exploratory study have shown that the mediator could be used to improve task performance [5], foster situation awareness [5], and lessen the disruptive effects of interruptions on users' emotional states [6]. The model may require additional considerations to properly address requirements of different users for different situations.

Subjective perceptions and preferences were collected at the end of the pilot experiment. They measured (1) subjective perception of the appropriateness of interrupting at specific interruption points, and (2) subjective perception of the importance of mediating interruption at each particular subtask. User preferences were generally inclined toward choosing the end of a subtask as the most appropriate interruption point within each subtask. Not surprisingly, subjective preferences were not supported by actual objective measures. To the contrary, an interruption point placed at the end of a subtask led to longer resumption times, partially because of the effort to decide on what to do next, but moreover because of the existing relationships between subsequent subtasks. Subjective preferences were not considered as a factor in the current implementation of the interruption mediator. However, they should not be neglected when designing user interfaces that give equal priority to user's satisfaction and comfort as to other performance measures.

6. Conclusions

In this paper, we have discussed a framework for a computer-mediated coordination of human interruptions in HCI. The conceptual framework is based on the new Interruption Taxonomy and uses Bayesian Belief Networks as a decision-support aid. The prototype version of the interruption mediator was implemented to explore the design space, and identify the limitations and potential adjustments to the proposed interruption framework. The interruption mediator have succeeded in recognizing interruption points sensitive to interruption based on the taxonomic factors that exhaustively describe the tasks and the interaction.

References

[1] Bailey B. P., Konstan J. A., and Carlis J. V. Measuring the Effects of Interruptions on Task Performance in the User Interface. *In Proc. of SMC 2000. IEEE* (2000) 757-762.

[2] Cutrell E. B., Czerwinski M., and Horvitz E. Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance. In *Proc. of INTERACT 2001*, IOS Press. (2001) 263-269.

[3] Finin, T.W., "GUMS - A General User Modeling Shell", In Kobsa A. And Wahlster W. (Eds), *User Models in Dialog Systems*, Springer-Verlag, Berlin Heidelberg, Germany. (1989).

[4] Gillie, T., and Broadbent D. What Makes Interruptions Disruptive? A Study of Length Similarity and Complexity. *Psychological Review*. Vol. 50 (1989) 243-250.

[5] Gievska, S., and Sibert, J. Empirical Validation of a Computer-Mediated Coordination of Interruption. *In Proc. of OZCHI 2004*. (2004).

[6] Gievska, S., and Sibert, J. Examining the Qualitative Gains of Mediating Interruptions during HCI. I*n the Proc. of HCI 2005* (2005)

[7] Latorella K. A. Effects of Modality on Interrupted Flight Deck Performance: Implications for Data Link. *Proc. of HFES*. (1998).

[8] Malone, T., and Crowston, K. The Interdisciplinary Study of Coordination. *ACM Computing Surveys*, Vol. 26,1 (1994) 87-119.

[9] Miyata, Y., and Norman, D.A. Psychological Issues in Support of Multiple Activiies. In Norman D.A., and Draper S.W. (Eds.) *User Centered Design*, Lawrence Erlbaum Associates, (1996) 265-184.

[10] McFarlane D. C. Comparison of Four Primary Methods for coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction*, Vol. 17, 1 (2003) 63-139.