

INFORMATION DOMAINS IN CSCW

Pippa Hennessy

Communications Research Group
Department of Computer Science
Nottingham University
Nottingham, NG7 2RD
England
email: pah@cs.nott.ac.uk

Abstract

Most abstract models of cooperative work concentrate on the *procedural* and *structural* aspects of group working. However, one of the fundamental processes that occurs when people work together is that of *information sharing*. This paper suggests that a complete model of cooperative work would benefit from explicitly incorporating the concept of *information domains* in such a model. It is proposed that this can be achieved by including *information domains* in such a model. Two projects (MacAll II and AMIGO MHS+) that have addressed this issue to a limited extent are outlined and compared. Finally, four areas of consideration are identified and discussed, and areas for further research are highlighted.

1 Introduction

Over the last few years there has been a considerable amount of research devoted to the subject of computer supported co-operative work (CSCW). Many areas have been addressed within this field, ranging from the development of underlying technologies to support CSCW, through studies of group working in general, to the proposal of abstract frameworks for modelling group communication using computers. This paper discusses an issue that has been touched on by many different researchers working in widely varying areas, yet has not been fully explored in its own right.

A key consideration for any approach to CSCW is that of *information*, in particular, the sharing of information. Underlying support for this has been developed (e.g. [11]), and there is much ongoing research into more efficient ways of managing the storage and access of information over large computer networks. However, a coherent model of information sharing within an organisation has yet to be developed.

This paper will demonstrate that, not only is a model of information sharing necessary, but that such a model can be integrated with models of procedural and structural aspects of group working in order to develop a complete and coherent description of CSCW. It will also identify and discuss some of the issues involved in modelling information sharing.

The concept of an *information domain* is introduced to facilitate the modelling of information sharing. An information domain is a grouping of people and/or roles, information objects, and other resources. This grouping represents a working group in real life, that is, it brings

together people, information, and necessary tools and services so that a goal or set of goals can be achieved.

The structure of the paper is as follows:

- Section 2 demonstrates the need for a model of information sharing, and suggests that the information domain (ID) could be the cornerstone for such a model.
- Section 3 reviews the work of two projects (MacAll II and AMIGO MHS+) that have included the concept to some extent.
- Section 4 discusses some issues involved in the development of a framework for describing IDs, and suggests some solutions to problems that arise.
- Section 5 concludes the paper by summarising the major issues involved in modelling information sharing, describing how the model could be integrated with procedural and structural aspects, and identifying areas for further study.

2 Why are information domains needed?

This section demonstrates the need for a concept of information domains in CSCW. There are three reasons for this need; the main reason is to facilitate the development of an explicit model of information sharing. The other two reasons are, first, to provide a series of contexts for naming objects, and second, to define a collaborating group and provide an explicit environment in which this group can work. The remainder of this section will discuss these reasons in more detail.

2.1 Information domains for information sharing

Many abstract models of group working have been put forward in recent years. These have been formulated using several different approaches, which can be categorised in many ways (e.g. [6], [17]). Group communication has been modelled from the point of view of the data involved (e.g. [13], [15]), the agents involved (e.g. [1], [7]), the existing communication structure (e.g. [22]), and the processes involved both at the level of individual turns in an activity (e.g. models based on speech act theory; [2], [8], [9]), and at the level of describing the activity itself (e.g. [5], [16]).

All the above models of group working describe its *procedural* aspects. Most of them also include some concept of the *structural* aspects of communication, for example, the AMIGO Activity Model [16] makes use of Tschritzis' and Gibbs' CANPLAY and HASPATH relations [22]. However, a striking omission from all these models is the notion of *information sharing*.

When a group of people forms with the intention of working together to achieve a particular goal or series of goals, some sort of context is implicitly or explicitly created within which this group will work. Members of the group will use this context to share information and tools pertinent to the tasks involved in attaining the goal(s). Goodman and Abel [10] call this context a *project space*. Their work involved looking at groups, focusing on patterns of communication and realisation of project spaces. Groups that they looked at were working groups within a department, rather than departments within an organisation, but it seems

likely that larger groups with less well-defined goals will also use contexts similar to project spaces to share resources.

It seems, then, that information domains are generally used in some form by groups of people working together. If communication within these groups is to be supported by a computer system based on a model of processes involved in group working, IDs will still be created and used to share information objects and resources. Therefore, it makes sense to provide support for this by including the concept of IDs in the particular model used, and thus providing an integrated support service.

2.2 Information domains for object naming

In addition to the advantages of using explicitly defined information domains to support information sharing, an ID may be integrated with an object naming scheme. Object naming is usually considered to be entirely local (the meaning of a name is private to individuals), or entirely global (a name means the same to all users). Most systems using a global naming scheme allow individuals to give private aliases to objects, but in a group working environment this does not give enough flexibility.

Sollins and Clark [21] propose a naming scheme that allows self-defining groups to jointly create and use names. The scheme relies on arbitrarily defined *contexts*, which are associations between groups of people and sets of names for objects. A name can be used to access the object it refers to, and it acts as a placeholder for the object. The attributes of contexts could easily be associated with IDs within a model of group working, as it is likely that groups defined by the two methods will overlap to a significant extent.

Kreifelts and Seuffert [14] describe the addressing scheme used in the DOMINO office procedure system. This relies on the definition of *organisational units*; which are groups of people within an office. Each organisational unit may contain other organisational units, or a set of projects, which are groups of people with specific goals. People may be referred to by their position in one of these units. A logical progression from this definition is to allow the inclusion of information objects in these groups, thus incorporating the information sharing aspects of group working.

Standard naming schemes such as that provided by the joint ISO/CCITT international directory standard ISO IS 9594, CCITT X.500 [4] can also be usefully integrated with information domains. The directory is represented as a tree, with general denominators such as countries at the top, and progressing downward through organisations and departments to names of people and objects. The global (*distinguished*) name of an entry is given by the entire path through the tree to its vertex, and its local (*relative distinguished*) name is given by the lowest of the identifiers, thus providing a name for the object at the level of a group. Organisations, departments, and groups can all be mapped to IDs.

2.3 Information domains for defining groups

The final reason for including information domains in a model of group working is to allow the definition of limits for groups. This can be useful in a number of ways, for example:

- easy identification of members of the group becomes possible,

- restricting access to information is made easier,
- tools and resources needed for task performance may be made available to group members.

There are two important consequences of explicitly defining groups using information domains. The most obvious of these is the possibility of constraining the scope of operations such as searching, modifying, and reading information objects. This is becoming increasingly important as the size and complexity of computer systems and networks increases. For instance, it would not be very efficient to search all the documents in a large organisation for one particular document. The search could easily be constrained by specifying which group is likely to have possession of the document, and therefore, which information domain it can be found in.

The other consequence of defining groups using information domains is that it becomes possible to associate activity management specifications with explicit groups. Benford [3] has identified the need for these *local management policies*, which further specify the performance of tasks within specific local environments. These environments could easily be mapped onto groups as defined by information domains.

3 Existing treatments of information domains

This section describes and compares the work of two projects, MacAll II, and AMIGO MHS+, that included a model of information domains in their models of group working. It should be noted that both projects covered a much wider range of issues related to group communication, but only the work relevant to the concept of IDs will be described here. These particular projects have been briefly compared in this context and others in another paper [12] as part of a wider review of European work on CSCW.

3.1 AME workspaces

The Activity Model Environment (AME) [19] was developed by the MacAll II group at Nottingham University, and sponsored by Digital. The project aimed to develop a model of group working within an organisation based on the concept of an *activity*. The resulting model was based on the office structure proposed by a previous project, MacAll [20]. This structure included the concept of a *workspace*, which was defined as a conceptual work area associated with a person playing a particular *role* (so a person playing several roles would use a different workspace for each role). A role is defined as a set of responsibilities assumed by any person playing that role. Examples of roles are: secretary, lecturer, accounts clerk, and department head.

The definition of workspaces in the AME was refined and extended somewhat. One workspace is associated with one role, and all instances (role-person bindings) of that role use the workspace. It contains a storage facility for *iunits* (information units – e.g. documents, etc.), and a set of trays that contain currently active messages. Tools, and access to external resources, are also contained within the workspace. Finally, the workspace keeps a record of current instances of its associated role.

Conceptually, all the work that a person does takes place in the workspace associated with the particular role that s/he is playing at any one time. All the resources that are needed for successful performance of a role are contained within its associated workspace, and accessible to all people playing the role. People playing other roles may see the contents of the workspace, but they may not access them.

3.2 AMIGO MHS+ environments

The AMIGO MHS+ project [18] involved several European institutions, and was funded by the Cost-11ter programme. It concentrated on describing group communication in terms of the shared access of information. These descriptions were developed to three different levels: an abstract datamodel, an underlying architecture to support this, and models for the extension of existing services. The first of these is most relevant to this paper, as it is the datamodel that outlines the concept of *environments* in the most detail.

The purpose of the AMIGO MHS+ environment is twofold; to define the scope of operations, and to allow the specification of resources associated with a group. An environment is initially defined by the creation of an entry in the *directory*. This entry specifies groups of communication entities, important information objects, message stores, and other resources that are associated with the environment. The directory entry may also include a textual description of the environment, and rules controlling access to information within the environment.

Groups of roles are formed by the creation of environments, which allows the sharing of information between all the people involved in an activity. In addition, boundaries of environments can overlap, allowing information sharing between groups.

3.3 Comparison and discussion

The AME workspaces are useful for making information available to role players (and localising information in a distributed system), thus helping them perform their tasks within activities. They also make possible a means of addressing messages without knowing the specific address for a role player; knowing the person's name and the role they are playing will be enough. However, the use of the information domain concept is unnecessarily limited by only assigning workspaces to groups of similar roles. No provision is made for including objects in more than one workspace, and workspaces may not overlap.

AMIGO MHS+ environments are much more flexible, in that any group of different roles may be included in an environment. Also, environments may overlap, and may even be completely nested within another environment, which is an important advantage. Like workspaces, they can also be used for addressing schemes, and due to the use made of directories, they may also be used to help provide a naming mechanism for objects and other entities. However, environments are only used for grouping entities, and not for any other purpose. For instance, access across boundaries of environments is not restricted.

There are also problems common to both projects. One of these is that the relationship with activities is not fully explored by either model of information domains. Workspaces are not used to provide a means of defining groups of roles involved in particular activities, which would be more useful from the information sharing point of view. On the other hand, envi-

ronments are created to define these groups of roles, but no knowledge about the procedures roles are expected to follow is included in the AMIGO MHS+ datamodel.

Neither project carried out a comprehensive study of the reasons for including information domains in a model of group communication, and the uses they may be put to. The result is a lack of integration with the procedural (and, to some extent, structural) aspects of group working and organisational structure. It seems clear that such a study needs to be attempted, in conjunction with the development of models of procedural and structural aspects of group working, in order to produce a complete framework for describing CSCW.

4 Issues involved in modelling information domains

This section will identify and discuss issues to be addressed when modelling information domains. The purpose for their introduction into models of group working is the starting point for this discussion, as the reasons for their use will strongly affect how they are modelled. The implications of these reasons will be outlined; then a series of issues will be identified and discussed.

Information domains for information sharing. The main implication here is that information domains should be structured in such a way that groups of people who *need* to share information, and groups who *want* to share information, are able to set up the appropriate IDs, or have them set up, and are able to maintain them. A consequence of the fact that there are many different types of groups (e.g. departments in an organisation, study groups in a university) is that there may have to be many different types of IDs.

Information domains for object naming. Here, the implication is that if a naming policy is adopted and used in conjunction with information domains; the way IDs are designed should take this into account. For instance, if there are different types of ID, it needs to be specified how each type affects the naming of objects. Obviously, this only applies where a naming scheme that is more complex than simple local or global naming of objects is used.

Information domains for defining groups. The internal structure of IDs needs to be specified, so that there is a definition of the possible constituents of a group, and so that constraints on these may be specified if necessary. Also, properties of boundaries of IDs need to be specified. For example, the types of access that are possible across boundaries may be defined, and if overlapping boundaries or nested IDs are allowed, the consequences of these properties should be fully investigated.

Following on from this analysis, there are two general considerations that need to be addressed when modelling IDs:

- Creation and maintenance of IDs should be facilitated as far as possible by the model.
- The way IDs are modelled should be conceptually acceptable to the end-user.

There are also two specific modelling issues where the reasons for including IDs have an effect:

- Specification of the internal structure of IDs.
- Definition of the properties of ID boundaries.

The remainder of this section will discuss each of these areas in turn, identify choices that must be made when developing a model of IDs, and propose some solutions to problems that may arise.

4.1 General considerations

4.1.1 Creation and maintenance of information domains

The main purpose of information domains is to facilitate as far as possible the sharing of information. Thus, it must be easy to create and maintain IDs, and this should not interfere unduly with performance of activities in general. Any future development of IDs needs to consider this aspect when specifying their structure. For example, guidelines on the possible contents of an information domain, and their specification, could be included (discussed further in section 4.2.1).

The process of creating an information domain is likely to follow an identifiable pattern, for example:

1. Create an ID and give it a name
2. Specify the contents of the ID and its maintainer
3. Notify the contents that they belong to the ID
4. Notify other entities of the ID's existence
5. Specify any management policies associated with the ID

This pattern will depend on the particular model of IDs that is developed, but once that has been specified, the tasks associated with creation will become apparent. The mechanisms of the procedural part of the model of group communication could then be used to define the *activity* of ID creation in order to provide additional support for the end-user.

When considering the maintenance of IDs, the degree that the ID and its contents are permitted to change will be decided upon. There are several possibilities here, ranging from no permitted change, through change of contents only, to change of both contents and properties of the ID. It is likely that any group will need to change its contents; to add a new member, or include a new information object. It is less likely that the basic properties of the group will change, but it is still possible (e.g. the work of a group suddenly becoming secret), and so that may also be considered.

The other consideration concerning ID maintenance is the degree of automation of this maintenance that is possible. The solution at the moment would be to define a role of *administrator* that has to be instantiated for every information domain, and provide a set of guidelines for the person playing the role to follow. However, there is a growing interest in the specification of activity management policies (e.g. [3]), and it is possible that this work could be applied

to the automation and support of ID maintenance. Also, the maintenance of an ID could be considered to be an activity in itself, and therefore the procedural aspects of the model could again be used.

4.1.2 Acceptance of the information domain concept

For efficient information sharing, end-users need to understand and accept the concept of an information domain. This can be achieved by defining IDs so that they reflect the formal and informal group structure that exists in real life. There are three implications of adopting such a structure for IDs:

- The structure of groups within an organisation is very dynamic; groups will often be formed, broken up, or changed in some way.
- Groups often overlap with, or exist entirely within, other groups.
- There are often different types of groups within an organisation.

All these implications should be considered when modelling information domains, in order to allow the end-users to assimilate the concept and use it effectively for information sharing.

The dynamic nature of groups indicates that the creation and maintenance of IDs (dealt with in the previous section) should be as flexible as possible, and should not interfere with normal group working. In addition, the increasingly distributed nature of computer installations will make it very difficult for a central entity to keep track of changes in group structure, although it would be possible. It would be more efficient to define IDs that are relatively autonomous and self-regulating.

The naming policy that is used could be affected by overlapping IDs, if it is sufficiently complex. If an information object exists in two overlapping groups with a different name in each, difficulties could arise. Another problem that needs to be addressed when considering overlapping IDs is that of management policies. If two groups with different policies and controls overlap, a conflict may arise. Further study is needed here, both into what happens in real life, and into possible ways of modelling these processes.

Organisations may contain many different types of groups that differ in many ways. For example, some groups may be more or less permanent (e.g. departments), whereas others may be formed solely for one short-term activity (e.g. entities involved in organising a trip); some groups may be secret, others may be open; and so on. A sensible categorisation of these groups is needed, from which it should be possible to see whether different types of ID need to be defined, or whether the various types of groups can be reflected by differing attributes of the same ID type.

4.2 Specific modelling issues

4.2.1 Internal structure of information domains

The general issues of creation and maintenance of IDs, and the users acceptance of the concept of IDs, must be considered when defining information domains. The process of creating an

information domain is likely to follow a particular pattern, as discussed earlier, and this may be partly identified from the contents and attributes of the ID. It therefore follows that these properties should be specified in a model of information domains, even if some are specified as optional. Examples of the types of properties that may be specified are:

- **Name** – an identifier by which the ID is referred to.
- **Description** – a textual description of the purpose of the ID.
- **People/Roles** – human entities included in the ID.
- **Information objects** – documents etc. included in the ID.
- **Storage information** – description of the organisation of information objects.
- **Resources** – tools and services available within the ID.
- **System agents** – automated entities that exist in the ID.
- **Management information** – this issue is discussed further in [3].

In addition to specifying which properties may be used, some guidelines on their use may be defined. For example, instead of just stating that a set of resources may be included; the format of the expression to include them may be specified as follows:

resource-component = <name> <description> <access-procedure>

These guidelines may be applied equally well to the processes of creation and maintenance of IDs. It would also be possible to specify constraints on the number or type of a particular entity that should be included in a type of ID. For example, it could be specified that a department has between 3 and 30 people, exactly one of which will play the role of department head.

The internal structure of IDs should reflect as far as possible the internal structure of real-life groups. That is, it will include all the entities that may be associated with a group in reality, and the guidelines and constraints on the specification of the contents of an environment should be sensible and easily understood. In particular, if different types of information domains are required and defined, their specified contents will reflect the IDs' purposes.

4.2.2 Properties of information domain boundaries

Creation, maintenance, and acceptance of IDs also need to be considered when defining the properties of ID boundaries. Any properties that may have to be refined at creation time, or during the lifetime of the ID, should be carefully specified, along with guidelines for their refinement. Also, the way ID boundaries are specified should allow groups to be formed in the same way that they are formed in reality.

There are two major properties that ID boundaries may possess that need to be considered:

- **Transparency** – controls the visibility of the contents from the outside.

- **Access** – controls the access of contents by entities outside.

There are four degrees of transparency and access that an ID boundary may possess. First, it may be completely opaque, so that no entities outside can see or access its contents. Second, it may be completely transparent, so that all its contents are visible to the outside but not accessible. Third, it may be transparent and all contents may be accessible. These three possibilities present no problem.

The fourth possibility is that some of the contents are visible to the outside, and others are invisible (some or all of the visible contents may be accessible). In this case, some mechanism for achieving this must be provided. One possibility is to make the ID responsible for taking requests to view or access contents, and referring to lists of visible and accessible contents when deciding which to display or allow access to. Whatever mechanism is chosen, it ought to be transparent to the end-user.

4.3 Summary

This section has examined the reasons for developing a model of IDs, and from these, four important issues have been identified and discussed. The overriding concern is to define information domains in such a way that they are useful, and facilitate, rather than interfere with, group working. In order to achieve this, the internal structure and boundary properties of information domains should be carefully specified, bearing in mind that the user must be able to accept the concept of an information domain in order to use it properly.

5 Conclusion

The previous sections have described the concept of *information domains*, outlining reasons for their incorporation into models of group working, describing past work on the subject, and identifying important issues that must be considered when modelling information domains. Although low-level support for information sharing has been investigated to some extent, most models of co-operative work developed so far do not take into account information sharing aspects of group working, but concentrate on procedural and structural aspects.

This paper has argued that there are three main reasons for including information sharing aspects in general models of group working. The first, and most important, of these is to facilitate and support efficient information sharing in a CSCW environment. Secondly, the integration of an object naming scheme that does not rely totally on global naming is provided for. Finally, the clear definition of working groups and the interaction of these groups becomes possible. It is suggested that the need for the inclusion of information sharing aspects in group communication models could be met by incorporating the notion of an information domain.

It has become apparent from looking at existing treatments of IDs that there are many issues to be considered and researched before a coherent model of information domains is developed. It seems likely that some of these issues may be resolved by investigating real-life group interactions, and modelling how information is shared and groups are formed in reality. In addition, extra support for the end-user should be provided so that the mechanisms for creating, maintaining, and using information domains do not inhibit the efficient performance of organisational processes.

The integration of procedural, structural, and information sharing aspects of modelling cooperative work should be relatively easy. The use of information domains in itself helps to integrate the latter two aspects, by providing a basic structure on which to build more detailed structural distinctions. Procedural definitions within the model will be used to define ID creation and maintenance processes, and IDs, once they have been created, can be used to contain all the entities involved in a particular activity.

There is much scope for further study of the information domain concept, and of information sharing in general. The three reasons for including IDs outlined in section 2 by no means exhaust the possible uses of information domains; further research would doubtless reveal more possibilities. There is a need for the investigation of the interaction between groups, and the different types of groups, that exist in real life. In addition to these problems, the issues highlighted in section 4 also need further consideration.

The whole concept of an information domain is relatively new, and there are many problems to be resolved. It seems clear that models of procedural and structural aspects of group communication are not enough to facilitate true computer supported cooperative work. It has been recognised that, at the most basic level, group working involves groups of people performing tasks, according to certain patterns. However, in order to do this, these people need information, and they need to be able to share that information.

References

- [1] L. Aiello, D. Nordi, and M. Panti. Modelling the office structure: A first step towards the office expert system. In *Proceedings of the second ACM SIGOA conference*, pages 25-32, 1984.
- [2] E. Auramaki, E. Lehtinen, and K. Lyytinen. A speech-act-based modelling approach. *ACM Transactions on Office Information Systems*, 6(2):126-152, 1988.
- [3] S.D. Benford. Requirements of activity management. In *Proceedings of the 1st European Conference on CSCW*, 1989.
- [4] S.D. Benford. Components of OSI: the OSI directory service. In *Connexions: the interoperability report*, pages 2-9, Summer 1989.
- [5] J. Bowers, J. Churcher, and T. Roberts. Structuring computer-mediated communication in COSMOS. In R. Speth (ed), *EUTECO '88 - Research into networks and distributed applications*, pages 195-210, 1988.
- [6] G. Bracchi and B. Pernici. The design requirements of office systems. *ACM Transactions on Office Information Systems*, 2(2):151-170, 1984.
- [7] K. Crowston, T.W. Malone, and F. Lin. Cognitive science and organisational design: A case study of computer conferencing. In *Proceedings of CSCW '86*, pages 43-61, 1986.
- [8] F. De Cindio, G. De Michelis, and C. Simone. Groups in a language/action perspective. *Cosmos Information Exchange Network*, pages 5-18, 1988.

- [9] F. Flores, M. Graves, B. Hartfield, and T. Winograd. Computer systems and the design of organisational interaction. *ACM Transactions on Office Information Systems*, 6(2):153-172, 1988.
- [10] G. Goodman and M. Abel. Collaboration research in SCL. In *Proceedings of CSCW '86*, pages 246-251, 1986.
- [11] I. Greif and S. Sarin. Data sharing in group work. In *Proceedings of CSCW '86*, pages 175-183, 1986.
- [12] P. Hennessy, S. Benford, and J. Bowers. Modelling group communication structures: an analysis of four European projects. In *Singapore International Conference on Networks*, pages 56-61, 1989.
- [13] J. Hogg. Intelligent message systems. In D. Tschritzis (ed), *Office automation*, pages 113-134. Springer-Verlag, 1985.
- [14] T. Kreifelts and P. Seuffert. Addressing in an office procedure system. In *Proceedings of the IFIP 6.5 international working conference on message handling systems*, 1987.
- [15] T.W. Malone, K.R. Grant, K-Y. Lai, R. Rao, and D. Rosenblitt. Semi-structured messages are surprisingly useful for computer-supported coordination. In *Proceedings of CSCW '86*, pages 102-114, 1986.
- [16] U. Pankoke-Babatz (ed), T. Danielsen, P.A. Pays, W. Prinz, and R. Speth. *Computer based group communication - The AMIGO activity model*. Ellis Horwood, 1989.
- [17] W. Prinz and R. Speth. Group communication and related aspects in office automation. In *Proceedings of the IFIP 6.5 international working conference on message handling systems*, 1987.
- [18] H. Smith, J. Onions, and S. Benford (eds). *Distributed group communication - the AMIGO information model*. Ellis Horwood, 1989.
- [19] H.T. Smith, P.A. Hennessy, and G.A. Lunt. The Activity Model Environment: an object-oriented framework for describing organisational communication. In *Proceedings of the 1st European Conference on CSCW*, 1989.
- [20] H.T. Smith, G.A. Lunt, D. Young, and J. Lawrence. MacAll: An interface design project. Final report on the MacAll I project, available from CRG, Computer Science Department, Nottingham University, 1987.
- [21] K. Sollins and D. Clark. Distributed name management. In *Proceedings of the IFIP 6.5 international working conference on message handling systems*, 1987.
- [22] D. Tschritzis and S.J. Gibbs. Etiquette specification in message systems. In D. Tschritzis (ed), *Office automation*, pages 93-112. Springer-Verlag, 1985.