Developing Methods and Techniques for the Design of Cross-modal Interfaces

Metatla O., Martin F.
School of Electronic Eng. & Computer Sci.
Queen Mary University of London
{o.metatla, f.martin}@qmul.ac.uk

Stockman T., Bryan-Kinns N.
School of Electronic Eng. & Computer Sci.
Queen Mary University of London
{t.stockman, n.bryan-kinns}@qmul.ac.uk

ABSTRACT
There exists relatively little work on the design of cross-modal interfaces, that is, interfaces which support collaboration between individuals that use different sets of modalities to interact with each other. In this paper, we examine the role of design patterns based on two phases of a design process. Firstly, we examine the role of participatory design workshops in identifying patterns that arise out of conflicting requirements between different interaction modes. We then describe how an analysis of these conflicts can lead to pattern-based solutions to interactional and implementation issues in the design of cross-modal displays.

Author Keywords
Guides; instructions; author's kit; conference publications; keywords should be separated by a semi-colon.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Cross-modal interaction is fundamental to human perception, involving the coordination of information received through multiple senses to establish meaning [10]. An example of this is when we both see and hear someone talking and associate the words spoken with the speaker, thus combining information received from two signals through different senses. In the design of interactive systems, the term cross-modal interaction has also been used to refer to situations where individuals interact with each other while accessing a shared space through different modalities such as graphical displays and audio output [11, 9]. In this paper, we examine how design patterns for cross-modal collaboration can be identified. We describe how we used an approach based on activity patterns [6] to uncover design patterns from two phases of a typical design process. First, we examine the role of participatory design workshops in identifying patterns that arise out of conflicting requirements between different interaction modes. We then demonstrate the application of activity patterns by reflecting on the evaluation phase of a cross-modal tool that supports collaborative diagram creation and editing by visually impaired and sighted coworkers. We also show how an analysis of the conflicts revealed by an activity theory-based analysis of these patterns can lead to solutions to interactional and implementation issues in the design of cross-modal displays.

BACKGROUND
Despite significant progress in the use of the audio and haptic modalities in interaction design, research into cross-modal interaction has so far remained sparse. Initial investigations have nonetheless identified a number of issues that impact the efficiency of collaboration in cross-modal settings. For example, an examination of collaboration between sighted and visually impaired individuals on an interactive puzzle game highlighted the importance of providing visually impaired collaborators with a continuous display of the status of the shared game [12]. Providing collaborators with independent views of the shared space, rather than shared cursor control, was also found to improve orientation, engagement and coordination in shared tasks. In another study, a multimodal system combining haptic devices with speech and non-speech auditory output was used to examine collaboration between pairs of visually impaired users on graph reading tasks [8]. Results showed that the use of haptic mechanisms for monitoring activities and shared audio output improves communication and promotes collaboration.

Although scarce, the literature on cross-modal collaboration has begun to generate insights into the knowledge that is needed to come up with effective designs to support interactions involving individuals with differing perceptual abilities across various domains. We propose to use design patterns as a means to capture such knowledge so that it can be effectively leveraged to provide solutions to support accessible collaborative working.

APPROACH

Theoretical foundation
We consider activity patterns [6] as a potential guiding framework for identifying and implementing design patterns for cross-modal displays. According to this framework, Alexander's patterns [1] could be appropriated to embody the principals of Activity Theory (AT) and hence could be used to analyse activity in terms of understanding tool-mediated work in its context [2]. AT views human activity in terms of a system of tool-mediated actions carried out by a subject (i.e. an individual or a group of individuals) in order to achieve a desired outcome. Actions are characterised in terms of how they are organised within a community context, and how they are regulated by internal rules and mediated by a division of labour. This unit of analysis is conventionally represented by a triangular model to show how its elements interact with each other (see Figure 1).

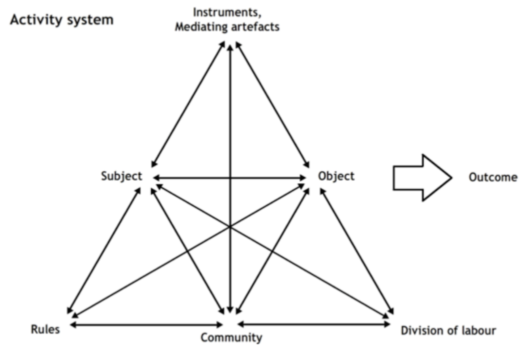

Figure 1. An activity system as conceptualized by activity theory.

According to [6], there are parallels between the design patterns principles as introduced in architectural design [1] and those of human activity as conceptualised by AT's unit of analysis. These include; the definition of a pattern in terms of three related components expressing the relationship between a given context, a problem and its solution, which is consistent with the method of AT; and the characterisation of a problem in a given context as being caused by a system of conflicting forces that arise in that context, which could be captured through AT's conceptual tool of contradictions. Additionally, the hierarchical levels of activity in AT can also be used in a similar way to the Alexandrian concept of scales to help structure the scope covered by activity patterns. The concept of activity levels in AT are hierarchically structured into three levels; activities, actions and operations where a given activity is realised through a set of concrete actions, which are in turn accomplished through a series of operations. This gives patterns a sense of scale from high-level activities down to low-level operations. According to [6], patterns could be written to reflect each element in a given activity system - the design of mediating artefacts; the work of a subject; the rules and procedures; and the roles within the division of labour or community of the work group - as well as organised into a coherent pattern language that preserves the unity of these elements within each of the three levels of activity.

In order to assess the feasibility of this organisational framework in supporting the process of identifying design patterns in cross-modal design, we applied it to data gathered from two phases of a typical process; requirements capture through participatory design workshops, and evaluation.

Initial participatory design workshops

The first stage of our approach involved setting up an initial workshop with 8 to 10 participants drawing from a network of users in the particular domain of focus. The workshop was organised around three main activities; focus group discussions, technology demonstrations, and audio-haptic mock-ups design. The aim of the focus discussions is to identify current best practice in the domain of concern, how current access technology supports this and a list of tasks which are either difficult or not possible using current access solutions. The technology demonstrations involve presenting a range of candidate technologies that could be used as a basis for designing solutions to the issues identified in the focus discussions.

Visually impaired users often have a good knowledge of the access solutions they personally use, but do not necessarily have direct access to or experience of other relevant technologies. It is important that the capabilities of a given technology are demonstrated without reference to an actual application. For example, in order to ensure an application-independent demonstration of the haptic devices, we used a custom program that allowed us to switch between different effects that could be simulated with these devices, such as vibration, spring effects and viscosity. The custom program allowed us to manipulate various parameters to demonstrate the range of representations and resolutions that could be achieved with each device in real-time. In the audio-haptic mock ups design phase of the workshop, we asked participants to think through new designs, having had hands-on experiences of the capabilities of the candidate technologies. In this phase, participants worked in small groups including one or two members of the design team to identify technology solutions to the problems that arose in the focus discussions. To close the session, participants presented the audio-haptic mock-ups they constructed with their group to the rest of the participants for further discussion.

An unexpected outcome of these initial workshops was that the invited users spontaneously agreed to sign up to a email list which from their point of view provided a forum for the sharing of best practice and workarounds, and for us provided a community forum for the discussion of design issues and a user group we could draw upon for participants in later formative evaluations.

We used the activity patterns approach to drive retrospective analysis of data gathered from the workshop. Participatory design activities generate a huge amount of data and patterns could help with the process of organising the themes that emerge from this data. The concept of contradictions can be a useful guide to identify the tensions that exist in activity systems constructed to model the requirements and scenarios described by workshop participants, and hence could lead to insights about design solutions that could resolve such contradictions.

Example: Conflicting requirements

In an example of such scenarios, while discussing his experience of working with sighted colleagues and clients, a visually impaired producer described his frustration with the inaccessibility of graphical and diagrammatic representations used in digital audio workstations. The visually impaired producer explained how his work with sighted colleagues often involves exchanging projects back and forth in order to complete sub tasks involving the manipulation of audio captured using inaccessible formats or requiring interaction with inaccessible audio plugins. In some cases, these accessibility issues have led to his exclusion from potential collaborative projects because the standard formats used are not readily accessible or would take too long to work with.

In a second scenario, a visually impaired participant who specialises as an accessibility trainer described a similar experience with inaccessible visual tools. In this case, the issue was specific to working on collaborative projects that were coded using an audio programming language known as Max/MSP, which is a visual programming language that uses diagrammatic representations as its main coding components. The visually impaired participant highlighted how inaccessible such programming languages are even though they are used to code audio, which could be considered a natural working modality for visually impaired individuals.

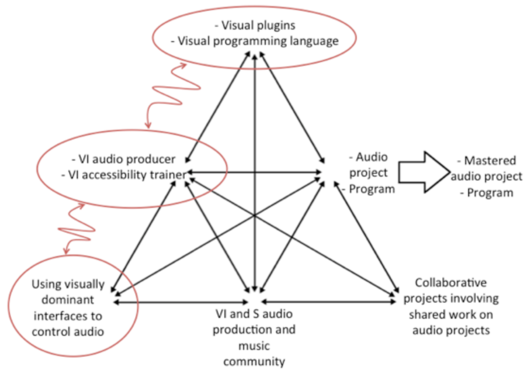

Figure 2. An activity system showing some contradictions captured through the PD workshop. Contradictions are highlighted with red circles and wavy arrows.

Both scenarios above could be captured by the activity system shown in Figure 2. Here, there is a clear contradiction between the subjects of the activity, i.e. the visually impaired audio producer and accessibility trainer, and the tools available to them as mediating artefacts in the context of their activities. Capturing these contradictions allows us to think about possible design solutions to eliminate them, which could eventually lead to the development of fully articulated design patterns that embody such solutions.

Identifying design patterns through evaluation

We used the activity patterns approach to reflect on the design of a diagramming cross-modal tool, which we evaluated with visually impaired and sighted users. The tool combines a visual diagram editor with auditory and haptic capabilities to allow simultaneous visual and non-visual interaction. That is, two coworkers collaborate on shared diagrams by accessing and editing them through the visual modality (for the sighted user) and the combination of audio and haptic modalities (for the visually impaired user) [9].

We deployed this collaborative tool in various workplaces including a local government office and a charity organisation where visually impaired and sighted coworkers access and edit diagrams as part of their daily jobs [9]. In the following, we describe an example of applying the activity pattern approach retrospectively to analyse data that we gathered from one of the field studies and how this helped the process of identifying potential patterns for cross-modal design.

Example: Consistency of interaction steps

In this example, a visually impaired manager (VI) and their sighted assistant (S) at a local government office edit an organisation chart to reflect recent changes in managerial structures. At one point during the interaction, the pair decides to create a connection between two nodes on the chart diagram to highlight a relationship between an existing and a new position. They do this while discussing how the tool should be used to create this relationship.

To create a connection between two nodes using the non-visual audio-haptic editor, the visually impaired user must 1) browse the chart to locate the first node and select it, 2) browse the chart again to locate the second node and select it, 3) select the type of connection they wish to use and 4) issue a command to create the desired connection. To do the same in the graphical editor, the sighted user must 1) select the type of connection they wish to use from the graphical tool bar, 2) select the first node on the chart, 3) drag the connection towards the second node using the computer mouse, and then 4) release the mouse to create the connection. Following the activity patterns approach, the actions of creating a connection between two nodes using the visual and non-visual editors can be represented as the two independent activity systems shown in Figure 3, which highlights a contradiction between the operational rules in the two activity systems; there is a mismatch between the interaction steps that each collaborator has to follow in order to create a connection between two nodes on the chart. Modelling the collaborative action of creating a connection in this manner has therefore uncovered a potential design flaw - which manifests itself as contradictions - that could hinder collaboration. Addressing this design flaw could lead to a design pattern that can eliminate the issues raised by the contradictions. For instance, the potential design pattern could describe the need to reconcile the two mediating artefacts in this context by ensuring *Equivalence and Consistency of Interaction Steps* between the visual and non-visual modalities.

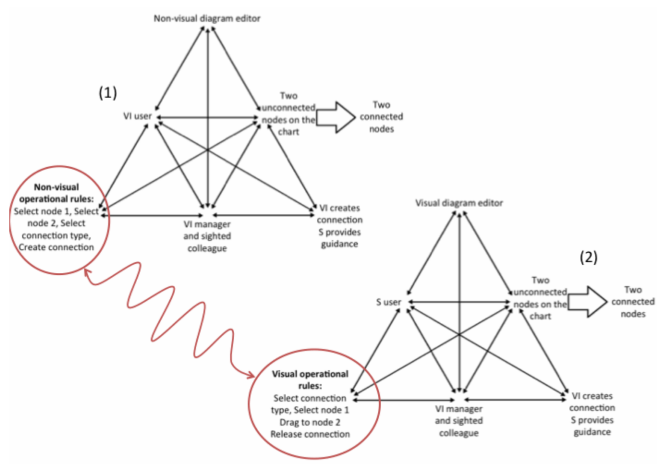

Figure 3. Two activity systems for creating a connection between two nodes using the non-visual (1) and the visual (2) editors. Contradictions are highlighted with red circles and wavy arrows.

Interaction and implementation design patterns

The conflicts highlighted when applying the activity patterns approach have also led us to identify a number of design patterns that could be used to guide the development of cross-modal software systems. Often, we found that interactional needs must be echoed in the actual implementation of such cross-modal systems. In particular, we explored how existing software design patterns (e.g. [4]) could be extended in order to accommodate needs that are specific to cross-modal interface implementation. Here too, the conceptual tool of contradiction has helped our analysis and guided the process of matching interaction and implementation design patterns.

Example 1: Managing input and output conflicts

For instance, cross-modal software applications must be able to manage multiple input and output streams through various controllers. The Observer pattern is particularly suitable for this kind of applications [4]. That is, input and output streams can be easily modelled as controllers [1] and observers respectively. However, this creates potential conflicts when users attempt to interact with shared content. Care must therefore be taken in order to avoid situations where changes in one modality trigger unwanted effects in another.

An example scenario where the above issue could arise is when editing the positions of objects on a interface. In our cross-modal application, moving diagram items can be accomplished using a mouse drag-and-drop in the visual modality. In audio, the same task could be accomplished by selecting an item and using the keyboard arrow keys to change its position. Visual feedback of movement is realised by refreshing the visual display and updating the coordinates of the item in real-time, whereas that of audio is realised by displaying audio feedback in response to each keyboard input stroke.

The refresh rate of mouse input is typically sampled at a high rate in order to ensure smoothness of movement, which in turn yields a high volume of data coordinates. But while graphical displays can cope well with such high frequency updates, an auditory display may end up with a fairly high quantity of sounds to display which, if not managed, could amount to pure noise. We have employed an analysis of contradictory conflicts using the activity patterns approach to derive a variation on the Observer pattern that we dubbed *Cross-Modal Observer*. This pattern is an example of how an interactional need - in this case having equal access to a shared space - requires carefully engineered implementation. The Cross-Modal Observer pattern builds on the original Observer pattern by introducing a reference to the source of the input controller that could then be used to filter out unnecessary data output streams. Thus, unnecessary output, such as displaying audio in response to actions issued through the graphical display, can be filtered out to enhance the usability of the auditory display as well as the overall performance of the cross-modal system.

Example 2: Supporting awareness

Another example of where an interactional need could be echoed in, and hence help establish a cross-modal implementation pattern is the need to support awareness. Maintaining awareness is critical in group collaboration [5] and should therefore be adequately accounted for in cross-modal collaborative systems.

[1]This term comes from the *Model-View-Controller [7]* software design pattern.

In a collaborative system, accessing shared resources concurrently can lead to inconsistencies in the underlying data model representation. In order to address this, a locking system can be used to synchronise the interaction by forcing any user-issued editing command to acquire a lock on a piece of data before changing it. In our collaborative application, this meant that every edit message must be preceded by another message carrying a lock request. Once finished with the data, a lock release message is then sent to the server in order to free resources.

When requesting/releasing a lock, the minimum information that a message must contain is that of the source as well as the specific item that needs to be locked/unlocked for editing. From the point of view of the interaction, this information specifies the subject and the object of the editing action. Since the lock request and release messages are sent both when an edit action is initiated and terminated, the locking system mirrors the interaction of the user with the system, and can thus be used to convey awareness information about each user's actions to other collaborators. We thus derived a design pattern which we dubbed *Lock Driven Awareness* to implement and support awareness in cross-modal collaborative displays. This pattern allows developers of cross-modal collaborative systems to leverage the locking technique, which is often needed to manage access to shared resources. Additionally, information about the source of an action and its object of interaction could be augmented to include more detailed awareness information, such as the type of action and how it affects the content: move, rename, delete etc.

DISCUSSION

The design of cross-modal collaborative systems presents a unique set of challenges because such systems must allow individual users to equally contribute to the shared tasks while accommodating their individual perceptual differences. To date, no research has examined how to capture the knowledge required to design technology that makes cross-modal collaboration easier.

In our approach, the participatory workshops played a key role in identifying barriers to collaboration in the respective domains and identifying potential solution stems for these problems. The concepts of conflicting forces and contradictions in AT have proved valuable in highlighting mismatches or incompatibilities in cross-modal interaction. Typical mismatches we have encountered include:

1. Mismatches in the series of actions required to achieve the same result through different interfaces to a system.
2. Mismatches between the representation of actions performed in one interface and the way in which those actions are represented in another interface.

Having used AT-based analysis to uncover the mismatches or contradictions, we have found that further analysing the contradiction with the aim of defining a pattern solution can lead to patterns at either the design or implementation level which can remove the issues raised by the contradictions. Furthermore, pattern solutions derived in this way can be sufficiently abstract to be considered as providing useful guidelines for design beyond the specific domain of interest. For example, ensuring *Equivalence and Consistency of Interaction Steps* between different interaction modes, and implementing *Lock Driven Awareness* to implement and support mutual awareness in collaborative systems, are design techniques which appear appropriate and applicable beyond the design domain of diagramming systems. Design patterns and pattern languages have been shown to facilitate the capture, presentation and communication of design knowledge [3]. We postulate that designers and application domain experts in cross-modal displays could benefit from using an activity patterns approach as both a means to identify and address design issues as well as a uniform representation for the design knowledge they generate.

CONCLUSION

We proposed that designers and application domain experts in cross-modal collaboration could benefit from using design patterns as a uniform representation for expert knowledge. To this end, we explored the question of how potential patterns can be uncovered from an iterative design process and suggested that activity patterns could be used as a structured method to address this question. One of the key benefits of using activity theory to identify patterns is the conceptual tool of contradictions, which can be a useful guide for designers to identify the tensions that exist in their designs when used in context and modelled as activity systems. We have exemplified how this approach was useful for us in managing requirement data from participatory design workshops as well as in the evaluation phase of a cross-modal collaborative tool. We plan to use this approach to both generate and articulate an initial set of patterns to form a pattern language for designing cross-modal collaboration, which we will then validate by applying the patterns in future design iterations and incorporating them in future studies and design and evaluation activities.

REFERENCES

1. Alexander, C., Ishikawa, S., and Silverstein, M. A pattern language: Towns, buildings, construction (center for environmental structure series).
2. Bertelsen, O. W., and Bødker, S. Activity theory. *HCI models, theories, and frameworks: Toward a multidisciplinary science* (2003), 291–324.
3. Borchers, J. O. A pattern approach to interaction design. *AI & Society 15*, 4 (2001), 359–376.
4. Gamma, E., Helm, R., Johnson, R., and Vlissides, J. *Design patterns: elements of reusable object-oriented software*. Pearson Education, 1994.
5. Gutwin, C., and Greenberg, S. Design for individuals, design for groups: tradeoffs between power and workspace awareness. In *Proc. of the 1998 ACM conference on Computer supported cooperative work*, ACM (1998), 207–216.
6. Guy, E. S. "...real, concrete facts about what works...": Integrating evaluation and design through patterns. In *Proc. of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, GROUP '05 (2005), 99–108.
7. Krasner, G. E., Pope, S. T., et al. A description of the model-view-controller user interface paradigm in the smalltalk-80 system. *Journal of object oriented programming 1*, 3 (1988), 26–49.
8. McGookin, D., and Brewster, S. An initial investigation into non-visual computer supported collaboration. In *CHI'07 Extended Abstracts on Human Factors in Computing Systems*, ACM (2007), 2573–2578.
9. Metatla, O., Bryan-Kinns, N., Stockman, T., and Martin, F. Supporting cross-modal collaboration in the workplace. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, British Computer Society (2012), 109–118.
10. Spence, C., and Driver, J. Cross-modal links in attention between audition, vision, and touch: Implications for interface design. *International Journal of Cognitive Ergonomics* (1997).
11. Winberg, F. Supporting cross-modal collaboration: Adding a social dimension to accessibility. In *Haptic and Audio Interaction Design*. Springer, 2006, 102–110.
12. Winberg, F., and Bowers, J. Assembling the senses: towards the design of cooperative interfaces for visually impaired users. In *Proc. of the 2004 ACM conference on Computer supported cooperative work*, ACM (2004), 332–341.