From tasks to conceptual structures: misfit analysis

Ann Blandford

School of Computing Science
Middlesex University
Bounds Green Road
London N11 2NQ. UK.
A.Blandford@mdx.ac.uk
Telephone: (44) 20 8411 6163
Fax : (44) 20 8362 6411

Thomas Green

Computer Based Learning Unit
University of Leeds
Leeds LS9 2JT, UK

Thomas.Green@ndirect.co.uk

SUMMARY

Ontological Sketch Modelling (OSM) is an approach to usability evaluation that focuses on structures rather than tasks, and assesses the goodness of fit between the user's conceptualisation of a system and that actually implemented by the designer. It provides a common representation that supports reasoning about users, domains and devices. In this paper, we introduce the rationale behind OSM, the approach itself, and a short case study of applying it to an electronic diary system.

KEYWORDS : Ontological Sketch Modelling, OSM, usability evaluation, misfit analysis, diaries.

INTRODUCTION

As is widely recognised, HCI is a complex area of study. There are a huge number of design and evaluation techniques, each of which is well suited to addressing some aspects of the problem, and inadequate for others. Task analysis in one shape or another dominated most techniques of HCI usability evaluation for many years, e.g. [1,7,11,14]. Task analysis techniques are typically useful for situations where there are a well-defined set of procedures users can adopt for achieving their goals. More recently, there has been more interest in work domain analysis to inform design, e.g. [3,15]. Such techniques are proving powerful in implementing systems that accurately reflect domain concepts. The approach we present here, Ontological Sketch Modelling (OSM), concentrates on the user's conceptual models of the device, the domain and working practices (that is: how the device fits in with the way the user works). Thus, it is concerned with domains *and* devices *and* users – but focuses on concepts rather than tasks or processes. Analysing the degree of fit between these can reveal potential problems of a semantic type that are not revealed by existing HCI approaches.

MODELLING IN OSM - AN OVERVIEW

In principle, when designing a new system, preliminary analysis identifies current use and future requirements, using observation, interviews, etc. That analysis yields a model, or specification, to be used as a basis for design. However, in the design process:

- the design team often find that some features can only be implemented by introducing new device concepts (e.g. handles for manipulating drawing objects, scroll bars for moving through a document)
- the design team may also realise that some things can be done 'better' if they introduce new device concepts (e.g. memory in pocket calculators, programmability concepts for VCRs), and those become part of the 'domain of the device'.

In practice, initial requirements gathering is generally incomplete. In addition, the co-evolving nature of technology adoption results in new concepts emerging through use of technology. These are all sources of misfits that may introduce usability difficulties. In principle, OSM can be used for both evaluation of existing systems and evolutionary design of new ones. Here, we focus on evaluation.

In OSM the modeller describes the visible entities, their attributes, and how they are linked within the device; and also describes the entities contained in the user's conceptual model and those embodied within the device that the user needs to be aware of if they are to use the device effectively. The resulting entities may be private to the device (the user cannot alter them), or they may be private to the user (the device does not represent them), or they may be shared (accessible to both the device and the user). All communication between the two worlds of user and device takes place through the shared entities. If the user-private entities do not fit well onto the shared entities, the device will have usability problems. Similarly, if the device-private entities are difficult to discover or understand, the user is likely to have difficulty learning to work with them.

Some of these difficulties may reflect substantial mismatches in representation; for example, the user of a video link may be compelled by the system to define the

Quality of Service in terms of bandwidth, while they think of it in terms of frame rate or image resolution. (In OSM terms, bandwidth is a *shared* entity, while frame rate and image resolution are *user-private*; changing bandwidth *affects* frame rate and image resolution, but only indirectly. There may also be *device-private* concepts that determine how bandwidth affects the related entities.)

Other misfits are less glaring, such as:

> *Inexpressiveness*: the user wishes to do something that is part of the conceptual domain but cannot be expressed using the device. *Example:* electronic organisers cannot indicate relative importance of engagements, nor links between them, which are important components of diary use [14,4] – these concepts are *user-private*.
>
> *Viscosity*: the user thinks of a single direct operation but the device requires an indirect operation, or a large number of operations [9]. *Example:* updating cross references in a standard word-processor is repetitive and tedious because the constraint on ordering of references is *user-private*, but is *affected by* insertion or deletion of references.

Misfits cannot be revealed by any approach to HCI that focuses solely on either the user or the device. Traditional task-centred user-modelling for HCI has some very effective results, but it cannot reveal misfits because it does not explicitly consider how the user's conceptual model of the domain and device relates to the model imposed by the device.

By focusing on entities and attributes, it is possible to consider the design and domain at different levels of detail without exponential explosion in the representation. For instance, when evaluating a drawing package for use in a particular setting (e.g. a school classroom), for every domain concept (e.g. a picture) there will be a limited number of intermediate concepts (e.g. a drawing-object, a line, a rectangle), a limited number of device concepts (e.g. the 'handles' on a drawing-object, the interface buttons) and a limited number of relationships (e.g. a picture consists-of drawing-objects) to be reasoned about. This contrasts with task-centred approaches, where the space of possible tasks becomes unmanageably large for devices such as drawing tools.

An Ontological Sketch Model describes a device and its domain of use in a form that is deliberately sketchy (i.e. precisely stated and structured, but not necessarily complete) but sufficient to indicate misfits. It thereby avoids the 'death by detail' that has characterised many previous evaluative techniques and has deterred human factors practitioners from using them [2]. Being a sketchy approach, OSM is potentially quick, both to learn and to use; moreover, modelling need only encompass those aspects of the device that the modeller considers relevant, and when appropriate the model can be relatively coarse-grained. A further advantage is that OSM models are re-usable. If part of a design is incorporated in another device, the relevant part of the old OSM model can be incorporated in an OSM of the new device. For example, when evaluating a diary design, the user concepts of events and appointments, with attributes such as importance, intended start time, likely duration, etc. can be re-used for evaluating alternative device designs.

RELATED APPROACHES

OSM has been much influenced by two seminal papers laying the foundations of misfit analysis (although not by that name) whose potential was never fully realised. Moran's [12] ETIT analysis mapped the 'external' task of the domain onto the 'internal' task of the device, from which an efficiency metric could be computed, essentially the number of device actions required to achieve one domain-level goal. Payne's [13] 'Task-Entity Analysis' listed conceptual entities named by users of electronic diaries, supplemented by entities inferred from psychological research (notably, the analysis of intentions), and showed that whereas paper-based diaries could express subtle but important differences between types of intention, electronic ones lacked the required expressiveness. Payne's version of misfit analysis was entirely informal and could not be attempted without knowledge of HCI and cognitive psychology; thus, his approach was not usable by professional designers.

OSM's choice of what to represent draws on previous work that explores representations of misfits. To predict misfit, an explicit description of the device or the information structures is needed. The representation of ERMIA (Entity Relationship Modelling for Information Artefacts [8]) can exhibit some kinds of misfits between users' conceptual models and the structures imposed by the device; OSM builds on that but is more expressive, while also being deliberately sketchier and less rigorous in detail. Since OSM includes the consequences of actions, it can be used for analysing a wider range of misfits than can ERMIA.

To represent the user's knowledge OSM draws on Programmable User Modelling (PUM: [5,16]). For a given device, PUM models of goal-based behaviour include a detailed list of operators and their preconditions. Cognitive modelling techniques are applied to the modelled user's knowledge to determine likely user behaviours. OSM demands much less rigour

than a full PUM knowledge analysis, and allows a wider range of concerns to be expressed within the model. In addition, OSM makes explicit the need for knowledge elicitation from users (or potential users), thereby grounding it much more firmly in real-world data.

CASE STUDY: EVENT SERIES IN DIARIES

Electronic diaries have been studied extensively over many years, and yet the current generation of such diaries are widely perceived as being difficult to use. Blandford and Green [4] present the results of a study with 16 users that investigated their use of such diaries, considering also their use of other time management tools. The findings presented there are not analysed in terms of OSM, but are informal; the main findings are that existing tools treat time management as a scheduling activity (whereas it is often very fluid and dynamic), that it is important to understand and appreciate the properties of different tools and move towards integration rather than expecting to create a single universal time management tool that satisfies all purposes, and that the requirements on individual and group time management often conflict in ways that require detailed attention when introducing group calendars into an organisation. Here, we focus on one particular feature offered by most electronic diaries: the capability to enter a series of events with the same properties (e.g. day of the week, description, time and duration).

All users in our study had access to the same shared electronic diary, which we will refer to here as MM. In MM, as in many similar diary systems, a repetitive activity is represented as an activity with a frequency. Each activity in the series inherits all attributes except date from the initial instance of the activity. Dates for series members are set according to user-defined parameters (e.g. daily, weekly, on the 4th Thursday of the month, or on the 15th of the month). It is not possible to set irregular intervals (such as "alternate Tuesdays and Thursdays"), but most regular frequencies can be defined.

This frequency feature was proposed by Kincaid *et al* [10], following their evaluation of earlier diary systems. It would probably not naturally emerge as a requirement if a domain analysis studied how users use paper diaries; for example, Beyer & Holtzblatt [3] discuss the fact that many users of paper diaries draw a line across several days to indicate a multi-day event, and propose that electronic diaries should therefore support the concept of a multi-day event (rather than a single-day event repeated over several days); this is a much more obvious feature of a paper diary than the fact that similar events may crop up at regular intervals. Indeed, because paper diaries do not support a 'recurrence' feature, people often do not transcribe recurrent events like birthdays into their paper diaries, but either keep a separate list or memorise them [4]. In our study, we found that many users reported avoiding using the frequency feature in MM. Elsewhere [6], we have conducted a task-oriented analysis of this feature, and shown why using the frequency feature is more error-prone than entering events singly, even though it is almost always more efficient. Here, we outline an OSM analysis of the same feature. Rather than presenting the OSM notation in the space available, we describe it using natural language.

In OSM terms, an event, which is a *shared domain-relevant entity*, is *created* in the diary through the *action* of inserting a new event or, alternatively, by applying the frequency attribute of a single event. An event is (typically) *deleted* by manipulating the single event. For events with a frequency attribute, options to delete the whole series or all future events in the series are also available. The event has various *attributes* (title, start time, duration, frequency, whether a reminder is sent tot he user, etc.), each of which can be *initialised* or *modified* (though many have default values).

In MM, a change to one member of the series will normally change all other members. Under certain conditions one member of a series may be detached and become an individual activity with no remaining links to the series. It is not possible to distinguish visually between a single activity and a series of linked activities most of whose members are off-screen, nor between a normal series and a pseudo-series in which one or more members have become independent. Thus the device embodies both hidden dependencies and hidden **in**dependencies, the former because changing one activity may cause invisible effects to others, and the latter because changing what is apparently a united series may omit activities that appear to be, but are not, linked. Here we have a severe mismatch between a conceptual entity, the series of activities, and its device representation.

The users' view of repeated activities, as derived from interview data, is considerably richer than the device-centred view of a single activity with a frequency attribute. Users did not report an activity that happens regularly, but a series of activities — for example, a lecture or seminar series or regular (monthly or quarterly) meetings — that is made up of a sequence of individual activities. That is: an activity-series *consists-of* activities. In this case, each activity inherits many attributes from the "prototypical activity" (such as location, participants, duration, purpose), but each may also vary in several respects. An activity series has a default "signature" (the features that, to the user, identify

it as a member of the series), but each event within the series may vary in terms of attributes other than just date.

The typical user's conceptualisation is of an activity-series; this is *user-private* (not explicitly represented within the device); an activity-series can be *added* to the diary by creating an activity with a frequency attribute, or by entering each event separately, and deleted by delete all instances. The series may have various attributes, such as expected participants, default location, default time, duration, day / date and purpose / function. The poor fit between the user's conceptualisation of a typical activity-series and that supported by the diary system is another reason for users electing not to make use of the frequency feature.

By partially implementing the abstract notion of a series of activities / meetings in a non-inspectable way, the developers of MM have introduced an inscrutable dependency structure. However, this discrepancy, or misfit, can only really be understood by laying out both user and device centred perspectives using the same representational structure.

The system designers apparently did not fully understand this feature, as illustrated by the fact that when an event has acquired a frequency attribute (e.g. weekly), even if there is only one event left in the series, the system asks the user whether to delete just the current event or all instances of the event. The underlying implementation appears to have an undocumented concept of whether the frequency attribute has been set attached to every event; this results in the surprising behaviour of the system when there is only one event remaining in the series.

CONCLUSIONS AND FUTURE WORK

OSM is at an early stage of development ; initial studies indicate that it has potential to support *reasoning about misfits*, an important property of systems that is missed by existing approaches to usability evaluation. Future work will address the scope of the technique – what is it useful for, and what does it miss ? – and also the usability, for any technique that remains a tool only for academic researchers is ultimately sterile.

ACKNOWLEDGEMENTS

This work is supported by EPSRC Grant GR/R39108.

BIBLIOGRAPHY

1. Annett, J. & Duncan, K.D. Task analysis and training design. *Occupational Psychology.* 41. 211-221. 1967.

2. Bellotti, V. Implications of Current Design Practice for the Use of HCI Techniques. In D. Jones & R. Winder (Eds.) *Proceedings of HCI'89,* 13-34. Cambridge University Press. 1989.

3. Beyer, H. & Holtzblatt, K. *Contextual Design.* San Francisco : Morgan Kaufmann. 1998.

4. Blandford, A. E. & Green, T.R.G. Group and individual time management tools: what you get is not what you need. To appear in *Personal and Ubiquitous Computing*.

5. Blandford, A. E. & Young, R. M. Specifying user knowledge for the design of interactive systems. *Software Engineering Journal*. 11.6, 323-333. 1996.

6. Butterworth, R., Blandford, A. & Duke, D. (2000) Demonstrating the cognitive plausibility of interactive system specifications. *Formal Aspects of Computing* 12. pp. 237-259.

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

8. Green, T. R. G. & Benyon, D. The skull beneath the skin: entity-relationship models of information artifacts. *International Journal of Human-Computer Studies*, 44(6) 801-828. 1996.

9. Green, T. R. G. The cognitive dimension of viscosity: a sticky problem for HCI. In D. Diaper, D. Gilmore, G. Cockton and B. Shackel (Eds.) *Human-Computer Interaction – INTERACT '90.* Elsevier. 1990.

10. Kincaid, C. M., Dupont, P. B. & Kaye, A. R. Electronic calendars in the office: an assessment of user needs and current technology. *ACM Trans. Office Information Systems*. 3.1. 89-102. 1985.

11. Johnson, P. J. *Human computer interaction,* McGraw Hill. 1992.

12. Moran, T. P. Getting into a system: External-Internal Task Mapping Analysis. *Proc. CHI'83.* 45-49. 1983.

13. Payne, S. J. Understanding calendar use. *Human-Computer Interaction.* 8. 83-100. 1993.

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

15. Vicente, K. *Cognitive Work Analysis.* Mahwah : LEA. 1999.

16. Young, R. M., Green, T. R. G. & Simon, T. Programmable user models for predictive evaluation of interface designs. In *Proceedings of CHI '89*, 15-19. 1989.