An Activity Based Approach to Context-Aware Computing

Tye Lawrence Rattenbury
John F. Canny

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2008-163
http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-163.html

December 17, 2008

An Activity Based Approach to Context-Aware Computing

by

Tye Lawrence Rattenbury

B.S. (University of Colorado at Boulder) 2001

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy
in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Canny, Chair
Professor Dan Klein
Professor Alice Agogino
Professor Anind Dey

Fall 2008

The dissertation of Tye Lawrence Rattenbury is approved:

Chair ____________________________ Date __________

____________________________ Date __________

____________________________ Date __________

____________________________ Date __________

University of California, Berkeley

Fall 2008

An Activity Based Approach to Context-Aware Computing

by

Tye Lawrence Rattenbury

Abstract

An Activity Based Approach to Context-Aware Computing

by

Tye Lawrence Rattenbury

Doctor of Philosophy in Computer Science

University of California, Berkeley

Professor John Canny, Chair

The notions of activity and context serve as unifying threads in this dissertation. As socio-cultural phenomena, both activity and context refer to complex and dynamic aspects of human existence. Prior context-aware computing (CAC) applications have failed to address this complexity and dynamism. We overcome these failures by defining activity and context relationally – as co-dependent and mutually producing phenomena. Our definitions explicitly highlight the evolving impact that activity has on context and that context has on activity. We then translate our definitions into a novel computational model. This model differs from prior work in its ability to automatically detect and evolve representations of activities and contexts. Finally, we demonstrate the utility of our computational model by evaluating a prototype CAC application. This application is driven by, and enables interaction with, automatically generated representations of the user's activities and contexts.

Professor John Canny
Dissertation Committee Chair

This dissertation is dedicated to Kirby.

Contents

Chapter 1

Introduction

The central threads running through this dissertation are the notions of activity and context. By analyzing and applying these notions, we contribute to the research area of Context-Aware Computing, and more generally to the field of Human Computer Interaction (HCI). More concretely, we translate the notions of activity and context into a novel computational model. This model drives a software application that both verifies our ability to model activity and context, and provides practical utility to its users.

1.1 Motivation: Why Activity and Context?

Over the last twenty-five years, HCI research has increasingly shifted its focus away from the individual computing paradigm – one user, one device – towards richer and more complicated environments and scenarios – e.g., collaborative groupwork, mobile computing, and augmented reality. Whereas the individual computing paradigm could be dissected into a series of simple steps, like "clicking on a button" (*John and Kieras* [1996]), which designers could optimize their systems around, richer and more complicated settings make such dissections and optimizations untenable. Instead, researchers in Computer Supported Cooperative Work (CSCW), Ubiquitous Computing (Ubicomp), and Context-Aware Computing (CAC) have employed more holistic conceptions of human activity to assess and design for the usability and utility of information technology (e.g., *Begole et al.* [2002]; *Christensen and Bardram* [2002]; *Convertino et al.* [2006]; *Dourish and Bellotti* [1992]; *Dourish and Bly* [1992]; *Greenberg* [1996]; *Gross et al.* [2003]; *Huang et al.* [2002]; *Huang and Mynatt* [2002]; *Isaacs et al.* [1996]; *Narendra and Moran* [2006]; *Pedersen and Sokoler* [1997]; *Sawhney et al.* [2001]; *Voida et al.* [2002]).

Activity does not occur in isolation, however. To fully understand an activity, one often requires an understanding of the context 'around' the activity. From a sociological and psychological perspective, context plays an important role in everything from effective learning situations within a person's lifetime (*Bateson* [1972]), to a more holistic sense of human development (*Leontiev* [1977, 1978, 1981]), to the systematic creation of sociological entities (*Giddens* [1984]; *Levine* [1971]). From a discourse or conversation analysis perspective, context is understood to be the necessary grounding that enables effective communication (*Atkinson and Heritage* [1984]; *Garfinkel and Sacks* [1970]; *Goodwin and Duranti* [1992]). And, as with activity, technology-centric research has often treated context as information that, if made explicit within technology, could improve human-computer interaction (*Ackerman et al.* [2001]; *Dey* [2000]; *Dey et al.* [2001]; *Dey* [2001]; *Dourish* [2001]; *Moran and Dourish* [2001]; *Winograd* [2001]).

A cursory pass through the literature cited above readily demonstrates that neither activity nor context is well-defined. Instead, these terms are applied to a variety of socio-cultural phenomena, at various temporal and spatial scales, and often without acknowledgement of prior applications – and hence without discussion of how to assess the relationship between applications. To enable a more precise discussion of activity and context, that synthesizes a variety of prior applications of these terms, we work with the following definitions:

Context is a set of markers (perceivable differences in the world) that enables a person's assessment of which activities, and which interpretations of these activities, are appropriate.

Activity is intentional behavior that produces both context markers[1] and the socially shared, i.e., cultural, mapping between context markers and appropriate activities.

Notice that these definitions are fundamentally subjective – they are determined by a person's capacity to associate markers with behavior. Moreover, these definitions are co-dependent. The emergent and dynamic aspects of activity give rise to the same qualities in context – i.e., which markers are perceived, and how they impact the assessment of various activities, is an emergent and dynamic phenomenon.

In unifying previous conceptions of context, our notion of context includes the following characteristics[2]: First, context is an emergent property of human activity (*Akman* [2000]; *Dourish* [2001, 2004]; *Goffman* [1974]). In other words, human activity produces context, but not in a procedural or mechanical fashion. Rather, context is produced by the complex interactions of human behaviors at different spatial and temporal scales, which makes it partially unpredictable and ineffable. Second, in addition to intra-personal behavioral interactions, context is determined by inter-personal interaction – i.e., it is social (*Garfinkel and Sacks* [1970]; *Giddens* [1984]; *Greenberg* [2001]). Finally, as behavioral interactions change with time, so does context. But while context is dynamic and evolving, this evolution exhibits stability (*Bateson*

[1] Note, however, that many context markers are not produced by human activity, e.g. sunrises and ocean tides.

[2] There are additional aspects to context which we ignore for the purposes of this dissertation. For example, the fact that context is ontologically different from the elements that compose it (*Bateson* [1972]).

[1972]; *Garfinkel and Sacks* [1970]; *Giddens* [1984]; *Leontiev* [1981]; *Svanaes* [2001]). In other words, within certain spatial and temporal regions, the stability of context permits a fairly decisive labeling of markers as either part of that context or not.

Likewise, our notion of activity draws heavily from previous conceptions of the term. Most notably, we draw on the notion of activity in Activity Theory (AT – *Leontiev* [1981]; *Nardi* [1996]). AT essentially posits that behavior is mediated by both material and social forces (*Wertsch* [1985a, 1991]). This mediation impacts the evolution and stability of behavior by creating a visible trace of its effects. AT also highlights the impact of environmental (i.e., contextual) factors on behavior. A more recent labeling of the contextual impact on behavior is "situated action" (*Suchman et al.* [1999]). Finally, AT explicitly treats human behavior at multiple granularities: from low-level, short duration operations to high-level, long-term, need-based activities (*Wertsch* [1979]; *Leontiev* [1977, 1978]).

While activity and context are clearly related by their relative importance to human existence, it is important to distinguish the nature of their complementary roles. We understand activity as the notion that captures the answers to three fundamental questions: (1) what are people doing? (2) why are they doing it? and, (3) how are they doing it? Alternatively, context captures the relationship between past activities and new and emerging activities. The question that context should answer is: what activities are appropriate for me (or others) to do right now? These descriptions of the roles of activity and context demonstrate their inseparability. They are co-dependent entities, playing out a dual relationship of mutual constraint and creation.

The idea that activity and context are connected is not new. Many researchers in the field of HCI have taken a similar position (*Abowd et al.* [2002]; *Ackerman* [2000]; *Akman* [2000]; *Crowley et al.* [2002]; *Dourish et al.* [1993]; *Halverson* [2002]; *Hollan et al.* [2000]; *Kuutti and Bannon* [1993]; *Winograd and Flores* [1986]). Likewise, other areas of computer science (*Bianchi-Berthouze et al.* [1999]; *Torre* [2001]) and other academic fields (*Allwood* [2000]; *Dewey* [2002]; *Hutchins* [1995]; *Hutchins and Klausen* [1996]; *Latour* [1995, 1999]; *Law* [1986b]; *Maturana*; *Star* [1997]; *Strauss* [1995]; *Suchman* [1987]; *Suchman et al.* [1999]) have recognized the benefit of understanding the co-dependence between human activity and context. Contemporary research on AT has also addressed this topic (*Blackler* [1993]; *Blackler et al.* [1999, 2000]; *Boer et al.* [2002]; *Chaiklin and Lave* [1993]; *Engeström and Blackler* [2005]; *Engeström* [2000, 1987]; *Engeström et al.* [1999b]; *Engeström* [1999]; *Halverson* [2002]; *Leontiev* [1977, 1978, 1981]; *Nardi* [1996]; *Russell* [1997]; *Spinuzzi* [2000, 2001, 2002, 2003]; *Wells* [2002]; *Wertsch* [1998]).

In this dissertation, we are concerned with the utilization of activity and context in technology; specifically, in software applications. It is therefore imperative that we ask how these notions have been used in prior work. Specifically, we are concerned with technology-centric research that explicitly includes the notions of activity and/or context. As we have argued above, the deep connection between these notions makes it difficult to treat them in isolation. However, we can roughly separate related work according to whether activity or context is treated as primary. On activity, there has been significant work in the areas of User Modeling and Artificial Intelligence (e.g., *Fischer* [2001]; *Huỳnh et al.* [2008]; *Kobsa* [2001]; *Webb et al.* [2001]; *Zukerman and Litman* [2001]; *Zukerman and Albrecht* [2001]) and in cognitive science approaches to HCI (e.g., *Card et al.* [1983]; *John and Kieras* [1996]). Most of this work has been primarily concerned with modeling some aspect of human behavior in a computational framework. The lacking focus on context in this work has prompted the general criticism that these models are idealized and over-constrained; they lack enough realism to be applicable in real end-user applications. It follows that an important question to address in this dissertation is: How can we capture and model human activity in a realistic way?

Complementing the work on activity is work focused on context. Again, most of this technology-centric research treats context as the set of information that could improve human-computer interaction. Such work is often associated with the area of CAC (see *Dey* [2000] for a comprehensive survey). What is important to highlight in prior CAC work is the approach that is taken for designing these systems. In general, the designer of these systems pre-specifies which markers (i.e., types of information) constitute context (*Dey* [2000]). The designer then specifies how to capture and utilize these context markers within the scope of the application they are building. However, the assumption that a designer can pre-specify context markers is likely flawed (*Dourish* [2001, 2004]). There are even arguments that capturing context is an inherently impossible endeavor (*Bellotti and Edwards* [2001]; *Grudin* [2001]). While we agree that pre-specifying which markers constitute context is a highly limited approach to CAC, we do not agree that attempts to capture context are doomed to fail. While it is almost certainly true that not every context marker can be captured, it is also almost certainly true that people having a meaningful interaction are not simultaneously aware of the exact same set of context markers. Rather, the ability to have a meaningful interaction hinges on sharing *enough* context markers. Hence, we should recognize that the success of CAC applications is not binary, but lives along a continuum where applications can improve provided they can better identify and monitor relevant context markers. This leads to another question that we seek to address in this dissertation: how can we effectively identify and track relevant context markers?

1.2 The Central Question of this Dissertation

In motivating this dissertation we have arrived at two questions that we would like to address: (1) how can we capture and model human activity in a more realistic way? and, (2) how can we effectively identify and track relevant context markers? In this section we translate these questions into a single, more precise research question that defines the primary contributions of this work.

To begin, we need to expose the implicit assumption in both of these questions. The assumption is that realistic activity modeling and more effective tracking of context markers is a worthwhile pursuit. While the ubiquity of activity and context discussed above certainly motivates additional studies of these notions, our motivation is more pragmatic. The ultimate objective, coming from a Computer Science, and more generally an engineering, perspective is to facilitate the creation of more beneficial technology. Here, "beneficial" is defined practically as utility to the end-users of the technology. Drawing on the explication of utility within the area of CAC (*Dey* [2000]), we can distinguish two categories: (1) improving existing technology, and (2) enabling the creation of new technology. As we show in this dissertation, our application contains elements of both of these categories.

With an understanding of why these questions deserve attention, we can begin to answer them. Recall that both of these questions came directly from criticisms of existing work. In the activity modeling domain, the lack of realism essentially removed any potential utility for end-users. In the CAC domain, the existing practice of specifying context markers a-priori limits the potential utility for end-users. Common to both sets of prior work is an underlying positivism about the structure of activity and the markers that constitute context (*Dourish* [2004]). But, as argued above, activity and context are emergent and dynamic social constructs. However, the stability of their evolution does not enable one to patently treat them as fixed in structure (and hence definable a-priori during the design of technology). Logically, to increase the realism of activity models and to more effectively track context markers, one needs to support dynamic and emergent structures.

These arguments lead to the central question answered by this dissertation: **Can we demonstrate the utility of a context-aware application that automatically tracks and models a user's activity and context?**

In the remainder of this dissertation, we describe how we *can* demonstrate the utility of an application that automatically tracks and models context and activity. Our application is a personal work awareness display. The content of the display indicates the various stable contexts created by the user as they carry out their various activities through and on the computer. The utility of this display is manifold. First, it demonstrates a clear advantage over related work awareness systems that often require users to manually create the content of the display. Second, it supports users in a real-time, context-aware fashion by dynamically reconfiguring itself as the user carries out various activities. And, third, by presenting users with a reflection of their own work practice, the display explicitly supports the reflexive nature of human activity.

Notice that while this application clearly demonstrates the utility of building applications that automatically model activity and track context, we are not claiming to have exhausted the space of such applications. Nor are we claiming to have identified the maximum or optimal amount of utility of the specific application we built. Rather, the application is a proof-of-concept. It was developed enough to be usable and useful to end-users. However, we openly acknowledge that it could be improved by continued iterative design.

1.3 Contributions of this Dissertation

There are three primary contributions made in this dissertation. As the first contribution, we provide an integrating survey of existing theories of activity and context, both within and outside of Computer Science. We have two goals in performing such a survey: First, we would like to exploit the many insights that have been produced with respect to these notions in the many fields concerned with human behavior. Second, as we are ultimately motivated by providing utility to the end-users of technology, we need to be sensitive to factors that exists "outside" of activity and context, but that might influence the utility of their application. By surveying various academic disciplines for their notions of activity and context, we are made aware of the various connections that activity and context have to other aspects of human behavior.

The second contribution of this dissertation is the translation of our notions of activity and context into a computational model. As we have argued above, this model does not define a-priori an explicit structure of a person's activities or identify a fixed set of relevant context markers. Rather, our computational model automatically builds and evolves a representation of a person's activities and which context markers are relevant to these activities.

The third contribution of this dissertation is the validation of an automatic approach to CAC. Again, we want to emphasize that this dissertation is not intended to provide the final statement on any outstanding open problems, or to end any lines of research. Rather, the validation provided in this dissertation should be understood as a clear step in a new direction of activity and context based technology design. Our ultimate objective is to inspire more research attention on technology that can effectively handle and respond to the nature of human activity and its relationship to context.

Chapter 2

Existing Theories and Models of Activity and Context

In this chapter, we survey existing theories and models of activity and context. Our goals in this survey are to: (1) to motivate and justify our computational translation of activity and context by more rigorously exploring their definition; and, (2) to identify aspects of human behavior that impact activity and context, but are not constituent elements of these notions.

2.1 Our Current Definitions of Activity and Context

In the introduction of this dissertation we defined activity as the intentional behavior that produces both context markers and the socially shared mapping between context markers and 'appropriate' activities. Superficially, i.e. using common sense interpretations of the terms, this definition appears to explicate our notion of activity. However, by describing the primary terms of this definition in greater depth, we can further motivate and justify the translation of our understanding of activity into a computational model. To guide this process, we survey existing literature related to the notion of activity and answer the following questions:

- What do we mean by "intentional" behavior?
- What do we mean by "production", regarding markers as well as mappings between markers and behavior?
- How does individual behavior lead to socially shared ideas/artifacts?
- What do we mean by "appropriate" activities?

Similarly, we re-visit our definition of context. Recall that we defined context as the set of perceivable markers that enables a person to act appropriately and to interpret the behavior of other people. As with activity, we hope to further motivate and justify our computational translation of this definition of context by answering more in-depth questions. Specifically:

- What do we mean by "markers", or "perceivable differences in the world"?
- What do we mean by "appropriate" activities? (note: this is the same question as posed for activity)
- What do we mean by "interpret" behavior?

In pursuing answers to these questions, we survey a fairly comprehensive set of theories and models of human behavior. Of course, our treatment of each of these theories and models is biased by our focus on activity and context. Accordingly, we do not provide a complete overview of each theory and model. Rather, we focus our discussion on how each theory or model of behavior explains, situates, and employs the notions of activity and context.

2.2 Existing Theories of Activity and Context

Again, we draw primarily from Activity Theory (AT) (*Bertelsen and Bodker* [2003]; *Engeström* [2000, 1987]; *Engeström et al.* [1999b]; *Engeström* [1999]; *Halverson* [2002]; *Leontiev* [1977, 1978, 1981]; *Nardi* [1996]; *Russell* [1997]; *Spinuzzi* [2000, 2001, 2002, 2003]; *Wells* [2002]). To temper this bias, we also discuss various related, and sometimes conflicting, theories. These include:

- Vygotsky's (with elaborations by Wertsch and others) conceptual development of tool-mediated action and the importance of genetic (developmental) perspectives on tool-use and skill acquisition (*Engeström et al.* [1999b]; *Luria* [1976]; *Scribner* [1985]; *Vygotsky* [1978, 1985]; *Wertsch* [1985b]; *Zinchenko* [1985]);

- Engeström's ideas on contradiction, expanding object transformations, multiple interacting activities, and his notion of social roles (*Engeström* [2000]; *Engeström et al.* [1999b]; *Engeström* [1999]);

- Actor-Network Theory (as described by Latour and Law) for its discussion of networks structured for power (both rhetorical and production), the roles embodied in these networks, and the notions of punctualization and black boxes (*Latour* [1995, 1999]; *Law* [1986a, b, 1992]; *Ryder*; *Stalder* [1997]);

- Distributed Cognition (as described by Hutchins) for its discussion of inter-person cognitive units (*Halverson* [2002]; *Hollan et al.* [2000]; *Hutchins* [1995]; *Hutchins and Klausen* [1996]);

- literary genres, speech genres, genre ecologies, and genre tracing (*Freedman and Medway* [1994]; *Russell* [1997]; *Bakhtin* [1979]; *Wertsch* [1985a, 1991, 1998]; *Spinuzzi* [2003]);

- Strauss (and others) for the theory of grounded action (*Addelson* [1995]; *Soeffner* [1995]; *Star* [1995, 1997]; *Strauss* [1995]);
- the frameworks of conversation and discourse analysis (*Atkinson and Heritage* [1984]; *Garfinkel and Sacks* [1970]);
- Bateson and Goffman's notion of frames and structures of interaction (*Bateson* [1972]; *Goffman* [1959, 1967, 1974]);
- Giddens's theory of structuration (*Giddens* [1984]);
- Feld's theory of social foci (*Feld* [1981]; *Wasserman and Faust* [1994]);
- Pragmatism (as developed by Dewey and James) for the notions of habitual behavior, inquiry, transaction, and pragmatic truth (*Dewey* [2002]; *Garrison* [2001]; *James* [1975]; *Miettinen* [2001]);
- Bateson's ideas on cybernetics (*Bateson* [1972]); and,
- theories on literacy and education (*Bruner* [1966]; *Leontiev* [1981]; *Luria* [1976]; *Scribner* [1985]; *Vygotsky* [1967, 1978, 1985]; *Wertsch* [1985a, b]).

Activity Theory

AT originated in post-revolutionary Russia during an unprecedented consolidation of the major fields of social science (*Wertsch* [1979]). AT is deeply connected to the wider socio-historical school of thinking which has its roots in Marxist psychology. Of the many influences that Marx had on the conception of AT, the most important is probably the emphasis on material and social mediation of behavior (*Leontiev* [1981]; *Wertsch* [1985a, 1991]). This mediation impacts the evolution and stability of behavior by creating a visible trace of its effects. The creation of this visible trace is often connected to Marx's conception of production and productive behavior (*Engeström*

[1999]; *Leontiev* [1981]). AT explains the influence of these material and social traces of past activities, which are major components of context, on present and future activity by emphasizing the situatedness of behavior (see Suchman's development of "situated action" for a more contemporary perspective, *Suchman et al.* [1999]). The situatedness of behavior is addressed by AT in multiple ways, depending on the granularity, or level, of behavior. These levels are organized hierarchically, from low-level, short duration operations to high-level, long-term, need-based activities (*Wertsch* [1979]; *Leontiev* [1977, 1978]). We begin our discussion at the lowest level.

The shortest duration behavior that AT acknowledges is the *operation*. Operations are the habitual, sub-conscious routines utilized by people to act in the world. As such, they are cued by conditions in a person's environment and are generally performed without significant conscious effort or use of conscious mental faculties. AT treats operations as triads consisting of a person or persons (referred to as the *subject(s)*), a set of mediating artifacts (referred to as *tools*), and the condition(s) in the world that cues the operation's performance. It is through the dependence on conditions that operations derive their situatedness. These entities are maximally connected. Schematically, this triad is usually represented as shown in Figure 2.1. Operations have been well studied in other academic fields (e.g. *Dewey* [2002]). In HCI and Computer Science in general, the operations from GOMS (*Card et al.* [1983]; *John and Kieras* [1996]) and the notions of low-level actions in traditional artificial intelligence (*Suchman* [1987]; *Winograd and Flores* [1986]) are nearly identical to AT operations.

An important distinction about the tools involved in an operation is that they are often assumed to encode the operation performance in their design (*Leontiev* [1977, 1978, 1981]). For example, the operation of hammering a nail is designed into the hammer itself by way of its weighting, its heavy and flat head, and its lighter and often ergonomically shaped handle. A further distinction concerning the tools

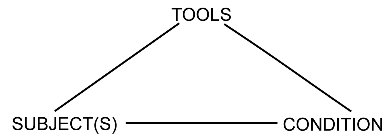

Figure 2.1. Schematic of an AT operation. Operations are the lowest level, finest granularity, of human behavior in AT, and can be commonly understood as subconscious habits.

involved in an operation is that they can control or determine the operation along a spectrum of magnitude. At one end of this spectrum the tool is essentially inert, permitting the person to have full determination over its use. An example of this type of tool is a semantically blank symbol, like x or c, used by a mathematician. At the other end of this spectrum are tools that completely specify their use – essentially appropriating the person as a power source. An example would be a doorknob on a door through which a person is trying to pass. Actor Network Theory also recognizes this varied relationship between people and tools (*Latour* [1986, 1995, 1999]; *Law* [1986a, b, 1992]; *Ryder*; *Stalder* [1997]). However, it does not restrict its analysis to the operation level when discussing this relationship.

The next level of behavior in AT is *action*. Actions, unlike operations, are generally conscious. In place of the conditions in an operation, actions have a conscious *goal*. The goal-formation process, however, is one of the least understood aspects of AT (*Leontiev* [1981]). However, it is widely acknowledged that goal formation is a situated process, made possible by feedback from one's material and social environment (*Leontiev* [1981]; *Wertsch* [1998]). Like operations, actions are maximally connected triads. Figure 2.2 illustrates this structure of actions. The re-use of subject(s) and tools is intentional. To understand how these entities are defined in actions, it is important to understand that actions consist of operations (*Leontiev* [1977, 1978, 1981]). Another way to say this is that in carrying out an action (i.e., achieving its goal), the subject(s) perform(s) a series of operations. Thus, the subject(s) in an action are just the union set of subject(s) in the constituent operations – likewise for the

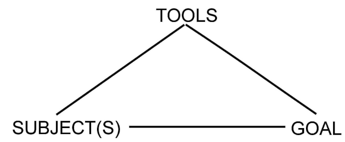

Figure 2.2. Schematic of an AT action composed of subject(s), tools, and a goal. Action is the middle level of behavior in AT.

tools of the action. Importantly, a single operation can be used, or rather re-used, in many different actions. And, sequences of operations for multiple actions are often interleaved in time. We refer to this phenomenon as multiplexing.

The action level of behavior is well studied in other paradigms. In both HCI and traditional Artificial Intelligence, the goal-driven perspective on human behavior has dominated both in theory and in practice – although in these settings, actions are often referred to as tasks (c.f., *Card et al.* [1983]; *John and Kieras* [1996]; *Suchman* [1987]; *Winograd and Flores* [1986]). Outside of these areas, which are under the Computer Science umbrella, goal-driven behavior has been an important paradigm for Cognitive Science, Psychology, Sociology, and Economics (*Simon* [1996]). Of course this list is not exhaustive, and blends between these fields have also employed the goal-driven perspective of human behavior.

So far, AT has not provided a discrepant or novel view of human behavior. The operation and action levels of AT correspond to well known, and well accepted, models. However, AT describes a third level of human behavior. It is at this level that AT differs most from other theories of human behavior.

The highest level of behavior in AT is, of course, *activity*. Activities, like operations and actions, are maximally connected triads which include subject(s) and tools. However, instead of the condition(s) of an operation or the goal of an action, activities have an object-motive pair. Figure 2.3 illustrates this structure. The process of adopting an object-motive pair, of incorporating it into one's behavior, is a strongly mediated by one's socio-cultural environment (*Bardram* [1998]; *Bertelsen*

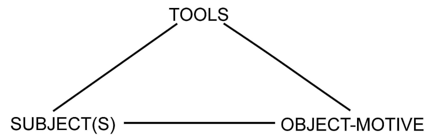

Figure 2.3. Schematic of an AT activity composed of subject(s), tools, and an object-motive pair. Activities are the highest level of behavior in AT.

and Bodker [2003]; *El'konin* [1972]; *Engeström and Blackler* [2005]; *Gal'perin* [1967]; *Leontiev* [1981]; *Wenger* [2000]). This adoption process is known as *internalization* in AT. The counterpart to internalization is *externalization*. Externalization is the process by which object-motive pairs are shared. This sharing process is mediated by language and material artifacts (e.g., in the tools that one teaches or leaves for others to use) (*Leontiev* [1981]).

Additionally, just as actions are composed of operations, activities are composed of actions. It follows that the subjects and tools involved in an activity are the union sets of subjects and tools involved in the constituent actions. An important characteristic of activities is that people are generally involved in many of them at any given time in their life. Even within specific settings, like the workplace, people have many ongoing, overlapping activities (*Chaiklin and Lave* [1993]; *Christensen and Bardram* [2002]; *Engeström* [1999]; *González and Mark* [2004]; *Mark et al.* [2005]). Practically, this overlapping plays out in actions that simultaneously service multiple activities and in the relative importance of activities changing with time (*Leontiev* [1981]). These phenomena are important aspects of multiplexing.

By now, subjects and tools should be clearly defined. Consequently, we focus our discussion on the object-motive component of an activity. The first point to stress is that "object" is intentionally singular. For each activity that a person is involved in (or working on), AT posits the existence of a single "thing" in the world which is the object of that activity. People orient to the objects of their activities in specific ways, with directed intentionality, often taking the form of a desired transformation for/to the object (*Leontiev* [1981]). This intentional orientation is the motive component of the object-motive pair.

The adoption of an object-motive pair has two major influencing factors. First, the possibilities a person perceives in terms of choosing an object-motive pair is limited by their socio-cultural situation (*Leontiev* [1981]). And second, object-motive pairs are adopted to satisfy a felt need (*Davydov et al.* [1983]; *Leontiev* [1977, 1978, 1981]). One way to understand the relationship between needs and object-motives is that needs provide the end goal, while object-motives provide the instrument for reaching that end goal (see Reiss for related discussion of ends, means, and basic desires, *Reiss* [2004]). To help clarify this point, we list some example objects-motive pairs and a likely need that they satisfy:

- creation a musical score – drive to create/be creative
- building of a house – drive to seek shelter/security
- creation of a legal case – drive to see/promote justice
- creation of a legal case – drive to display legal prowess
- writing of a research paper – drive to understand a problem/phenomenon
- execution of a planned vacation – drive to explore/for adventure
- deepening of a relationship – drive to experience a close personal connection
- killing of an animal – drive to eat (satisfy hunger)
- harvesting of a subsistence garden – drive to harmonize with one's environment

Notice that we have provided stereotypical needs in association with our example object-motive pairs. In reality, the subject(s) in an activity have unique, although generally highly related, motives in comparison to the people who have adopted similar object-motive pairs in the past. This historical similarity is the basis for the stability in the evolution of human behavior (*Cheyne and Tarulli* [1999]; *Leontiev* [1977, 1978, 1981]; *Miettinen* [2001]; *Scribner* [1985]; *Vygotsky* [1978, 1985]; *Wertsch* [1979, 1985a, b, 1991, 1998]). Likewise, the 'momentum' of object-motive pairs, or classes of object-motive pairs, is fueled by the tendency for people to adopt similar object-motive pairs to the people around them. This process is deeply connected to human development and education (*Leontiev* [1977, 1978, 1981]; *Scribner and Cole* [1981]; *Scribner* [1985]; *Vygotsky* [1978]; *Wertsch* [1985b]).

AT is not an alternative to or in conflict with the traditional HCI task-perspective, but complementary to it. Activities are at a different level and also represent a different kind of structure (*Kuutti and Bannon* [1993]). They are realized by the actions that comprise them, and in return they provide *stable* context for making sense of those actions (*Nardi* [1996]). For example, suppose a defense lawyer is preparing the opening statement in a case where she believes the defendant is guilty. If her primary motive is to see justice done, she might prepare a speech that subtly argues for an appropriate punishment in spite of trying to defend the client. Conversely, if her motive is to demonstrate her skill at navigating the legal system, she might prepare a speech with the intent of proving her client innocent based on a loophole in the law. Although trite, this example illustrates how the activity level of human behavior provides a handle on the relevant context for interpreting the actions of people.

So how does AT answer, or help to answer, the questions posed by our definitions of activity and context? To the first questions, regarding intentional behavior, AT offers a clear answer. Intentionality in behavior, according to AT, is captured by the object-motive pairs of a person's activities. Essentially, a person's need-based drive to interact with, and possibly transform, certain objects in their life is AT's explanation of “intention”. Similarly, at the action level, AT translates intentionality into conscious goals. As the other theories and models we survey below show, this is not a comprehensive interpretation of “intention”. It is limited in its explicit connection to an activity object and to conscious goals. Other theories and models that do not recognize an object as the defining component of activity help us overcome this limitation.

In terms of the production of context markers and the production of the mapping between markers and behavior, AT offers a partial answer. Namely that through the externalization process, or the process of sharing object-motive pairs socially, people produce various linguistic and material representations of their intentional behavior. These linguistic and material representations can, and almost always are, treated as context markers that influence present and future activity. For example, past dissertations are externalizations of many people’s academic activities. These dissertations, and the associated meta-discussion of the process of their production, have directly and indirectly steered the production of this dissertation. These dissertations are a material trace that constrains, and will continue to constrain, present and future dissertations.

The notion of constraint is connected to the idea of mappings between markers and behavior. One way to understand this connection is to think of constraints as a form of inclusion or exclusion relation between markers and behaviors. For example, the markers that distinguish an office setting have certain exclusion relations to “unprofessional” behavior and “unprofessional” genres of discourse. Sometimes these links are reified in various forms, e.g. legal definitions and consequences of “harassment”. These relations are often left tacit/implicit as well, e.g. the expectation that co-workers should occasionally have lunch together.

To briefly summarize, externalization is the process identified by AT that produces context markers and the mapping between these markers and behavior. What is missing in AT's treatment of externalization, however, is a more nuanced discussion of how markers and mappings can be grouped in meaningful ways, as they are not all ontologically equivalent or universally applicable. As we discuss below, other theories and models address these meaningful groupings more directly, providing a more complete answer to this question about activity.

Regarding the connection between individual behavior and socially shared ideas and artifacts, AT provides perhaps the best explication of this process through its dual notions of internalization and externalization. Again, internalization is the process by which individuals adopt object-motive pairs from their social and material environment. Externalization is the process by which individuals share their object-motive pairs. So AT explicitly describes the process by which ideas and artifacts (as the material traces of object-motive pairs) link individuals together to form social structures. As with the notion of "intention", AT's explanation of how individuals are linked to social structures is limited by the explicit focus on object-motive pairs. Using other theories and models, discussed below, we overcome this limitation and describe a more generic process by which individuals link to social structures.

Regarding the term "appropriate" as it modifies behavior, AT offers only indirect, partial answers. One partial answer is that "appropriate" might mean triggered by one's environment – through the conditions that impact operations and through the social and material factors that influence action goals. Another partial answer is that "appropriate" is implicitly enforced by the socio-cultural factors that render certain object-motive pairs visible, and hence options for adoption, while rendering other object-motive pairs invisible, and hence unavailable for adoption. Yet another partial answer is that within the scope of an action's goal, or an activity's object-motive pair, certain behaviors enable progress towards these ends while others do not. Those that enable progress we could treat as appropriate, while those that do not we could treat as inappropriate. However, each of these partial answers does not fully explicate the term "appropriate". Aspects of this term that are left unexplored by AT are how this term, and the process of applying this term, are part of social discourse and hence directly accessible and contestable (in addition to the implicit impacts that are captured by AT's conception of this term).

AT has only partial answers about the identity of context markers. While they are clearly implicated in the conditions of operations, it isn't clear what they are. Likewise, the process of externalization produces material and linguistic traces of object-motive pairs, which are often markers. But are all markers related to object-motive pairs? As we discuss below in our survey of other theories and models of behavior, it is likely the case that most context markers that influence behavior were produced as the externalization process for some object-motive pair at some point in history. However, the connection between a context marker and an object-motive pair is often lost as cultures and behaviors evolve.

Finally, regarding the interpretation of behavior, AT does not offer any definitive answers. Clearly, interpretation is coupled with the notion of appropriate. However, what exactly does it mean to interpret behavior? Is this a purely conscious process – i.e., an AT action? Or, can interpretation happen at the operation or activity levels? Is it functionally separate from "acting" – i.e., from AT operations, actions, or activities? If not, how is interpretation triggered by environmental conditions, or how does it impact the goals of actions, or how does it service the object-motive pair of an activity? While we could posit reasonable answers to these questions in line with AT, we rely instead on other theories and models of behavior that more explicitly address these, and other, issues regarding interpretation.

Vygotsky's (and others') Development of Mediated Action and a Genetic Perspective on Human Behavior

Lev Vygotsky was one of the founding contributors to the socio-historic school of thinking that grew out of Russia during the communist revolution. His theories on human behavior significantly impacted AT, as should be clear from the discussion below. His major contributions were to the field of developmental psychology (within which he is often placed on equal ground with Jean Piaget). We address Vygotsky's impact on psychological development in greater depth when we talk about literacy and education. For now, we focus on his conception and development of the idea of mediated action.

The root of mediated action for Vygotsky was his revision of the commonly understood, at the time, connection between a stimulus and a response (understood from the traditional Pavlovian perspective). What Vygotsky changed in this simple model was to augment the stimulus and response connection with a mediating mechanism. The mediating mechanisms for Vygotsky were of two types: tools and signs/symbols (*Vygotsky* [1978]; *Wertsch* [1985b]). The difference between these forms of mediation is that tools are externally oriented while signs are internally oriented. The implication of this mediation mechanism is that people can disrupt, and redirect, stimuli. The importance of signs/symbols, or more generally semiotic mediation, has been further explored by contemporary socio-cultural theorists and scientists (e.g., *Wertsch* [1979, 1985a, b, 1991, 1998]; *Zinchenko* [1985]).

In Vygotsky's work, mediation is the key element that enables the development of "higher psychological functions" such as memory, language, or logical reasoning (*Leontiev* [1981]; *Vygotsky* [1978]; *Wertsch* [1985b]). This mediation mechanism facilitates the internalization of social behavior in individuals. Vygotsky understood the internalization process as a transition from the inter-psychological plane, where people are explicitly interacting and collaborating to perform a behavior, to the intra-psychological plane, where an individual can perform the behavior "on their own". The internalization process is marked by a transition of speech exhibited by the individual. The inter-psychological plane is characterized by social speech, which is a primary mechanism supporting the coordination of behavior. As the internalization process proceeds people exhibit ego-centric speech; where social speech used to coordinate the behavior, it is now spoken out-load, to the self, to help the individual coordinate their own behavior. In this stage, individuals are their own "coach". Once the behavior has been fully internalized, the speech that was once social, and then ego-centric, is also internalized. Vygotsky referred to it as "inner speech". What is interesting about this transition of speech is its systematic abbreviation during the process of internalization – more and more components of the original speech, starting with the subject components of the sentences, are omitted.

Along with this micro-genetic and onto-genetic internalization of behaviors, Vygotsky also emphasized two other levels of development: the phylogenetic and the socio-cultural (*Leontiev* [1981]; *Scribner* [1985]; *Vygotsky* [1978]; *Wertsch* [1985b]). His basic argument is that internalized behaviors, and their social counterparts, are part of a larger socio-cultural milieu which is reified in the language and tools of these behaviors. The connection to AT on this point is almost direct. In addition to this socio-cultural milieu, Vygotsky also emphasized phylogenetic development. Although, as is widely acknowledged, he understood that these forces of development have not disappeared. However, they are largely eclipsed by socio-cultural influences.

We turn now to the questions posed by our definitions of activity and context. Instead of addressing every question, we only discuss those whose answers are elaborated on by Vygotsky and his theories on mediated action and genetic development. As with AT, Vygotsky's theory on internalization has clear implications for the connection between individuals and socially shared ideas and artifacts. Whereas AT treats the internalization process in complete generality, Vygotsky provides some additional structure to this process by foregrounding terms like inter- and intra-psychological planes, social speech, ego-centric speech, and inner speech. By explicating the process by which individuals adopt and appropriate socially shared ideas and artifacts, we have a better understanding of how individuals are linked to these social elements. Fundamentally, this link is, for the individual, an educational or developmental experience that is driven by the individual's appropriation of a new mediating tool or sign. More to the point, Vygotsky strongly emphasized the social interaction that initiates this internalization process. In his theories, that fact that people's behaviors are coordinated, i.e., made public and social, makes various mediating tools and signs available for others to internalize. And, it is through social interaction and coordinated behavior that these tools and signs (artifacts and ideas) are given the necessary framing so that individuals who are internalizing them also know how to use them (i.e., what behaviors they mediate).

The learning of behaviors that are mediated by internalized tools and signs is another aspect to the notion of "appropriate" behavior. An overly simplified description of this aspect of appropriate behavior is the following: through internalization, individuals build up a "toolbox" of mediating tools and signs that they know how to effectively use. The effective use of these tools and signs is in relation to a set of behaviors, or classes of behaviors. These behaviors, which are effectively enabled by an individual's toolbox of mediating tools and signs, are the appropriate behaviors associated with the socially shared ideas and artifacts (tools and signs) that they are aware of. In actuality, multiple behaviors are often being carried out simultaneously through an orchestration of many tools and signs. This multiplexing enables complex social interactions like ironic or sarcastic dialogue or pantomiming for story-telling.

Finally, the idea of behaviors being socially coordinated and then internalized along with mediating tools and signs offers another partial answer to the question concerning the interpretation of behavior. In some sense, an individual can, through his/her own practical knowledge of how to perform that behavior and of which mediating tools and signs effectively enable that behavior, understand the behavior of others. In fact, this imitation-based connection to the behavior of others is known to be deeply related to biological human instinct and is often acknowledged as the root of human sympathy and empathy (*Tversky et al.* [2002]).

Engeström's ideas on Contradiction and Expanding Objects

Whereas Vygotsky is often attributed as a founding contributor to the socio-historical school, which fueled the development of AT, Engeström is often considered the most contemporary AT theorist. However, Engeström's work has taken AT in new directions. Leontiev and other founding AT theorists focused on individual behavior and development, and its connection to the socio-cultural milieu. Alternatively, Engeström focuses on group behavior, in the form of group activities (*Engeström* [2000, 1987]; *Engeström et al.* [1999b]; *Engeström* [1999]). Engeström's concern with group activity is its developmental dynamic: How does it unfold in time? How does the relationship between involved people and artifacts evolve?

Engeström's focus on group activities and their temporal dynamics is reified in his expansion of the original AT activity triad (see Figure 2.3). Three new elements were introduced: rules, community, and division of labor. These three elements are, as Engeström convincingly argues, important components in understanding human behavior (*Engeström* [1987]). They stem from the formation of a species in their struggle for survival. However, these three elements seem arbitrary in modern human life. For example, elements like physical characteristics of the environment or history of related activities can have as much impact on human behavior as rules or division of labor.

So, while we do not agree with Engeström's choice of which additional elements to add to the AT triad, we are grateful for his explicit consideration of other elements/factors that impact human behavior. While many of these elements/factors are implicitly included in Vygotsky's and Leontiev's exposition of human behavior (*Leontiev* [1981]; *Vygotsky* [1978]), Engeström's work makes some of these explicit. How can an influencing factor of human behavior be made more explicit? To answer this question, one must first develop an ontology of such factors. The building consensus is that these factors are not independent, clearly identifiable, aspects of the world. Rather, these factors are epistemological idealizations that orient a human observer to a perspective on human behavior. These perspectives highlight certain qualities of human behavior while hiding others (*Agar* [2004]; *Bateson* [1972]; *Levine* [1971]). However, the fact that an observer notices certain qualities and not others is not evidence that the unobserved qualities are not present (*Agar* [2004]). It is in this way that we need to understand what it means to make influencing factors of behavior explicit. The process of making these factors explicit is a process of taking an observational perspective. And so, while rules, community, and division of labor are important qualities of human behavior, their existence is not independent of all the other elements/factors/qualities of human behavior. Consequently, we choose to take a more agnostic and situational position where we leave influencing factors out of our explicit model of behavior (as with the AT triad); and, when we apply our model to concrete situations, we employ a series of perspectives to highlight influencing factors (such as rules, community, and division of labor) as they are, or are not, relevant in the situation.

We turn now to Engeström's ideas on the temporal dynamics of human behavior. For Engeström these dynamics are driven by a series of encounters with contradictions. Foremost are contradictions between the subject(s) predicted progress (as captured in the goals of their actions) and their actual progress. Engeström claims that these contradictions, and their resolutions, largely determine the course of an activity (and hence the patterns of tool usage). Second, their are also inter-activity conflicts caused by the natural multiplexing that occurs when many people try to coordinate their behavior. Finally, because Engeström is biased toward social components like rules and labor hierarchies, his studies are quite sensitive to patterns around social roles.

For Engeström, the notion of social roles extends beyond common categories like manager, staff, technician, etc. He explores the interplay between these common roles, often specified at the start of a group activity, and the contradictions that force the specification of these roles to evolve as the group activity unfolds. In addition to role evolution, Engeström focuses on the evolution of the object of the activity (see the discussion of AT above for more details on the notion of object). And, as with role evolution, Engeström assumes that contradiction, between people's vision of how the activity should unfold, drives the evolution of the object in a group activity (*Engeström* [2000, 1987]; *Engeström et al.* [1999b]; *Engeström* [1999]). This evolution of the object of the activity is understood by Engeström as a learning/developmental process for the people in the group. As people learn and develop, the object of the group activity "expands". It takes on new meanings in addition to the current and historical meanings that it had.

The contradiction-based evolution of the group activity object is a more explicit model of the socially coordinated behavior that initiates Vygotsky's internalization process. However, we are forced to question whether socially coordinated behavior is often organized by a single, shared object. How often do people working together share a single vision of what they want to accomplish? The fact that this is frequently not the case (*Carroll et al.* [2006]; *Hill et al.* [2001]; *Hinds and Baily* [2003]; *Hinds and Mortensen* [2005]; *Mohammed et al.* [2000]; *Mohammed and Dumville* [2001]) is acknowledged in Engeström's focus on contradiction. So, it is perhaps more appropriate to think of multiple interacting activities, rather than a single group activity.

Engeström handles this perspective partially by exploring the influence that multiple, ongoing, group activities can have on each other. However, he still assumes that each of these group activities has a single object.

With some of the major aspects of Engeström's theories and models of activity described, we turn now to the questions posed by our definitions of activity and context. First, we address the production of context markers and the mappings between these markers and behaviors. As with Leontiev's conception of AT, Engeström's theories on activity include the idea of externalization. This externalization process for Engeström, as with most of his understanding of human behavior, is driven by contradiction. And, like Leontiev's AT, the externalization process is focused on the objects of activities. The difference is that Leontiev's activity objects are generally, although not always, individual whereas Engeström's activity objects are shared. By being explicitly shared, Engeström's externalization process can produce context markers and mappings between markers and behavior that are immediately social and shared. The contradictions that drive this process imply that these markers and mappings are contested in their production.

The next question that Engeström's theories on activity address are how individual behavior produces socially shared ideas and artifacts. As discussed above, the externalization process in group activities naturally produces socially shared ideas and artifacts. Less superficially, Engeström's theories on group activity reveal the nature in which socially shared ideas and artifacts are produced. They are contested outputs of a process that is driven by contradictions. Furthermore, these socially shared ideas and artifacts are not the products of an isolated group activity that can be separated from other activities or the other aspects and factors that influence human behavior. From Engeström, we are at least aware of rules, communities, and divisions of labor. But, as we discussed above, these are only three of the possible influencing factors of behavior. Other perspectives highlight other relevant factors (e.g., an architectural perspective might emphasize the spatial organization of the environment and its impact on behavior).

Regarding the determination of "appropriate", the group nature of Engeström's activities provides a clear perspective on the social aspects of this term. Furthermore, as with the production of socially shared ideas and artifacts, the social determination of "appropriate" can be a contested process, driven by contradiction. This characterization of the process by which behaviors can become "appropriate" is more in-line with other sociological theories like Actor Network Theory, which we discuss next.

Finally, Engeström's theories have something to say about context markers and the process of "interpreting" behaviors. As with almost every other aspect of Engeström's theories, "interpreting" behavior is a process driven by contradiction. As a group of people externalize their activity, the materials and language they use become markers for the group to interpret the state of the activity. These markers are one of the foci around which contradictions form – especially when the interpretation of the marker, and hence of the activity, is not universally accepted (*Engeström* [1987]).

Actor Network Theory

Actor Network Theory (ANT) provides a similar perspective on the organization of people and artifacts to AT (*Latour* [1995, 1999]; *Law* [1986a, b, 1992]; *Ryder*; *Stalder* [1997]). ANT refers to these organizations as networks, and every element in these networks (people and artifacts alike) as actors. However, unlike AT, ANT is focused on networks being structured for and by power struggles (*Ducheneaut* [2003]). Thus, actors are often referred to as allies (e.g., a researcher trying to gain clout in their research community might rely on more senior members of the community to vouch for and support them publicly). This can be performed with the actual people, or indirectly by citing and relying on their publications. This process of gaining allies, called "interessment", is part of the "translation" mechanism designed to build "obligatory-passage-points" (*Law* [1986b, 1992]; *Ryder*) – i.e., a network which must be defeated or deconstructed to retake the power gained by its creation.

An interesting connection between AT and ANT is covered by the ANT notion of "black boxes". Essentially, ANT recognizes that well-structured networks can become black boxes ("punctualized") – points of passage that are no longer questioned or generally challenged. In ANT, punctualization is a social process whereby the people who have culturally accepted a network cease to question its inner workings. In AT, this notion is captured on a more individual level. For the subject(s) in an activity, as their habit-based skills increase, many patterns of behavior (involving relationships between subject(s) and tools) will move from the action level down to the operation level. The classic example of this process (*Leontiev* [1978]) is the development of habitual reflexes while driving a car. People generally start driving by consciously controlling the steering wheel, shifting mechanism, and gas/break pedals. After some practice, these movements become sub-conscious operations. Interestingly, both AT and ANT describe situations where these black boxes are called into question. For AT, an unexpected condition can push operations back to the conscious action level – e.g., when road conditions are particularly hazardous. In ANT, a power struggle can result in various black boxes being re-opened to re-construct the network of actors/allies.

As in Engeström's theories of group activity, ANT recognizes that each actor in a network does not perform the same role. Rather, the varying capacities, in both type and endurance, of each actor relegates him/her to a specific role within the network. For example, a doorknob has no power, but acts as the gate-keeping mechanism or human passage. A person must, unless the door is already open, interact with the doorknob in the fashion dictated by the doorknob. A more quintessential example of the special roles of different actors in ANT is that of the "immutable mobile" text. In tracing the spread of knowledge, ANT theorists have often (e.g., *Brown and Duguid*

[1995]) recognized the powerful role that written text has played. Both its mobility, in book or scroll form, and its permanence (or immutability) enables a transfer of ideas across time and space in a way that was not possible before. In this way, both the local actor-network that produces texts, and the subsequent actor-networks that consume and use this text are connected in an interesting and vital way.

For us, two key aspects of ANT should be highlighted. First, ANT does not distinguish between levels of networks. While networks are clearly composed of and intersect with other networks, there is no notion of levels like operation, action, and activity from AT. Likewise, there is no notion of networks being oriented primarily to environmental conditions, goals, or an object-motive pair. The reason for this is likely that every actor in the network is considered to have agency, human or otherwise. Thus, no single actor or sub-group of actors in the network is allowed sole control of the network dynamics. Actors' behaviors are overlapping, often coordinated, sometimes conflicted, through time. This point provides an interesting contrast to AT's notion of intention in behavior (where the subjects of the activity are clearly dominant over the tools). We are not the first researchers to recognize this difference. In fact, many researchers have sought new theories and models of behavior to help bridge this gap between ANT and AT (e.g., *Blackler et al.* [2000]; *Brown and Duguid* [1991]; *Carroll et al.* [2006]; *Engeström et al.* [1999a]; *Engeström* [2000]; *Nardi et al.* [2002]; *Shaw and Gaines* [1999]; *Wenger* [2000]; *Zager* [2002]). The second key aspect of ANT for us is the notion of punctualization, or black-boxing – essentially that an entire actor-network can be treated as single entity. This potential regard for an actor-network has implications for what is externalized by the network as it acts (to borrow the term from AT).

We turn now to the questions posed by our definitions of activity and context that are addressed by ANT. First, as noted above, the fact that every actor in an actor-network is assumed to have some amount of agency (i.e., influence over the dynamics of the network) calls into question the clear ego- and human-centric notions of "intention" in behavior that we have discussed with AT and other theories above. Within some actor-networks, the human actors may not have clear goals or motives that drive the dynamics of their behavior. Rather, they may be simply responding the agency of the non-human actors in the network (recall the doorknob example above). While this perspective does not exclude the possibility, and likely certainty, that the human actors have goals and motives while they are behaving, it calls into question whether it is these goals and motives that are the true drivers of the behavior at the time. Ultimately, we think that "intent" can, and sometimes should, be accepted as coming from the non-human actors in a situation.

Regarding the production of context markers and the mappings between these markers and behaviors, ANT provides some additional processes, beyond externalization, that we need to consider. Namely, the notions of translation and punctualization help to highlight other methods of production. While each of these notions is clearly connected to externalization, they emphasize different aspects of this process. Translation provides a clearer perspective, than the more localized notions of internalization and externalization, on how socially shared ideas and artifacts can, and are, produced. Specifically, one can think of their production as an iterative design/modification process. Accordingly, each internalization and externalization combination "translates" the idea/artifact into a (slightly) new idea/artifact, which addresses different needs than the previous idea/artifact satisfied. Similarly, punctualization, or "black-boxing", of networks emphasizes the potential production of markers and mappings that represent potentially vast actor-networks (e.g., "the government" or "my car"). Often, these punctualized networks represent the coordinated efforts of more people and artifacts than can be adequately addressed by Leontiev's or Engeström's notions of activity. Yet they clearly exist and represent aspects of human behavior. Accordingly, we rely on ANT to expand our understanding of the production of context markers and the mappings between these markers and behaviors to include processes that are either larger than AT activities, or fall between the different levels of behavior specified by AT.

ANT provides two key answers to the question concerning the link between individual behavior and the production of socially shared ideas and artifacts. First, within an actor-network, each actor's behavior is coordinated to produce a network level effect. In this way, the multiple co-dependencies between actors in the network help to constrain individual actors' behaviors and to combine these behaviors in a productive way (at least in functional actor-networks). Second, ANT describes an inter-network dependency referred to as obligatory-passage-points. Obligatory-passage-points are punctualized networks that must be addressed (referenced, consulted, etc) in present and future behavior. In this way, individual behavior is coordinated (through specific "passages") to produce and recognize ideas and artifacts that are socially shared. The scope of these ideas and artifacts depends on the temporal and spatial scope of the obligatory-passage-points that create and maintain them.

The notions of intra- and inter-network constraints on behavior also have impact on the notions of "appropriateness" and "interpretability" of behavior. Clearly, those behaviors that are allowed/permitted (i.e., not ruled out) by other actors within a network or by obligatory-passage-points are candidates for being labeled "appropriate". The "interpretability" of these behaviors is constrained, both consciously and subconsciously, by a person's recognition of these constraints. Interestingly, the notion of punctualization has important implications for the recognition of these constraints because it generally indicates a lack of awareness of how to deconstruct, or better understand, black-boxed constraints like obligatory-passage-points.

Finally, ANT provides an interpretation of "marker" that we find useful. Again, markers are perceivable differences in the world. Their importance is in recognizing that people's activities are sensitive to the perception of these differences – i.e., people act differently depending on what markers they are around. In this sense, any actor in an actor-network is potentially a "marker" that, according to ANT theorists, can influence and constrain one's behavior. This "agency" of markers, as understood from the ANT perspective, is useful in trying to describe how context, understood as sets of these markers, influences a person's activities.

Distributed Cognition

Like ANT, Distributed Cognition (DCog) focuses on the patterns of people and artifacts in the performance of various tasks/actions (*Halverson* [2002]; *Hollan et al.* [2000]; *Hutchins* [1995]; *Hutchins and Klausen* [1996]). The DCog perspective is that people (AT subjects) offload their cognitive burdens on artifacts (AT tools) in the world. There is less emphasis on power struggles, since the non-human artifacts are considered subordinate, and more focus on the cognitive contributions that artifacts make in the performance of various actions. The capacity for artifacts to store and enable various cognitive operations is an acknowledged component of AT (e.g., tools are often designed to encode the operations they are intended to enable, *Leontiev* [1978]). DCog-based studies, generally qualitative in nature, have demonstrated this characteristic of tools in multi-person settings such as ship navigation rooms (*Hutchins* [1995]) and airplane cockpits (*Hutchins and Klausen* [1996]).

In many ways, DCog mirrors ANT by focusing on networks of human and non-human actors whose behaviors are coordinated. However, as noted above, DCog networks are explicitly focused on a goal. The goal is the driving motivation for the DCog network. This focus, along with the focus on the cognitive offloading that happens between human and non-human actors in the network, provides some minor additions to the questions posed by our definitions of activity and context.

Specifically, offloading cognition is a specific type of externalization that lends itself to understanding the current state of the process (i.e., progress toward the goal) and hence the "appropriateness" of the behaviors of the actors within the DCog network. Assessment through interpretation can happen within the behaviors that are being coordinated to reach the goal. Likewise, the nature of the externalizations, that they are cognitive middle-steps, enables a level of interpretation that is vital toward the education and training of the actors within the DCog network (*Hutchins* [1995]).

Cognitive offloading also has clear implication for the production of markers and the mapping between these markers and behavior. On short time scales (order of seconds), people in DCog networks are creating markers in the form of externalized cognitive middle-steps, serving to coordinate their next behavior. As an example to illustrate this point, consider the use of pencil and paper when doing long division. By marking the middle steps of the problem, a person creates their own markers to which they need to respond. It follows that context markers can be, in addition to actors in an actor-network or the externalized linguistic and material traces of AT object-motives, cognitive middle-steps recorded within a DCog network as it operates.

Literary Genres, Speech Genres, Genre Ecologies and Genre Tracing

Another important class of theories and models that inform our discussion of activity and context relate to the notion of genres (*Freedman and Medway* [1994]; *Russell* [1997]; *Spinuzzi* [2003]). Through Engeström's theories on group activities and the structural dynamics specified by ANT, we have already introduced the notion of roles for human actors (subjects). The counterpart in the material world to the notion of roles is genres, which can be generically understood as a fuzzy classification of linguistic and material artifacts based on their perceivable characteristics. For example, poetry is often separated from prose as hammers are often separated from saws. The connection between ideas in AT and ideas in genre theory have been well-studied (e.g., *Erickson* [2000]; *Freedman and Medway* [1994]; *Russell* [1997]; *Spinuzzi* [2000, 2001, 2002, 2003]; *Wertsch* [1979, 1985a, b, 1991, 1998]).

We are not concerned here with the theory of genres and related topics, per se. Rather, we are interested in how these theories can inform our understanding of activity and context. Genres and related notions provide a rich analytical framework for understanding how mediational artifacts (tools and signs/symbols, see the discussion of Vygotsky above), and more generally any externalized material trace of behavior, can and often is treated as belonging to an abstract class, or type. In other words, people often perceive and understand material artifacts, linguistic or otherwise, from the perspective of an abstract class or type. Thus, while a person's behavior is often mediated by specific material artifacts, the interaction with these artifacts is itself mediated by the perception and awareness of an abstract classification/typology over these artifacts. Importantly, these classifications and typologies are often fuzzy. The fuzziness is often the result of behaviors and artifacts belonging to multiple classes/types, promoting a simultaneous multiplexing of interpretations. The theory of genres, and the related notions discussed here, help to explicate this secondary mediational process.

Of course, genres themselves can be classified – examples include literary genres (*Breure* [2001]; *Freedman and Medway* [1994]; *Karlgren and Cutting* [1994]), speech genres (*Bakhtin* [1979]; *Cheyne and Tarulli* [1999]; *Hagtvet* [2003]; *Wertsch* [1979, 1985a, b, 1991, 1998]), informal/ad-hoc genres (*Delin et al.* [2000]; *Karjalainen and Salminen* [2000]; *Millen et al.* [2005]; *Paolillo* [1999]; *Proctor et al.* [1998]; *Spinuzzi* [2000, 2001, 2002, 2003]; *Yoshioka et al.* [2000]; *Yoshioka and Herman* [2000]; *Yoshioka et al.* [2001]), and conversational genres (*Atkinson and Heritage* [1984]; *Bakhtin* [1979]; *Erickson* [2000]; *Goffman* [1959, 1967, 1974]; *Palen and Dourish*

[2003]; *Searle* [1969]; *Turoff et al.* [1999]). What is common to all of genre theory in general is the notion of a duality of structure: a genre, or abstract typology of specific artifacts, exists only in as much as people continue to recognize, contribute to, and maintain it. But, to enable people to maintain it, a genre must then constrain and guide their behavior. In this way, genres have a dual existence – they constrain the behavior which creates them.

Notice that this same duality of structure applies to context markers when we understand their influence on behavior as constraints and guides that exist only as long as people continue to create and perceive them. The continual re-creation of genres (and context markers) leads to their characterization as emergent phenomena – the practically endless feedback loop that generates them could just as easily have generated a (slightly or drastically) different set of genres. We re-visit the idea of a duality of structure below in relation to the theories of structuration and frames.

Focusing on speech genres, we highlight an additional idea that furthers our understanding of activity and context. Specifically, Bakhtin's development of the theory of speech genres is heavily reliant on the notion of dialogism and its more general version, multi-vocality (*Bakhtin* [1979]). Dialogism is the idea that written text, like spoken speech, is also dialogic in nature. Although the "other" to whom the text is addressed may be separated by time and space, and also may be known only stereotypically by the author, the text is still written and conceived with this other in mind (even if only implicitly). In fact, most text has at least three contributing perspectives/voices. The first is the author's – what he/she/they want(s) to say. Second is the perspective of the intended audience of the text – for which the author is generally trying to find the most effective message. And third, the author, through the reconstruction of one or more literary and speech genres, is engaged in a dialogue with these genres. As Bakhtin convincingly argues, one cannot hope to understand a text, written or spoken, without assessing the many voices employed simultaneously by the author. Multi-vocality is yet another form of multiplexing.

Related to both the theories of literary and speech genres are Spinuzzi's theories on genre ecologies and genre tracing (*Spinuzzi* [2000, 2001, 2002, 2003]). Spinuzzi's theories are a meta-theory about genre that captures inter- and intra-genre dynamics. Genre ecologies are the sets of genres employed in a given setting or situation. For example, in the traditional office work setting, certain speech genres are employed in combination with certain literary genres as well as certain computational/technological genres. By using specific artifacts from each of these genres in concert, office workers are able to carry out their various tasks and job functions. An important aspect of how these genres are used, however, is that each office worker often needs to modify (appropriate) the specific artifacts in the genres to fit their particular needs. In other words, people often redesign tools from prototypical genres to better fit their local, situated needs and goals. However, these adaptations tend to preserve much of the existing genre class structure, enabling a tracing of the genre through space and time as it is modified and re-created.

We turn now to the questions posed by our definitions of activity and context. We address only those questions whose answers are impacted by the theories of genres we have discussed. First, as markers and the mappings between markers and behaviors exist in their own genres (see, for example, the discussion of frames below), they are generally subject to the processes that determine genre (re-)creation. The key aspect of this genre (re-)creation process is captured, in slightly different ways, by the notions of duality in structure, multi-vocality, and genre tracing. Basically, people continue to create material artifacts – traces of their activities/behaviors – that are simultaneously constrained by existing genres and extend the lives of these genres by contributing another specific artifact to their existence. Importantly, however, Spinuzzi recognizes that these genre re-creations are not identical replications – they represent the local, situated appropriation of these genres by individuals as they carry out their activities.

Moreover, the constraints generated by existing genres on the people trying to re-create them provide yet another way to understand the term "appropriate" as it modifies behavior. An "appropriate" behavior is one that effectively re-creates the genres (literary, speech, or otherwise) which are employed by the person in the behavior. As Spinuzzi points out, effective re-creation does not imply replication. Rather, one must simply maintain enough categorical resemblance between their specific instantiations of genres and the abstract definition of these genres so that observers, the actor included, can correctly associate the specific genre instantiations with the right genres.

Of course, it follows that our discussion of the theories of genre provides a partial answer to the question of how to interpret behavior. Behavior that preserves the intended genre associations to specific elements in/of the behavior is interpretable by any other person who is knowledgeable about the genre. In fact, it is this idea of knowledgeability that enables the creation of socially shared ideas/artifacts to begin with (whether they are understood as genre instantiations, traces of object-motive pairs, or context markers).

The Grounded Theory of Action

Strauss's grounded theory of action (*Addelson* [1995]; *Soeffner* [1995]; *Star* [1995, 1997]; *Strauss* [1995]) describes not only a model of human behavior but prescribes a method for studying human behavior. The essential claim in the grounded theory is that human action (here "action" is analogous to AT action) is highly influenced by the surrounding situation. This influence results in two seemingly contradictory characteristics: first, that human action is dynamic and unique, and second, that human action tends to follow frame-like patterns (*Goffman* [1974]). The first characteristic has been recently popularized by Suchman (*Suchman* [1987]). The second characteristic has been recognized for some time in the social sciences (*Bateson* [1972]; *Dewey* [2002]; *Goffman* [1974]; *Leontiev* [1978]) and more recently in computer science (*Winograd and Flores* [1986]; *Bardram* [1998]; *Fjeld et al.* [2002]; *Nardi* [1996]; *Nardi and Redmiles* [2002]; *Shaw and Gaines* [1999]). Based on this model, the grounded theory prescribes a qualitative method for studying behavior that leaves ample room for the observers to reconstruct their understanding of what is happening and why – to build an emic understanding along with an etic one. This method is related to current conceptions of ethnography (*Agar* [2004]; *Blomberg et al.* [2003]).

For us, the grounded theory of action provides further support on two points. First, that the context markers comprising a "situation"/"environment" are fundamental drivers of human behavior. And second, that while patterns exist in the mapping between context markers and behavior, we should remain open and sensitive to possible changes that could happen in this mapping, even on short time scales (e.g., seconds or minutes). These points have important implications for two of our activity and context definition questions.

First, regarding the definition of "appropriate" behavior, the grounded theory of action emphasizes the impact of situational and environmental factors on present action. People are, in many ways, responding to their environment. And, while AT acknowledges this aspect of behavior primarily at the operation level, the grounded theory of action emphasizes that the responsive aspects of behavior extend completely into the AT action level of behavior as well. The situated action idea, also emphasized by Suchman (*Suchman* [1987]), reveals that many behavior-orienting goals are themselves responses to the environment – to markers in the world that are not in the control of the people who are responding to them. In other words, people set goals to handle immediate issues or contingencies. AT does not explicitly exclude these ideas.

However, the tone of AT is more inline with the perspective that people can and often do set goals based only on markers that they control (like the object-motive pairs of their activities). Notice as well that ANT is implicitly aligned with the grounded theory of action in that it definitely recognizes the agency that external artifacts and people can have on behavior.

The second question addressed by the grounded theory of action relates to the idea of "interpretability" of behavior. Specifically, all people observe the behavior of the people around them. All people are lay ethnographers and qualitative social scientists (*Giddens* [1984]). And, in recognizing the ubiquity of observation, we can include in our understanding of "interpretability" the tension between emic and etic perspectives – between the interpretation of behavior by the actors involved and the interpretation by outside observers. As a research method, the grounded theory of action can choose to emphasize the emic perspective. However, in understanding how all people observe the behavior of those around them, we need to be open to the various styles of observation, and hence interpretations, that actually happen. Thus, our notion of interpretability needs to include the entire spectrum from the emic to the etic.

Conversation and Discourse Analysis

In addition to fairly generic and comprehensive theories and models of behavior, we need to look at more localized and specialized theories. Two such theories are conversation analysis (CA) and discourse analysis (DA) (*Atkinson and Heritage* [1984]; *Garfinkel and Sacks* [1970]). The primary objective of these theories and fields of research is to explicate the important markers and influences of markers on human interaction as it unfolds moment to moment. For example: What are the ways in which an interaction is initiated? How are each of these initiations realized materially? And, how does each of these initiations shape the ensuing interaction? The markers used in CA and DA include tone, pitch, pauses, vocabulary, utterances, and body language in addition to more information based markers like Clark's common ground. Through the inclusion of such markers, CA and DA often extend beyond traditional linguistic paradigms (e.g., *Searle* [1969]).

Instead of discussing each of these theories and fields of research in detail, we simply highlight some key, high-level points that impact our discussion of activity and context. Following the grounded theory of action above, we understand behavior to be situated, or occasioned. Garfinkel and Sacks describe conversational language similarly: "the mastery of natural language is throughout and without relief an occasioned accomplishment" (*Garfinkel and Sacks* [1970]). This means that while activity and conversational language exhibit formal structure (i.e., genres) every enactment/reification of these structures is situated to practically serve in the present moment. Moreover, the genre structures of activity and conversational language have the following three qualities. (1) They exhibit uniformity, reproducibility, repetitiveness, standardization, typicality, etc. In other words, they are meta-level patterns that exist through the regularity of their application/usage. (2) These genre structures exist independently from any specific actors/people. One way to understand this is that the distributed and redundant storage of a genre structure, through its internalization by many people through time and space, makes it robust to the local behavior of specific actors/people. (3) People that enact genre structures are aware of their distributed and redundant storage. It is through this awareness – Clark's *common ground* – that people are able to effectively communicate a message by enacting a genre structure. They are, in the Bakhtinian sense, one of the voices involved in an utterance. Hence, through the mutual awareness of genre structures, communication is both enabled and made "accountable" (i.e., interpretable). Garfinkel and Sacks refer to the constructive analysis of an accountable behavior as a "gloss". In many ways, glosses are the unit of data in both CA and DA (the phenomena which can be treated singularly).

Clearly, CA and DA contribute to our understanding of activity and context. Most importantly, the notion of "accountability" is related to our notions of "appropriate" and "interpretable". As with other theories discussed earlier, "appropriate" is meant to capture the idea that certain behaviors should be performed given the presence of certain markers or the decision to enact certain genres. To state this negatively, the presence of certain markers or the decision to enact certain genres constrains the behaviors that one perceives as available to perform (provided they are trying to effectively respond to the markers or to effectively enact the genre). Appropriate behaviors are accountable behaviors; they fit in the socially acknowledged schemas triggered by markers and genres. Moreover, by fitting into socially acknowledged schemas, observers and receivers can rely on the schemas to help interpret behaviors, activities, and messages. In fact, the ability to interpret a message is included in the notion of accountability.

One other question that is addressed by CA and DA is the definition of a marker. Clearly, markers can include extra-linguistic things such as tone, pitch, pauses, vocabulary, utterances, and body language. These markers have very localized importance in time and space. For example, a listener's nod is over in a fraction of a second, but it informs the speaker that the listener is paying attention and that they should continue talking. Outside of the conversation, the nod might mark an entirely different message and behavior, and hence require a different interpretation.

Frames and Structured Interaction

Combining the analyses of CA and DA with a wider perspective on interaction moves us into the space of frames and structured interaction (*Bateson* [1972]; *Goff-*

man [1959, 1967, 1974]). Here we include forms of non-verbal interaction such as spectatorship and interactions that last over multiple sessions (e.g., group projects). Frames are, according to Bateson, the necessary background information enabling a message recipient to properly interpret the message. For example, a person waving their arms in the air (the message) means something different if (1) they are at a sporting event, (2) they are standing on the side of the highway, or (3) they are on stage in front of an audience. Each of these sets of framing information has important implications for the interpretation of an arm-waving message.

What Bateson and Goffman understood is that sets of framing information are not arbitrary. Rather, they are often organized into genre classes based on what information they usually include, what messages are likely to be sent given that framing information, and how such messages are supposed to be interpreted. For both Bateson and Goffman, the term frame often refers to the genre class, and not to particular enactments of the genre. Consequently, frames exhibit a duality of structure: they constrain how people act – in what messages they send and in how they interpret the messages that they receive – and they exist through people's enactment or usage of them. While Bateson was more interested in the ontology of frames, and hence in their abstract qualities, Goffman was interested in studying actual frames, focusing on, for example, what types of people were involved and what rituals of interaction they exhibited.

Interestingly, through the duality of structure frames act simultaneously at the level of activity (because people are enacting them) and at a meta-level (because they constrain how they are enacted). Goffman builds up an extensive dramaturgical metaphor to help explain what a frame is. Through this metaphor, Goffman is able to distinguish between within-frame and out-of-frame behavior, as well as how multiple frames can be simultaneously interlaced as people act. The most common example is the use of a "back-channel" frame. The applicability of multiple frames to a given sequence of behavior is an important aspect of behavioral multiplexing, related to Bakhtin's notion of multi-vocality. The dramaturgical metaphor also highlights the concepts of keying – the method by which a frame is made explicitly "turned on", and can henceforth be used to interpret the messages being communicated – and of roles.

Goffman also acknowledged the internalization of frames. This process is an important aspect of acculturation, resulting in a person's ability to sufficiently navigate and react to the enactment of all the "primary frames" of his/her social group. Moreover, through sufficient internalization and practice, people can start to nest frames within one another – e.g., to act like a cop (one frame) within the telling of a story (another frame) at a social gathering in a restaurant (yet another frame). Obviously, this capacity enables one to "fabricate" a frame – i.e., to use it in an unauthentic way.

At this point, we return to the questions posed by our definitions of activity and context. First, we address the question of how individual behavior leads to socially shared ideas and artifacts. Obviously, frames are social enterprises – they are both known and understood by many people and they often require multiple people when enacted. In other words, an individual's behavior is coordinated, although not dictated, by frames. Hence, an individual's behavior reifies, and has the potential to modify, socially shared ideas and artifacts that are part of frames.

Regarding the question of appropriate behavior, frames provide a very specific understanding. "Appropriate" behavior is within-frame whereas inappropriate behavior is out-of-frame or causes a frame break. This understanding of appropriate behavior is, of course, connected to the interpretation of behavior. As noted above, frames provide a vital mechanism for interpreting the within-frame behavior of other people. However, there is often a break-down in one's ability to interpret this frame-breaking behavior unless there is a higher, more encompassing frame in place. In other words, interpretation is facilitated within a frame and is challenged by the frames themselves when behavior is out-of-frame or frame-breaking.

Finally, regarding context markers, frames provide us with another perspective on the potential ontology of context markers. Following Bateson and Goffman, context markers are very often markers of frames – produced both during the keying and maintenance of a frame, and simply recognized in one's environment when it promotes the enactment of certain frames over others.

Giddens's Theory of Structuration

Giddens's theory of structuration (*Giddens* [1984]) builds on and generalizes the idea of frames and interaction rituals. In structuration, structures are socially coordinated frames or sets of frames that span various amounts of time and space. The smallest (in a time and space sense) structures are social structures which are unique to sub-cultures and the face-to-face interactions that they have. Large structures are system structures which span entire cultures and dictate more than just face-to-face interaction (e.g., the practices of the business world). The primary driver of structures is internalized, supportive routines (i.e., habits). These routines re-create the structures they support.

In many ways, Giddens's structures are the natural enlargement of Goffman's frames into the larger social enterprises that dictate human culture (e.g., work, sports, religion, etc.). Like frames, structures exhibit a duality of structure and are often multiplexed in their application. In additional to simply recognizing routine behavior, Giddens highlights the importance of notions like roles and resources in the re-creation of structures. Roles for Giddens are identical to those discussed above in Goffman frames and in ANT. Resources, alternatively, are understood not simply as some quantity of something, but as enabling different types of power relations to build within the structure. Authoritative resources are rooted in human agency, and hence are immediately under the influence of certain people; allocative resources are material in nature, and tend to be owned and controlled by some people.

Turning to the questions posed by our definitions of activity and context, we note that much of the discussion from the theory of frames also holds for Giddens's theory of structuration. By replacing the notion of frames with the notion of structures, we observe several relevant points. First, an individual's behavior reifies and has the potential to modify socially shared ideas and artifacts that are (and are part of) structures. Second, the notions of "appropriateness" and "interpretability" can be defined relative to structures, with the acknowledged emphasis on these terms being grounded in people's routines, in their praxis. And, finally, that context markers can be markers of structures – produced during the re-creation of structures and simply recognized in the environment when they promote the enactment of supporting routines for the structures.

Feld's Theory of Social Foci

Social Network Analysis (SNA) is the field of research that analyzes patterns in how people group together: who is talking to whom?, who is talking the most?, who is central to the group?, who is peripheral?, etc. (*Wasserman and Faust* [1994]). One of the guiding theories behind SNA is Feld's theory of social foci (*Feld* [1981]; *Wasserman and Faust* [1994]).

Feld attempted to understand what induces social groupings. The resulting theory of social foci is rather generic. It basically assumes that joint activities are organized around shared entities – psychological, physical, legal, etc. Engeström's shared objects in group activities are an example of social foci. Interestingly, Feld highlights the idea of two people sharing multiple foci. The "multiplexity" of their relationship correlates to the strength of the social tie for practical reasons (like having more or less common ground) It also indicates the potential diversity of interactions that the people can share. Notice that Goffman frames, AT objects, Vygotsky tools and signs, Bateson's cybernetic feedback information, and all types of genres are social foci.

With respect to our activity and context questions, social foci provide only a partial answer to the identity of context markers. Being socially shared, foci provide an interesting mechanism whereby people can collectively react to their environment, or coordinate their activities. While we do not assume that all context markers are drivers of social grouping (e.g., a dessert spoon), we should recognize that some context markers are social foci (e.g., a "cult classic" movie) and that this changes the nature of their influence on how people perceive them.

Dewey and James' Pragmatism

Pragmatism is a philosophy that emphasizes practical consequences as the basis for one's epistemology (*James* [1975]). One of its more extreme tenets is that truth is whatever one believes and can practically rely on – in both a predictive and descriptive sense. What is important for our present concern on activity and context is the idea that practical consequences can be the basis of a person's epistemology. This idea provides additional insight into the process of internalization. Whereas Vygotsky focused on the internalization of mediating artifacts like tools and signs (i.e., *how* things are done), Pragmatism focuses on the internalization of practical consequences (i.e., *what* things were done). While we do not want to lose our emphasis on mediation and mediating processes, we seek to include the idea that people develop an internalized understanding of the consequences of their actions. Moreover, this internalized understanding of practical consequences is a major component in people's epistemology – people *know* based on, at least partially, what they believe will happen.

Given this perspective on epistemology, we should ask if the internalization process for practical consequences differs from the internalization process for mediating artifacts. Here, Dewey's concepts of "transaction" and "inquiry" are important to consider (*Garrison* [2001]; *Miettinen* [2001]). For Dewey, meaningful human behavior extends beyond simply interacting with the world. By performing "inquiries", a person can both practically influence the world and learn about these practical consequences. Inquiries are a form of "transaction" whereby a person transforms him- or herself by synthesizing both physical and intellectual artifacts that are used and produced *while* they are practically changing the world around them. This transformation process parallels Engeström's theories on the transformation of an AT object-motive pair (see the discussion on AT above).

Finally, another characteristic of Pragmatism that is important to our discussion of activity and context is the idea of habit. Both Dewey and James recognized the strong influence that habits have over the behavior of people (*Dewey* [2002]; *James* [1975]). This idea complements AT's recognition that moment-to-moment behavior is largely the performance of operations (and that operations are essentially habits). From Dewey and James' perspectives, the practical consequences internalized by a person often result in the formation of habits. These habits form the driving engine of a person's epistemology, morality, and behavior. A simplified model is this: at some level, a person wishes to see certain practical consequences occur in the world; this desire leads that person to perform various habits that historically produced these practical consequences. It follows that habits are involved in both an individual's intent to perform behaviors and in his/her sense of appropriateness and interpretability of these behaviors.

Turning to the questions posed by our definitions of activity and context, we see that Pragmatism furthers our understanding on several of them. The concepts of inquiry and transaction highlight an aspect of intention in behavior that we have not yet addressed – that in addition to choosing a behavior/activity based on a desired impact on one's surrounding environment, people often choose a behavior/activity based on a desired impact on themselves. By internalizing practical consequences (both internal and external), people can choose behaviors/activities that satisfy goals/needs for both aspects of their existence.

The other questions addressed by Pragmatism concern the notions of appropriateness and interpretability. Habits are involved in both an individual's intent to perform behaviors and in his/her sense of appropriateness and interpretability of these behaviors. As both Dewey and James noted, a person's sense of appropriateness (or more abstractly, their values) is rooted in their habits. The behaviors people know how to do comfortably (i.e., habitual behaviors) are the basis for their understanding of and connection to the world. It follows that a person's ability to understand and determine whether a behavior is appropriate or not depends on whether they have already incorporated that behavior into their habitual conduct, and/or how that behavior compares to the behaviors in their habitual conduct that they might have performed in a similar situation/environment. In other words, a person's history of activity impacts their ability to interpret behaviors and to judge them as appropriate or not.

Cybernetics

Cybernetics is a rich field of study that has fallen in and out of popularity since its inception in the mid 1900's (*Wiener* [1948]). We will not concern ourselves with the entire field. Rather, we focus on Bateson's early theories on cybernetics (*Bateson* [1972]) and on how these theories impact our definitions of activity and context.

The overarching model of a cybernetic system in Bateson's theories is of a feedback circuit. Basically, a cybernetic system contains entities, in the most abstract sense of the term, that produce output. The output of every entity impacts the systems at large, and these impacts can be recorded as differences between that state of the world before the output and the state of the world after the output. These state differences, broadly understood as information, are fed back into each entity within the system. The feedback loop could continue indefinitely; and the boundary around a cybernetic system is always fuzzy. Basically, the goal in placing such a boundary is to isolate groups of input-output entities that have a high impact on one another. The quintessential example of a cybernetic system is an ecosystem.[1]

The cybernetic system paradigm is important to our analysis because it highlights the most essential criteria for conceiving of two entities together rather than apart. This criteria is whether or not the output from either entity can impact the input of other entity. Chaos theory might push one into the safe but debilitating perspective that all things are connected. However, the level of influence of two entities can be reasonably estimated as zero provided one limits the temporal and spatial ranges where this estimation is assumed to apply.

The important unifying characteristic of state difference is its capacity to be perceived by the entities in a cybernetic system. As Bateson repeatedly points out, for information to be perceivable it must be different from its surroundings. The amount of perceptible difference is a measure of the amount of information from the cybernetic system's perspective.

[1]Related to Bateson's theories on cybernetics are Maturana's ideas on autopoiesis and structural coupling (*Maturana*; *Winograd and Flores* [1986]). Maturana recognizes both the dependence of behavior on the environment and the dependence of the environment on behavior. By structurally coupling with the world, an organism is both influenced by and influences their environment. Autopoiesis is the process of managing these influences to stay alive, to self-preserve. Maturana extends his discussion of autopoiesis from the microscopic interactions of chemical systems within single-celled organisms to the thought processes of more complex animals, including humans.

The cybernetic system model is important for us because it informs our understanding of what context markers are (and could be). From a cybernetic perspective, a context marker is a perceivable state difference that (1) is input for at least one entity whose (2) behavior is changed by this state difference. In the language of this dissertation, context markers are any piece of information that impacts the activities of people when, and if, they are perceived. It follows that context markers include everything from small-scale outputs (e.g., the markings on a sheet of paper made while solving a long-division problem) to the material traces of object-motive pairs (e.g., a copy of a dissertation) to differences in one's environment that they did not produce (e.g., rock formations during a hike in the mountains).

Theories on Literacy and Education

Many of the theories we cited above talk to varying degrees about individual development. Often the conception of development is around two forms: literacy and intellectual education. Accordingly, we discuss some references on these topics to provide a more complete picture of how they, and individual development in general, influence activity and context.

Basically, the conceptions of literacy and education embedded (implicitly or explicitly) in the theories and models we have discussed share the idea that these forms of development are highly social (*Bateson* [1972]; *Bruner* [1966]; *Dewey* [2002]; *Leontiev* [1981]; *Luria* [1976]; *Vygotsky* [1967, 1978, 1985]; *Wertsch* [1985a, b]). People learn language and other higher psychological functions by participating in social settings where knowledgeable teachers can guide students in their development (*Brown and Duguid* [1991]; *Lave and Wenger* [1991]; *Scribner* [1985]). What is vital in these learning situations is that instruction be provided at the right level. Instruction should not be so hard that the learner cannot keep up, but hard enough that the learner is engaged and doing something they otherwise could not do (*Wertsch* [1985b]). An appropriate level of instruction lies within the learner's *Zone of Proximal Development* (ZPD) (*Vygotsky* [1978]; *Wertsch* [1985b]).

While it is beyond the scope of this thesis to explore these ideas thoroughly, we note that the idea of the ZPD is a vital concept in understanding an effective internalization process. Basically, people cannot develop without the proper guidance and without having internalized the necessary prerequisite skills. In other words, the process of internalization (of frames, routines, perception of markers, and mappings between markers and behavior, etc.) is dependent on who is trying to internalize these behaviors and practices. Hence, we need to be sensitive to a subjective perspective on activity and context in addition to a generic one.

2.3 Discussion of our Definitions of Activity and Context

In the remainder of this chapter, we discuss our definitions of activity and context in light of the theories surveyed above. Specifically, we provide definitive answers to the questions raised by these definitions in preparation for our translation of these definitions into a computational model. Additionally, we identify a number of factors that impact activity and context, but are not constituent elements of these notions.

First, recall our definitions of activity and context:

Activity is intentional behavior that produces both context markers and the socially shared, i.e., cultural, mapping between context markers and appropriate activities.

Context is a set of markers (perceivable differences in the world) that enables a person's assessment of which activities, and which interpretations of these activities, are appropriate.

From these definitions, we posed the following questions to answer by surveying existing literature:

- What do we mean by "intentional" behavior?
- What do we mean by "production", regarding markers and mappings between markers and behavior?
- How does individual behavior lead to socially shared ideas/artifacts?
- What do we mean by "appropriate" activities?
- What do we mean by "markers", or "perceivable differences in the world"?
- What do we mean by "interpret" behavior?

Combining the relevant contributions of each theory and model, we may now answer these questions. Regarding intention, we recognize the ego-centric perspective of AT and related theories but we also emphasize the idea that intent can be applied by external agents (ANT). More importantly, ego-centric notions of intent should include both internally- and externally-focused motives (e.g., Dewey's notions of inquiry and transaction and AT's coupling of needs satisfaction with the externalization of object-motive pairs).

In terms of the production of markers and the mappings between markers and behavior, the theories and methods we discussed have highlighted a number of interesting perspectives. Both AT and Engeström's theories on group behavior have emphasized the process of externalization, which can be driven and characterized by its many intra- and inter-process contradictions. While we recognize externalization as the general process by which people create material traces of their activities, we do not agree with AT and Engeström's focus on the externalization of object-motive pairs. For example, DCog highlights a number of mid-cognitive-process externalizations that are used to organize behavior on short time scales. Ultimately, we seek a definition of activity that recognizes production over an entire range of scales/granularities. Furthermore, ANT highlights the historical forces that apply to the material productions of behavior both through the notion of transformation and in the notion of punctualization, where entire networks start to be treated as singular entities. Yet another perspective on this production process comes from genre theory: the material productions of behavior do not exist simply in their actual instantiations; rather, they create abstract orderings and classifications around themselves. These genres of material productions are deeply social; they exist because many people are aware of and act according to them.

To summarize, the production of markers and mappings between markers and behavior (1) occurs at multiple granularities (in terms of time, space, and man-power), (2) are a deeply historical process whereby current production modifies and transforms past productions, and (3) both markers and mappings engender genres that enable people to treat them both in their particular and actual instantiations as well as in the abstract. In other words, production processes are subject to the duality of structure exhibited by the genres of markers and mappings. Genres are maintained by their ability to constrain what is produced.

Building on the idea of constrained production, the theories and models of behavior we surveyed highlight various constrained mechanisms and processes that link individual behavior to socially shared ideas and artifacts. The basic insight is that existing ideas and artifacts influence individual behavior. And, consequently, individual behavior often maintains the socially shared status of these ideas and artifacts. This is essentially the same idea as duality of structure. For Vygotsky, Engeström, and AT, this process is captured in the notions of internalization (which explains how socially shared ideas and artifacts get into a position where they can influence individual behavior) and externalization (which explains how individual behavior maintains the socially shared status of these ideas and artifacts). Importantly, for all of these theories, the internalization process is a social one. Hence, individuals often experience the social aspects of ideas and artifacts first-hand.

Similarly, Bateson and Goffman's notions of frames and Giddens's notion of structure can be aptly treated as (complex) social ideas/artifacts. Both frames and structures exhibit the duality of structure discussed above, and both frames and structures are internalized into an individual's routines/habits through social experiences. Consequently, when individuals perform their routines/habits, they are (implicitly or explicitly) re-creating/maintaining the status of the frames and structures they have internalized.

Individual behavior is connected to socially shared ideas and artifacts through the processes of internalization and externalization. Interestingly, the processes of internalization and externalization is often fraught with contradictions. Individuals do not smoothly, and without conflict, internalize and externalize these ideas and artifacts. Often, the internalization process is complicated by what the individual has already internalized and by how effective the social experience is at presenting the new ideas/artifacts[2]. Similarly, the externalization process can be complicated by competing demands (e.g., trying to produce something that satisfies two different object-motive pairs, frames, or structures). The potential for these processes to exhibit conflicts is acknowledge in ANT in two places. First, in ANT's explicit acknowledgement that any actor in a network can constrain the other actors in the network. And second, in ANT's notion of obligatory-passage-points, which illustrate how the material products of past activities can impact present and future behavior.

[2]Here the connection to learning, education and development is made clear in the "effective" internalization.

Extending the idea of constraint, the theories and models discussed above provide similar insights for the notion of appropriate activities. The basic idea is that the constraints that help to link individuals to the social world are the same constraints that map appropriate activities/behaviors to context markers. In other words, the appropriate activities/behaviors relative to context markers are internalized through actual social experience and then externalized (i.e., reproduced) to maintain the mapping between these activities/behaviors and their associated context markers.

Loosely speaking, appropriateness is determined by what has been done before (and has been internalized). It follows that appropriate activities/behaviors are: AT operations that responds to environmental conditions in a historically consistent way, AT actions that effectively service the subject's object-motive pair, AT activities that are centered on socially acknowledged and accepted object-motive pairs, behaviors that are mediated by a specific tool or sign in a historically consistent way, etc. Historical consistency based judgments are emphasized in Giddens's ideas on routines and routinization, and in Dewey and James' emphasis on habits.

Moreover, the theories and models discussed above highlight different types of constraints on which activities/behaviors are deemed appropriate. These types of constraints do not exist independently from one another, but rather are different aspects of the influence that past activities/behaviors have on present and future activities/behaviors. From AT, we acknowledge the influence of an individual's needs (which influence their adoption of object-motive pairs), the influence of an object-motive pair on goal formation, and the influence of environmental conditions on operations. From Vygotsky we acknowledge the influence of mediational mechanisms (tools and signs) on the behavior of people. From Engeström we acknowledge the influence of group dynamics, particularly conflicts and contradictions, and more external factors like existing rules of interaction, communities of stakeholders in the group activity, and the existence of roles and divisions of labor. From ANT we acknowledge the constraints that both human and non-human actors can have on one another and the use of punctualization to create obligatory-passage-points that constrain how different actors can behave. From DCog we acknowledge the creation of and response to markers of cognitive processing that are externalized and shared among members of a DCog network. From genre theory we acknowledge the constraints that existing genres place on their enactment and maintenance through actual instantiations. From the theory of grounded action we acknowledge the capacity for constraints to have different localized influences; and that these localized influences require both emic and etic orientations to elicit. From CA and DA we acknowledge the influence of accountability on behavior – the constraints that effective communication place on coordinated activities. And finally, from the theory of frames we acknowledge the boundaries that frames place on activities/behaviors, labeled superficially as within-frame or out-of-frame.

Complementary to the notion of appropriate behavior is the notion of interpretability. Obviously, assessing whether a behavior is appropriate or not is a form of interpretation. Furthermore, the ability to interpret behavior is rooted in the same capacities that enable one to assess whether the behavior is appropriate. The basic idea is that to interpret an activity/behavior, people need to situate that activity/behavior in their knowledge (conscious or tacit) of the world. People can interpret activities/behaviors that are similar to activity/behaviors they have already internalized. It follows that interpretation of behaviors is largely social, based on historical consistency, can occur at all granularities, and depends on both actual instantiations of behavior and on the genre classes that the behaviors are associated with.

Finally, we address the question of context markers. As with differing perspectives on constraints, the theories and models discussed above describe a number of different perspectives on what context markers are (and can be). Seeking to synthesize all of these perspectives, we acknowledge that context markers can be the environmental conditions that impact AT operations, the material traces of object-motive pairs, any actor in an ANT network, any externalized cognitive middle-step in a DCog network, any controllable difference in the world that impacts communication (including both linguistic and extra-linguistic signs), identifying attributes/elements of frames and structures, and the entities around which joint-activities are organized (whether they be an AT object or not). Generally speaking, a context marker is any perceivable difference in the state of one's self or one's environment that impacts their behavior – regardless of temporal or spatial scale, or whether it was human-generated or not.

2.4 Key Characteristics of Activity and Context

In previous sections of this chapter we discuss our definitions of activity and context using existing theories and models of human behavior. This discussion clarified all of the major points in our definitions. However, these descriptions are not easily translated into a computational model of activity and context. In this section, we distill the key characteristics of our definitions of activity and context. These characteristics are the basis for our computational model of activity and context.

Summarized as a list of key characteristics, we treat activity and context as:

observable – Both activity and context are observable as sets of coordinated markers (activities produce externalized, material traces of behavior; context includes these traces as well as naturally occurring markers). Furthermore, relevant markers (of activity and context) are not discretely separable from irrelevant markers (that identify neither activity nor context). Rather, markers have a continuous range of relevance, some being more indicative of, or influential on, activity and context than others.

inseparable – As discussed earlier, activity markers are all, to varying degree, context markers. Moreover, aside from naturally occuring markers, the degree to which a marker is treated as an activity marker, a context marker, or both depends on the perspective of the observer.

dynamic – The processes of internalization and externalization, which are deeply rooted in social interaction, result in the evolution of activity and context. Consequently, the markers that accurately identify a person's activity and context evolve.

multiplexed – While each activity and context can be analytically isolated, and the markers that strongly identify each activity and context separated, in actual behavior activities and contexts are multiplexed. Many activities can be worked on simultaneously with the same course of action, and many contexts can be simultaneously applied to create a complex and nuanced set of frames for interpretation.

ontologically rich – Both activity and context exhibit inter-structure relationships that generate genre-like classification systems. Furthermore, both activity and context exhibit intra-structure regularities that can be interpreted as typologies of roles. It is through this ontological diversity that entities like rules, community and division of labor (see the discussion of Engeström above) influence activity and context.

In the next chapter, we describe how these characteristics can be captured in a computational model.

Chapter 3

A Computational Model of Activity and Context

In the last chapter, we surveyed a fairly comprehensive set of theories and models of human behavior to better explain our definitions of activity and context. In this chapter, we translate these definitions into a computational model.

3.1 Defining the Variables of the Model

We begin by listing the key characteristics of activity and context outlined in the last chapter:

observable – Both activity and context are observable as sets of coordinated markers (activities produce externalized, material traces of behavior; context includes these traces as well as naturally occurring markers). Furthermore, relevant markers (of activity and context) are not discretely separable from irrelevant markers (that identify neither activity nor context). Rather, markers have a continuous range of relevance, some being more indicative of, or influential on, activity and context than others.

inseparable – As discussed earlier, activity markers are all, to varying degree, context markers. Moreover, aside from naturally occurring markers, the degree to which a marker is treated as an activity marker, a context marker, or both depends on the perspective of the observer.

dynamic – The processes of internalization and externalization, which are deeply rooted in social interaction, result in the evolution of activity and context. Consequently, the markers that accurately identify a person's activity and context evolve.

multiplexed – While each activity and context can be analytically isolated, and the markers that strongly identify each activity and context separated, in actual behavior activities and contexts are multiplexed. Many activities can be worked on simultaneously with the same course of action, and many contexts can be simultaneously applied to create a complex and nuanced set of frames for interpretation.

ontologically rich – Both activity and context exhibit inter-structure relationships that generate genre-like classification systems. Furthermore, both activity and context exhibit intra-structure regularities that can be interpreted as typologies of roles. It is through this ontological diversity that entities like rules, community and division of labor (see the discussion of Engeström above) influence activity and context.

Each of these characteristics highlights an important aspect of our computational model. Accordingly, we organize the description of our model by addressing each characteristic in turn. As our goal is to build a computational model, each characteristic might introduce variables and/or constraints on existing variables. We deal with the observable characteristic of activity and context first.

The ability to observe activity and context is a necessary prerequisite to modeling them. Until some technique or tool or perspective renders a phenomenon observable, one can only guess at its inner structure. Conveniently, both activity and context are observable through the same mechanism: the measurement/sensing of markers. The possible set of markers is at least as large as the set of measurable differences in the world; including, for example, what clothes people are wearing, what time of day the clothes are worn, where people live, what objects they have in their hands, what words they use, who they communicate with, what food they eat, when they eat, what documents they read, who wrote the documents, who else has read the same documents, etc. Clearly, the set of possible markers is infinite. However, in practice many of these markers are of such low relevance (i.e., they do not reliably identify any activity or context) that they can be ignored. While we do not assume that ignoring irrelevant markers will result in a finite set, we do assume that the set of markers of non-trivial relevance is countable. Accordingly, we let the natural numbers, $\{j \mid j = 1, 2, 3, \ldots\}$, denote the set of possible markers. Of course, no computational model cannot explicitly model an infinite number of markers. So, we maintain a finite, but growing set of markers: $\{j \mid j = 1, 2, 3, \ldots, M\}$. As new markers are encountered, M is appropriately increased.

Additionally, the second characteristic allows us to treat activity and context markers similarly – i.e., we do not need to index activity markers separately from context markers. An objection to this argument is the existence of naturally occurring context markers that are not, and could not be, activity markers – e.g., most astrological events, most meteorological events, most non-human biological/ecological events, etc. While we do not claim that these markers are ontologically equivalent to their human-producable counterparts, we find it unnecessary to distinguish these two classes of markers from a modeling perspective. Ultimately, what matters is that both classes of markers can be indicative of various activities *and* contexts, in spite of their ontological differences.

Similarly, we require indices for activities and contexts. The inseparability of activity and context, the second key characteristic, makes it reasonable to treat activity and context as part of a singular unit/structure. Treating activity and context in a single unit does not imply that we are unaware of, or trying to mask, the difference between these two concepts and the phenomena they signify. However, we assume that distinguishing between aspects of activity and context, given a set of markers, depends on the interpretive perspective being applied. Hence, from a computational modeling perspective, we prefer to maintain an agnostic interpretative perspective and treat activity and context as inseparable aspects of holistic structures. We refer to these structures as *activity/context units*, or ACUs. ACUs are distinguishable from one another based on the markers that are regularly produced or available when they occur. We index ACUs similarly to how we index markers, using the set $\{k \mid i = 1, 2, 3, \ldots, K\}$. And, as with the marker indices, we maintain a growing list of indices. As new ACUs are found in our computational model, K is appropriately increased.

Recall that markers can be more or less indicative of various ACUs. We define $Y_{kj} \in (0, L) \subset \mathbb{R}^{++}$ as the extent to which marker j indicates the occurrence of ACU k. Another interpretation of Y_{kj} is as the relevance weight of marker j to ACU k. To simplify future discussion, we refer to Y_{kj} simply as a *marker weight.* (Note: L is a theoretical upper bound on the potential value of Y_{kj}. For all practical purposes, we can assume L is significantly larger than any range of values we might encounter.)

It follows that every ACU, indexed by k, has an associated vector of marker weights: $\vec{Y}_k = (Y_{k1}, Y_{k2}, \ldots, Y_{kM})$. When new markers are encountered, and M is increased, we can maintain backwards compatibility of the marker weight vectors by treating the newly added Y_{kj} weights as having always existed with a value very close to zero. Similarly, as new ACUs are added to the model, we can assume they always existed and that their associated marker weights were all set to a value very close to zero.

We turn now to the third characteristic of activity and context: that they are dynamic structures that evolve. While the coupled processes of internalization and externalization are drivers of the evolution of both activity and context, we do not need to model these processes explicitly. Instead, we can focus on modeling the changes in the marker weights for every ACU in the model. These changes happen at two different levels. At the higher level, ACUs are created and abandoned/terminated. We can model these changes using the vectors of marker weights. Prior to being created or observed, an ACU will have a vector of marker weights all equal to some trivial value close to zero. After being observed, the associated vector of marker weights will contain some non-trivial, i.e., significantly greater than zero, values. The second level of evolution happens during the "lifetime" of an ACU and corresponds to shifts in how indicative markers are of the ACU. These shifts can be adequately modeled by allowing the marker weight values associated with each ACU to change.

In our model, the vector of marker weights for each ACU are allowed to change at discrete intervals (e.g., once per day). Changes in marker weights can, as noted above, handle both levels of ACU evolution. To ensure that ACUs maintain a recognizable amount of structure between changes, we bias the re-calculation of the marker weights on the existing weight values. More specifically, we define $\hat{Y}_{kj}$ as the existing marker weight of marker j in relation to ACU k. Without loss of generality, we can assume $\hat{Y}_{kj} \in (0,1] \subset \mathbb{R}^{++}$ for all k and j, and that $\sum_j \hat{Y}_{kj} = 1$ for all k. To handle the amount of bias between the existing marker weights for ACU k and the new marker weights, we utilize a parameter: $\alpha_k \in \mathbb{R}^{++}$. Using a probabilistic framework, we let:

$$\begin{aligned} P(\vec{\hat{Y}}_k \mid \vec{\overline{Y}}_k) &= (1/Z)exp(-\alpha_k KL(\vec{\hat{Y}}_k \,||\, \vec{\overline{Y}}_k)) \\ &= (1/Z)exp(-\alpha_k \textstyle\sum_{j=1}^{M} \hat{Y}_{kj} \log(\hat{Y}_{kj}/\overline{Y}_{kj})) \end{aligned} \tag{3.1}$$

where $KL(\cdot||\cdot)$ is the Kullback–Leibler divergence, Z is the necessary normalization constant to make this distribution integrate to one, and $\overline{Y}_{kj} = Y_{kj}/\sum_{\hat{j}} Y_{k\hat{j}}$. By allowing for different values of α_k, we can variably stress how similar the new marker weights, Y_{kj}, are to the existing marker weights, $\hat{Y}_{kj}$.

So far, we have addressed the structure of our ACU representations and how our model constrains the evolution of these representations. Multiplexity, the fourth characteristic of activity and context, however, forces us to address how ACUs occur in people's behavior: ACUs occur in overlapping mixtures. I.e., more than one ACU can occur simultaneously; creating a complex of (potentially conflicting) factors that influence people's behavior. Moreover, people often switch between many ongoing activities and contexts (*González and Mark* [2004]; *Mark et al.* [2005]) which is captured in different sets of markers being observed during different time segments. To model multiplexing and switching of activities and contexts, our computational model relies on an additive mixture of ACU representations to explain actual behavior.

Specifically, we assume that there exists a set of mixture weights, $\{X_k \in (0, L) \subset \mathbb{R}^{++} \mid k = 1, \ldots, K\}$ one for each ACU. Our computational model allows these mixture weights to change in time – i.e., which ACUs are influencing behavior change according to the mixture weights. To handle changes in the mixture weights, our computational model divides time into discrete segments. During each time segment, behavior is observed as counts of actual marker occurrences – the very markers that are indicative of, and relevant to, the ACUs that influenced the behavior. For each marker, we define f_{ij} as the number of times marker j was present/occurred during time interval i (from time t_i to time t_{i+1}); where $i \in \{1, \ldots, N\}$. (Note that for many markers and many time segments, f_{ij} equals zero.) To model the mixture of ACUs that gives rise to these observed marker occurrences in time interval i, we assume that a generative process exists that takes $(\sum_{k=1}^{K} X_{ik}Y_{kj})$, the weighted average of the marker weights for each ACU, as an input and produces the observed count f_{ij}.

Since both X_{ik} and Y_{kj} are in $\mathbb{R}^{++}$, it follows that $(\sum_{k=1}^{K} X_{ik}Y_{kj})$ is also in $\mathbb{R}^{++}$. To convert non-negative, real-valued parameters into non-negative, discrete-valued counts, we use the Poisson probability distribution where F_{ij} is the Poisson random variable that takes on the observed value f_{ij}:

$$P(F_{ij} = f_{ij} \mid X_{i1}, \ldots, X_{iK}, Y_{1j}, \ldots, Y_{Kj}) = \frac{(\sum_{k=1}^{K} X_{ik}Y_{kj})^{f_{ij}} exp(-\sum_{k=1}^{K} X_{ik}Y_{kj})}{f_{ij}!} \quad (3.2)$$

The final characteristic of activity and context, their ontological diversity, is a characteristic that we ignore in our current computational model. While notions like genres and roles are clearly important aspects of activity and context, they are higher-order structures. Roles generally capture regularities in the relationships between markers within an ACU and genres generally capture regularities between ACUs. In our computational model, we are primarily concerned with determining the degree to which markers are indicative of ACUs; as opposed to trying to model the regularities of marker-to-marker or ACU-to-ACU relationships. We note, however, that recent research has had success with modeling such high-order structures (e.g., *Clancey* [2002]; *Doane and Sohn* [2000]; *Hudson et al.* [1999]; *John et al.* [2002]; *Lemon and Gruenstein* [2004]; *Liao et al.* [2001]; *Patterson et al.* [2003]; *Perkowitz et al.* [2004]).

To handle the structure of roles effectively, one needs to segment ACUs into smaller, coherent chunks and then determine whether these chunks are repetitive (or, conversely, novel). Some techniques from Natural Language Processing might be effective: e.g., novelty detection (*Allan et al.* [2003]; *Kumaran and Allan* [2004]) and semantic chunking (*Hearst* [1994]; *Liu et al.* [2004]). To handle genres, one needs to look across people for shared aspects of each individuals' ACUs. A number of studies could inform this search (e.g., *Bardram* [1998]; *Chaiklin and Lave* [1993]; *Ducheneaut*

[2003]; *Engeström* [2000]; *Fjeld et al.* [2002]; *Goffman* [1959, 1967, 1974]; *Hill et al.* [2001]; *Hutchins* [1995]; *Hutchins and Klausen* [1996]; *Latour* [1986, 1995]; *Lave and Wenger* [1991]; *Law* [1992]; *Millen et al.* [2005]; *Mukherjee et al.* [2003]; *Muller et al.* [2004]; *Palen and Dourish* [2003]; *Shaw and Gaines* [1999]; *Wasserman and Faust* [1994]; *Wells* [2002]; *Yoshioka et al.* [2000]; *Yoshioka and Herman* [2000]; *Yoshioka et al.* [2001]).

3.2 Formal Model Description

With all of the variables and constraints (equations 3.1 and 3.2) defined, we can combine them into a unified framework using a simple Bayesian network that is graphically depicted in Figure 3.1. In this model, we assume that F_{ij} and $\vec{\hat{Y}}_k$ are observed random variables. We further assume that the values of α_k, $k \in \{1, \ldots, K\}$, are fixed a-priori. (Recall that α_k corresponds to the amount of bias we place on the current ACU representation, $\vec{\hat{Y}}_k$, when we calculate the new representation, $\vec{Y}_k$.) In our model, the unknown parameters are $\vec{X}_k$ and $\vec{Y}_k$.

There are two problems with the model depicted in Figure 3.1. First, it has the unfortunate risk of building a completely trivial set of ACU representations. Namely, if $K \geq M$, then

$$Y_{kj} = \begin{cases} 1 - \epsilon \text{ if } k = j \\ \epsilon / M \text{ else} \end{cases}$$

is a computationally optimal solution (with ϵ close to zero). However, it is devoid of any interesting structure – treating each marker as indicative of a different ACU. Second, it is possible for two or more ACU representations to become essentially indistinguishable. The model could then choose any possible mixture weights over

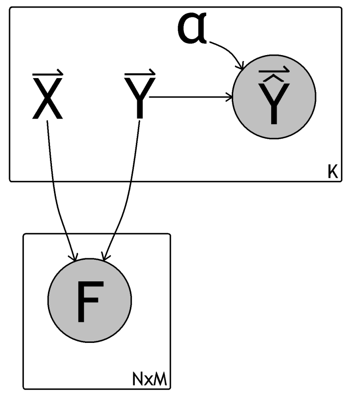

Figure 3.1. Standard Bayesian network depiction of our computational model of activity and context. $\vec{Y}_k$ is the vector of weights capturing how strongly each marker indicates the occurrence of ACU k. $\vec{X}_k$ is the vector of mixture weights, over all time segments, for ACU k. $\vec{\hat{Y}}_k$ is the previous marker weight vector for ACU k. α_k is the parameter determining the amount of bias placed on the new ACU representation, $\vec{Y}_k$, by the existing representation, $\vec{\hat{Y}}_k$. F_{ij} is the observed random variable that denotes the number of times marker j occurred (was observed) during time interval i.

these similar ACUs and still be equally likely to produce the same observed marker counts, f_{ij}.

We address the first risk in two ways. First, we constrain K to be smaller than M – i.e., the number of existing ACUs is forced to be smaller than the number of observed markers. Second, because our computational model does not have a global optimum, we initialize the computation to avoid the trivial ACU representation. We address the second risk by forcing ACU representations that are too similar to merge. We discuss our initialization and merging procedures in section 3.4.3.

3.3 Utilizing Our Computational Model

In this section we describe how our model of activity and context is utilized from a computational perspective. First, since all of the random variables are observed, we are only interested in parameter estimation (i.e., in determining the best values for X_{ik} and Y_{kj} – recall that we set the α_k parameter values a-priori[1]. For now, we walk through the general process of parameter estimation.

As with many applications of Bayesian networks, we use the maximum likelihood estimation process to determine optimal (in this case locally optimal) parameter values. Following standard methodology, we work with the log-likelihood and attempt to solve:

$$\begin{aligned}\arg\max_{X_{ik},Y_{kj}} \quad & \textstyle\sum_{i,j} \Big(f_{ij} \log(\sum_{k=1}^{K} X_{ik} Y_{kj}) - (\sum_{k=1}^{K} X_{ik} Y_{kj}) \Big) \\ & - \textstyle\sum_k \alpha_k \sum_j \Big(\hat{Y}_{kj} \log(\hat{Y}_{kj} / \overline{Y}_{kj}) \Big)\end{aligned} \tag{3.3}$$

where $\overline{Y}_{kj} = \frac{Y_{kj}}{\sum_{\hat{j}} Y_{k\hat{j}}}$. Removing constant terms, the final optimization problem can be written as:

$$\begin{aligned}\arg\max_{X_{ik},Y_{kj}} \quad & \textstyle\sum_{i,j} \Big(f_{ij} \log(\sum_{k=1}^{K} X_{ik} Y_{kj}) - (\sum_{k=1}^{K} X_{ik} Y_{kj}) \Big) \\ & + \textstyle\sum_k \alpha_k \Big(\big(\sum_j \hat{Y}_{kj} \log(Y_{kj}) \big) - \log(\sum_{\hat{j}} Y_{k\hat{j}}) \Big)\end{aligned} \tag{3.4}$$

To solve this problem, we derive iterative update equations in the same style as Lee and Seung (*Lee and Seung* [2001]). To do this, we employ two approximations. First, assuming $a, \hat{a}, b > 0$,

$$\begin{aligned} f_1(a) &\triangleq \log(a+b) \\ &= \log\Big(\tfrac{\hat{a}(\hat{a}+b)}{(\hat{a}+b)\hat{a}} a + \tfrac{b(\hat{a}+b)}{(\hat{a}+b)b} b \Big) \\ &\geq \tfrac{\hat{a}}{\hat{a}+b} \log\Big(\tfrac{\hat{a}+b}{\hat{a}} a \Big) + \tfrac{b}{\hat{a}+b} \log(\hat{a}+b) \\ &= \log(\hat{a}+b) + \tfrac{\hat{a}}{\hat{a}+b} \big(\log(a) - \log(\hat{a}) \big) \\ &= \tfrac{\hat{a}}{\hat{a}+b} \log(a) + \kappa_1 \\ &\triangleq g_1(a, \hat{a}) \end{aligned} \tag{3.5}$$

where κ_1 is a constant relative to a. Note that $f_1(\hat{a}) = g_1(\hat{a}, \hat{a})$. The second approxi-

[1] We discuss how the α_k parameters are set at the end of this section

mation is similar. Again, assuming $a, \hat{a}, b > 0$,

$$\begin{aligned} f_2(a) &\triangleq -\log(a+b) \\ &\geq -log(\hat{a}+b) - \frac{a-\hat{a}}{\hat{a}+b} \\ &= -\frac{a}{\hat{a}+b} + \kappa_2 \\ &\triangleq g_2(a, \hat{a}) \end{aligned} \tag{3.6}$$

where κ_2 is a constant relative to a. Again, note that $f_2(\hat{a}) = g_2(\hat{a}, \hat{a})$.

Both $g_1(\cdot, \cdot)$ and $g_2(\cdot, \cdot)$ are commonly referred to as auxiliary functions (*Lee and Seung* [2001]). They are useful because optimizing an auxiliary function simulatenously optimizes its associated cost function. To prove this, let $a^{t+1} = \arg\max_a g_1(a, a^t)$. The following inequalities hold: $f_1(a^{t+1}) \geq g_1(a^{t+1}, a^t) \geq g_1(a^t, a^t) = f_1(a^t)$.

Now we derive the iterative update equations that solve our maximum likelihood problem. First, consider X_{ik}. Removing terms that do not depend on this parameter, we can reduce the log likelihood to:

$$f(X_{ik}) \triangleq \sum_j \Big(f_{ij} \log(X_{ik}Y_{kj} + \sum_{\ell \neq k} X_{i\ell}Y_{\ell j}) - X_{ik}Y_{kj} \Big) \tag{3.7}$$

Using the approximations above, we can define an associated auxiliary function as:

$$\begin{aligned} g(X_{ik}, X_{ik}^t) &\triangleq \textstyle\sum_j \Big(f_{ij} \frac{X_{ik}^t Y_{kj}}{X_{ik}^t Y_{kj} + \sum_{\ell \neq k} X_{i\ell}Y_{\ell j}} \log(X_{ik}) - X_{ik}Y_{kj} \Big) + \kappa_t \\ &= \Big(\textstyle\sum_j \frac{f_{ij} X_{ik}^t Y_{kj}}{X_{ik}^t Y_{kj} + \sum_{\ell \neq k} X_{i\ell}Y_{\ell j}} \Big) \log(X_{ik}) - \big(\textstyle\sum_j Y_{kj} \big) X_{ik} + \kappa_t \end{aligned} \tag{3.8}$$

where κ_t is a constant relative to X_{ik}.

To optimize $g(X_{ik}, X_{ik}^t)$, we take the derivative relative to X_{ik} and set it to zero.

$$\Big(\sum_j \frac{f_{ij} X_{ik}^t Y_{kj}}{X_{ik}^t Y_{kj} + \sum_{\ell \neq k} X_{i\ell}Y_{\ell j}} \Big) \frac{1}{X_{ik}} - \Big(\sum_j Y_{kj} \Big) = 0 \tag{3.9}$$

Rearranging the terms, we arrive at the following iterative update equation

$$X_{ik}^{t+1} \leftarrow \Big(\sum_j \frac{f_{ij} X_{ik}^t Y_{kj}}{X_{ik}^t Y_{kj} + \sum_{\ell \neq k} X_{i\ell}Y_{\ell j}} \Big) / \Big(\sum_j Y_{kj} \Big) \tag{3.10}$$

Similar derivation produces the iterative update equation for Y_{kj}:

$$Y_{kj}^{t+1} \leftarrow \frac{\sum_i \frac{f_{ij} X_{ik} Y_{kj}^t}{X_{ik} Y_{kj}^t + \sum_{\ell \neq k} X_{i\ell} Y_{\ell j}} + \alpha_k \hat{Y}_{kj}}{\sum_i X_{ik} + \frac{\alpha_k}{Y_{kj}^t + \sum_{\hat{\jmath} \neq j} Y_{k\hat{\jmath}}}} \tag{3.11}$$

A simple algorithm for calculating the maximum likelihood estimates for each of the parameters in our model is:

- initialize X_{ik}^0 and Y_{kj}^0 for all i, k, and j
- loop until convergence:
    - for all i
        - for all k
            - compute X_{ik}^{t+1} using Y_{kj}^t and $X_{i\hat{k}}^t$ if $\hat{k} \geq k$ or $X_{i\hat{k}}^{t+1}$ if $\hat{k} < k$
    - for all j
        - for all k
            - compute Y_{kj}^{t+1} using X_{ik}^{t+1} and $Y_{\hat{k}j}^t$ if $\hat{k} \geq k$, $Y_{k\hat{\jmath}}^t$ if $\hat{\jmath} \geq j$, $Y_{\hat{k}j}^{t+1}$ if $\hat{k} < k$ or $Y_{k\hat{\jmath}}^{t+1}$ if $\hat{\jmath} < j$

Finally, we address how the α_k parameter values are set a-priori. From the Y_{kj} update equation, it is unclear how to semantically interpret what values would be appropriate for the α_k parameters – even in relation to the other parameters in the model. To facilitate this interpretation, it is useful to transform the α_k parameters using the following definition:

$$\alpha_k \triangleq \Big(\sum_i X_{ik}\Big)\beta_k \tag{3.12}$$

which changes the Y_{kj} update equation to:

$$Y_{kj}^{t+1} \leftarrow \frac{\left(\frac{\sum_i \frac{f_{ij} X_{ik} Y_{kj}^t}{X_{ik} Y_{kj}^t + \sum_{\ell \neq k} X_{i\ell} Y_{\ell j}}}{\sum_i X_{ik}} + \beta_k \hat{Y}_{kj}\right)}{\left(1 + \frac{\beta_k}{Y_{kj}^t + \sum_{\hat{\jmath} \neq j} Y_{k\hat{\jmath}}}\right)} \tag{3.13}$$

From this update equation, we can interpret β_k as the relative weight assigned to the current ACU representation, $\hat{\vec{Y}}_k$, found during the last model update. Empirically, we have found that our computational model can accurately track activities and contexts when β_k is proportional to the lifetime of an ACU (as measured in model update cycles – i.e., each time a new $\vec{Y}_k$ is calculated). In other words, more recently observed ACUs should have smaller associated β_k parameter values than ACUs that have been around longer. This method for setting the β_k (and α_k) parameters allows new ACU representations to change more, while also allowing older ACUs to maintain their representations.

3.4 Discussion

We divide the discussion of our computational model of activity and context into three subsection. First, we address the computational complexity of the model. Second, we discuss some additional theory on text analysis that further justifies the form of our model. Finally, we describe the actual application domain of our computational model. Specifically, we describe the markers available in our domain and how we collect the observed occurrences of these markers.

3.4.1 Computational Complexity

Our model of activity and context has some two beneficial computational features. First, the computational complexity of the update equations are driven by the sparseness in the observed random variables – i.e., computation can be omitted when $f_{ij} = 0$. Second, these update equations can be vectorized, utilizing cached, partial computations to speed up the overall computational costs.

Our model has a run-time complexity based on the amount of observed data and the number of iterations required for the parameter values to numerically converge. In every application of our model, we have found that a fixed number of iterations, typically between 5 and 20, is sufficient to reach numerical convergence. However, for the sake of completeness, we represent the number of iterations to reach convergence as T_c. We also need to represent the number of non-zero observations of f_{ij}. Accordingly, we set $T_f \triangleq \sum_{i=1}^{N} \sum_{j=1}^{M} 1_{(f_{ij}>0)}$. Note: in the worst case, T_f is $O(NM)$, but, in almost all of our applications of the model, T_f is often $O(N)$.

Overall, the run-time complexity of our model is $O(T_c K^2 T_f)$. So, in the worst case, the values of the parameters in our model can be calculated in $O(T_c K^2 NM)$. However, as noted above, when T_f is $O(N)$, the run-time complexity of our model is $O(T_c K^2 N)$.

In terms of space requirements, to represent all of the non-zero, observed random variable values and the parameter values, our model requires $O(T_f + NK + MK)$ memory. When the algorithm is vectorized by caching partial computations, the additional space requirements are $O(K + N + M)$, which is negligible relative to the cost of representing the parameters and variables in the model. Furthermore, additional space savings can be made by writing the variables to disk, and reading in only those values needed during each update step.

3.4.2 Additional Justification: Text Analysis

In spite of the theoretical justification presented above, modeling activity and context computationally might still appear unjustified. However, we claim that similar computational models have already been tested in a different domain: text analysis. The correspondence between ACUs and text-based topics/themes has been thoroughly explored in many academic disciplines. In particular, it is a deep meta-theme within the socio-historical school and in contemporary literary theory.[2] The notion of "language as symbolic action" is fundamental in Vygotskian psychology (*Vygotsky* [1985]; *Wertsch* [1985a, 1991, 1998]) and in the work of leading literary theorists (e.g., *Bakhtin* [1979]; *Burke* [1950]).

The notion of "language as symbolic action" provides yet another way to understand activity and context. Specifically, an ACU coordinating the behavior of specific people is analogous to a "theme" coordinating the words in a specific text or corpus. At a deeper level, the use of language and text is rooted in the physical and social experiences of people as they develop (*Lakoff and Johnson* [1980]; *Lave and Wenger* [1991]; *Vygotsky* [1985]). Consequently, text use bears the markers of these experiences which results in the natural categorization of language into speech genres (*Bakhtin* [1979]) and other praxis/use-based ontologies (e.g., *Atkinson and Heritage* [1984]; *Millen et al.* [2005]). It follows that processing texts necessarily results in the processing of various markers of human activity and context. Our claim is that effective text analysis relies on identifying (often implicitly) the activity and context markers in the text – e.g., the use of term-frequency-inverse-document-frequency (tfidf) to identify "meaningful" terms.

In fact, recent experiments support this connection. The GaP algorithm (*Canny* [2004]), which is closely related to our computational model, is one of the most accurate algorithms for modeling and smoothing text corpora. GaP also gives some of the most accurate results to date on standard keyword retrieval tasks. Naturally, adaptations of other topic extraction algorithms might be useful for activity/context analysis – e.g., Latent-Dirichlet Allocation (*Blei et al.* [2003]), Non-Negative Matrix Factorization (*Xu et al.* [2003]), and adaptive subspace iteration (*Li et al.* [2004]).

[2] Flavors of this connection have also been analyzed in Linguistics (*Searle* [1969]) and in discourse and conversation analysis (*Atkinson and Heritage* [1984]).

3.4.3 Applying our Computational Model: The Personal Computing Environment

All of the preceding discussion in this dissertation treats activity and context fairly generically. In this section we describe the particular domain where we apply our computational model of activity and context: the personal computing environment.

The personal computing environment can be generically understood as a single user/person connecting to various electronic resources (including information, tools/applications, or other people) using the traditional input devices of a mouse and keyboard. The electronic resources being accessed are processed by the computer and often, although not always, displayed visually on a monitor.

The activity and context markers in this environment include: running applications, application states (including display windows and CPU and memory usage), received and sent electronic messages to other people, folder and file accesses, web browsing/searching, the semantic content of different files and web pages, duration of information accesses, frequency of input (mouse, keyboard, etc), time of day, day of week, time of year, etc.

The motiviation behind choosing the personal computing environment is to address acknowledged problems faced by many people when they interact with their computers. Specifically, by building ACU representations in the personal computing environment, we hope to help people: (1) access relevant information easier and (2) reflect on, and hence pro-actively adapt, their behavior (at least the behavior that impacts their computer interactions). In the remaining chapters of this dissertation, we describe how we leverage the ACU representations in our model to address these objectives. In the remainder of this chapter, we describe in greater detail how we observe markers in the personal computing environment; and how these marker observations are managed by our computational model.

So how do we observe, and record, the marker occurrences on the personal computing environment? We have designed and built logging software that captures the following marker occurrences: file and folder accesses, application start and end times, the focus window of the computing environment (where the mouse and keyboard interaction is directed), sent email messages, and web browsing behavior (which web sites and when). For more discussion of the actual markers that are observed, see the system description section in the next chapter.

As noted above, the raw marker occurrence events are aggregated into segments of time. We experimented with a number of different segment durations, from 30 seconds up to 30 minutes, including multiple durations at once and both overlapping and non-overlapping segments. We found that the resulting ACU representations did not differ significantly based on the time segment durations. Ultimately, we settled on 30 second, non-overlapping time segments because it captures enough marker co-occurrence observations to build reasonable ACU representations while keeping the number of non-zero observations low (which affects the run-time complexity of our computational model).

Following the connection between ACUs and text-based topics or themes discussed above, it is reasonable to think of these logs of marker occurrence observations as text corpora where time segments correspond to documents and the markers observed during the time segment correspond to words in the document. Again, it is well known that "themes" or "topics" can be discovered from texts using appropriate algorithms (e.g., *Blei et al.* [2003]; *Canny* [2004]; *Xu et al.* [2003]). More importantly, these algorithms are known to be robust to actual document boundaries – i.e., they find reasonable themes and topics for a text corpora so long as reasonable length segments of text are preserved (e.g., sentences and paragraphs). This robustness characteristic might explain why we found little variation in the activity/context representations that resulted from varying the time segment durations.

The logging software that captures the time segments of marker observations writes these counts to a growing matrix that corresponds to f_{ij} elements. A number of house-keeping steps need to be performed just prior to running the model. First, we include only the last six (6) weeks of data. There are two reasons for restricting the observed data based on time. First, since the run-time complexity of our computational model is dependent on the number of non-zero observations, too much data could make the model practically intractable. Second, it is hard for people to maintain enough working memory of what they have been doing to reliably interpret ACU representations that are older than a few weeks.

The final preparatory work that needs to be done concerns the existing ACU representations, $\hat{Y}_{kj}$. Specifically, we normalize these values such that $\sum_{j=1}^{M} \hat{Y}_{kj} = 1$ for all k. We also use the life-time of each ACU representation, as measured in number of model updates, to set the α_k parameters. (Recall the α_k parameters were translated into β_k parameters to improve their interpretability.)

Now, with all the input data in place, we can describe the logistics of applying our computational model. Most importantly, our model is updated once per day, in the middle of the night. Each update calculates a new set of ACU representations using the observed data and the existing ACU representations. To allow for the creation of new ACU representations, we embed our computational model in a greedy search framework. Specifically, at the beginning of every update, we allow the model to start with many more potential ACU representations than it actually needs. To describe this process concretely, we need to introduce some additional notation. Let N' be the number of time segments used during the last model update; and N be the number of time segments used in the current model update ($N \geq N'$). Similarly, let K' be the number of ACU representations outputted by the last model update; and K be the number of ACU representations outputted in the current model update ($K \geq K'$).

The greedy search strategy starts by segmenting the time segments since the last model update (i.e., time segments $N'+1$ to N) into a active usage sessions – where an active usage session starts with a user-driven interaction with the computer (mouse or keyboard input) and ends at the last user-driven interaction before a sufficiently long period of without any user-interaction (in this case, we used two minutes as the minimum inactivity duration that marks the end of an active session). Other research has shown that such active sessions are frequent and often of short duration – the median active session duration is less than five minutes (*Rattenbury et al.* [2008]). Let N'' be the number of active sessions covering time segments $N'+1$ to N. We initialize the new model update, given the old model update, using the following equations:

$$K = K' + N'' \tag{3.14}$$

$$Y_{kj} = \begin{cases} \hat{Y}_{kj} \text{ if } k \leq K' \\ \sum_{i \in \mathbb{T}_\ell} f_{ij} / |\mathbb{T}_\ell| \text{ if } k = K' + \ell \end{cases} \tag{3.15}$$

where $\ell \in \{1 \ldots N''\}$ indexes the active sessions covering time segments $N' + 1$ to N, and $\mathbb{T}_\ell \subset \{N' + 1 \ldots N\}$ is the set of time segment indices corresponding to the ℓ-th active session. After the initialization is complete, at the end of each of the outer-most loops of the algorithm, we check whether $\overrightarrow{X}_k$ is "too similar" to $\overrightarrow{X}_{\hat{k}}$ or whether $\overrightarrow{Y}_k$ is "too similar" to $\overrightarrow{Y}_{\hat{k}}$, for every $k \neq \hat{k}$. If the vectors are too similar to one another, we merge the vectors together by averaging their values, storing these values in one of them, and then we delete the other. Similarity between vectors can be measured in a number of ways. We tried the standard Euclidean cross-product, the Jensen-Shannon distance, and both of these metrics/distances with the vectors converted to $\{0, 1\}$ using a simple thresholding scheme. Empirically, the best results were produced when the vectors were converted to $\{0, 1\}$ and compared using the Jensen-Shannon distance.

By including this greedy search strategy, we change the run-time complexity from $O(T_c K^2 T_f)$ to $O(T_c K^2 (T_f + NK + MK))$. In practice, the greedy search strategy removes most of the newly added ACU representations in the first 10 outer-loop iterations.

In the next chapter, we discuss the primary application of our computational activity and context model in the personal computing environment. An important distinction to highlight is the between *offline* and *online* model updates. The model update described immediately above is what we refer to as offline. The primary objective of the offline model update is to calculate the ACU representation vectors. Alternatively, the online model update assumes a fixed set of ACU representations. The online model update calculates, in real-time, the mixture weights that best explain which ACUs are occurring (i.e., which ACUs that a person is working on/involved in). These real-time mixture weights are used in the application, described below, to provide immediate feedback to the user about what he/she is doing.

3.5 Related Work

The related work for our model of activity and context falls in two categories. The first category includes models that have a similar mathematical or algorithmic structure. The primary objective in discussing this category of related work is to highlight and further justify the mathematical structure of our model. The second category of related work includes models that, to varying degrees, attempt to capture activity and/or context. We discuss work in this category to highlight the novel aspects of our model.

3.5.1 Mathematically and Algorithmically Similar Models

If we look at the relation between F_{ij}, X_{ik} and Y_{kj} we see that our computational model can be thought of as a modified non-negative matrix factorization (NMF) problem where $F_{ij} \approx \sum_{k=1}^{K} X_{ik} Y_{kj}$. The inclusion of $\hat{Y}_{kj}$ significantly alters the relationship between our model and the standard NMF model presented by Lee and Seung (*Lee and Seung* [2001]). Other extensions to the standard NMF formulation attempt to regularize the computation to produce sparse or independent vectors (e.g., *Hoyer* [2004]).

The modeling of F_{ij} as a Poisson random variable has clear connections with the Gamma-Poisson (GaP) model (*Canny* [2004]). The major differences between GaP and our model of activity and context are (1) the inclusion of $\hat{Y}_{kj}$ and (2) the lack of a Gamma prior distribution on X_{ik}. In practice, the difference in values when X_{ik} is treated as a random variable, with a Gamma prior, or when X_{ik} is treated simply as a parameter are negligible provided N, the number of time segments of observed data, is sufficiently large. Moreover, dropping the Gamma prior simplifies the model by decreasing the number of parameters that need to be a-priori or empirically set. Buntine and Jakulin provide a detailed discussion of the relationships between GaP and other computational models (*Buntine and Jakulin* [2005]).

An important aspect of our computational model is that all of the random variables and parameters are assumed to be non-negative or positive. We explicitly avoided algorithms and methods that result in negative mixture weights – e.g., Latent Semantic Analysis (*Landauer et al.* [1998]) and eigen decomposition (*Eagle* [2005]). In the theories and models of human behavior that we discussed in chapter 2, the notions of negative influence by a marker pertain more to the appropriateness of certain, labeled activities or contexts (e.g., *Goffman* [1959, 1967, 1974]). However, it is unclear how one can passively observe these negative influences. Rather, it is often through some form of trial and error that people come to learn about these negative constraints (*Bateson* [1972]). The complexity of building a computational model that could actively test these relationships is beyond the scope of this dissertation. Accordingly, we formulated a model of activity and context that relies solely on positive relationships.

One final point to note is that our model and related mathematical models are all locally optimal (as opposed to being globally optimal). Our ultimate concern with the interpretability of the ACU representations renders this point moot. It is of little concern whether an ACU representation is globally optimal. What matters is whether or not the ACU representation is interpretable.

3.5.2 Models of Activity and/or Context

There are many algorithms and systems that utilize action (task) based information to maintain models about the user, their activity, and/or the state of the world. Most of these systems and algorithms are published in the User Modeling or Artificial Intelligence communities. We briefly discuss some of these systems. For a shorter summary of user modeling in HCI applications, with some discussion of current challenges, see Fischer (*Fischer* [2001]). For a brief summary of user modeling research see Kobsa (*Kobsa* [2001]). For a survey of recent application of Machine Learning and Natural Language Processing to user modeling problems see Webb et al and Zukerman et al (*Webb et al.* [2001]; *Zukerman and Litman* [2001]; *Zukerman and Albrecht* [2001]).

The related work on modeling activity and/or context falls roughly into four categories:

1. systems that model behavior as a dynamic distribution over words, commands,

or word senses (*Billsus and Pazzani* [1999, 2000]; *Bueno and David* [2001]; *Cheng* [2000]; *Lieberman* [1995]; *Linton et al.* [1999]; *Magnini and Strapparava* [2001]; *Teo* [2001])

2. systems that model action/task goals, usually in very restricted domains (*Christensen and Bardram* [2002]; *Clancey* [2002]; *Crowley et al.* [2002]; *Eliassi-rad and Shavlik* [2003]; *Faaborg and Lieberman* [2006]; *Fleming and Cohen* [2001]; *Franklin et al.* [2002]; *Horvitz et al.* [1998]; *Horvitz and Paek* [1999])

3. systems that model learning trajectories, generally for tutoring support (*Bull et al.* [2001]; *Bunt and Conati* [2003])

4. systems that model behavior by categorizing the user's actions (*Begole et al.* [2002]; *Oliver et al.* [2002]; *Liao et al.* [2001]; *Patterson et al.* [2003])

We briefly discuss each of these related research projects.

Billsus and Pazzani (*Billsus and Pazzani* [1999, 2000]) describe a news story classification system that determines whether or not a new news story would be of interest to the user. Their system maintains both a short-term and long-term model of the user's news interests. The short-term model generally looks for novel articles (that are sufficiently different from recently viewed articles) while the long-term model maintains a list of keywords that are of interest to the user.

The Metiore system (*Bueno and David* [2001]) maintains a model of a user's information uses by explicitly having the user group information retrieval (IR) queries into long-term information need groups. For each information need group, the system stores a list of attribute keywords and how they relate to this information need (e.g., a user could point out that all of the relevant documents to a need contained a certain keyword). Results from a new IR query can then be sorted based on the attributes' relations to the user's explicitly declared information need.

Cheng's Knowledgescapes system (*Cheng* [2000]) attempts to model tacit knowledge in work groups. Like our model, Cheng's model attempts to discover implicit coordinating factors that drive behavior – where we rely on the notion of ACUs, Cheng relies on the notion of information needs. However, unlike our model, the Knowledgescapes system heuristically assumes that a user's information need changes after a sufficiently long period of inactivity or anytime the user submits a different IR query. It also assumes that only one information need at a time can impact the user's behavior.

Letizia (*Lieberman* [1995]) is a program that builds a model of a user's information interests by logging web-browsing behavior. The model of the user's interests is short term (i.e., at the AT action/task level as opposed to the AT activity level) and consists of a list of keywords. Given a user interest model, Letizia suggests a priority ordering of the links on the current web-page being browsed.

Linton et al. build models of user expertise in Microsoft Word (*Linton et al.* [1999]). Relying on both automatically gathered log data (recording things like "save", "copy", and "paste") and "naturalistic inquiry" – similar to the qualitative methods in Contextual Inquiry (*Beyer and Holtzblatt* [1998]) – they build distributions over commands that span the scale between novice and expert users. Using these models, they claim to improve learning situations by instructing users on which commands they should learn to use.

Magnini and Strapparava (*Magnini and Strapparava* [2001]) describe a system similar to those discussed above. However, instead of storing keywords in their user model, they store word senses as presented by MultiWordNet (*Pianta et al.* [2002]). The objective of their system is to build a language independent user model so that news stories from multiple languages can be sorted for relevancy according to the user's interests.

Teo's system models the user's interests by parsing their textual ouput, specifically their personal web pages (*Teo* [2001]). The system then measures the informational overlap between documents and the model of the user's interests. Documents with high overlap are presented to the user for review.

Christensen and Bardram (*Christensen and Bardram* [2002]) work explicitly with AT's conceptualization of human behavior. Their system, built for hospital personnel, seeks to efficiently organize and provide information about current patients and their required services. Although this system has proven utility, it has some shortcomings. First, the activities that this system can handle are all pre-specified and hard coded into the software. Therefore, the system cannot detect an activity that some employee has not previously entered into the database. Thus it disregards improvisational, spontaneous, or coping activities which are pervasive in knowledge work (*Bardram* [1998]; *Engeström* [2000]). Secondly, the system is entirely query driven. Hospital employees must actively engage with the system to receive and benefit.

Crowley et al. describe a system for gathering data on human behavior and using it to model "activity" (*Crowley et al.* [2002]). They specifically discuss an application of their system to multimedia presentations. Utilizing position and sound sensors to track users and recognize their actions, the system maintains a computational representation of a user's presentation "activity" – e.g., whether they are talking or writing/drawing something to show their audience. In their model however, the ordering of steps is critical since their temporal patterns are captured in Finite-State Machines. Thus their approach is actually much closer to traditional task modeling in AI and HCI, than the activity and context model we have defined.

Like Crowley et al., Franklin et al. focus on modeling people making presentations (*Franklin et al.* [2002]). Their system gathers data on people making presentations, and attempts to reconcile this data with preset "plans" (like AT actions) modeled as ensembles of finite state automata. The data points are observations providing the automata states, which roughly correspond to AT operations. As with Crowley et al., this work is actually much closer to traditional task modeling in AI and HCI, than the activity and context model we have defined.

Eliassi-rad and Shavlik (*Eliassi-rad and Shavlik* [2003]) describe a general system architecture for building software agents that help with information retrieval and extraction. The user provides some initial if-then rule which the agent uses to generate its own training examples. They envision users running multiple agents simultaneously. However, it is up to the user to invoke the appropriate agents given their current IR needs. Depending on how how the user's interests change over time, setting up and invoking agents might be more work than simply searching from scratch.

Faaborg and Lieberman (*Faaborg and Lieberman* [2006]) built a system that detects goal-oriented behavior during web browsing. For example, a user can create a simple program that copies recipies from a recipe web page and store them in a local cookbook. The program is created by example: after the user has performed the copy operations a few times, the system detects this copy pattern and stores it for future use. To detect that only recipes should be copied, Faaborg and Lieberman's system relies on the knowledge repositories that can both identify the recipe and provide a short list of possible uses for the recipe (e.g., copying to a local cookbook or storing the ingredients in a shopping list). The system combines the ConceptNet (developed at MIT) and TAP (The Alpiri Project – developed at Stanford) knowledge repositories.

Fleming and Cohen apply user modeling to mixed-initiative interfaces (*Fleming and Cohen* [2001]). Their model of the user includes the following items: a simple model of the user's knowledge, a measure of the user's willingness to be interrupted, and a measure of how interpretable the system's current request will be for the user.

These three items are combined to create a dynamic utility measure used to determine whether the system should request the user's help in addressing some current problem/issue.

Clancey's Brahms system (*Clancey* [2002]) models people's "workframes". Workframes are related to Schank and Abelson's "scripts" in natural language processing and Barker's "behavioral settings" with their associated "action patterns" from sociology. Clancey's primary concern is modeling the stereotypical actions in a given setting. Although the Brahms system can provide models of the actual dynamics of how actions and operations are carried out, it does not provide support for learning these dynamics directly from observed data. The "workframes" in Brahms need to be manually created.

The Lumière project (*Horvitz et al.* [1998]) anticipates a user's goals and information needs within the context of using Microsoft Excel. Although they explicitly target actions and goals, they also maintain a long-term user profile which they use to determine a user's level of expertise, and hence what types of goals and needs the user might have.

Similar to the Lumière project is Horvitz and Paek's "computational architecture for conversation" (*Horvitz and Paek* [1999]). Their system acts as an office receptionist. It attempts to infer people's goals (e.g., what information they are trying to learn from the receptionist); and then support the satisfaction of these goals. Their underlying architecture is a pre-specified goal hierarchy.

Bull et al.'s research on the I-help system uses simple models of students and tutors to improve electronically mediated learning environments (*Bull et al.* [2001]). Each person known to the system has an associated agent which maintains a dynamic model of what the person knows or needs to learn and which other people this person likes to work with – these models are built from self- and peer-reports. The agents then negotiate the assignment of tutors to students to maximally meet everyone's needs.

Bunt and Conati use Bayesian networks to model students as they learn various mathematical concepts (*Bunt and Conati* [2003]). Their objective is to automatically provide targeted exercises designed to expose and correct students' misunderstandings. The student models are scaffolded from manually generated curricula and specifications of correct mathematical reasoning.

The Work Rhythms system (*Begole et al.* [2002]) is a desktop system that attempts to model whether a user is busy or not. Using this binary distinction, Begole et al. claim that they can improve office interactions. By logging low-level sensor events like keyboard and mouse movement and combining that with known scheduling information from calendars and task managers, they attempt to predict effective times for people to have scheduled or impromptu meetings.

Like the Work Rhythms system, Oliver et al. describe an architecture designed to model whether people are busy or not (*Oliver et al.* [2002]). The difference is that they break their "busy" state into pre-specified behaviors like communicating-face-to-face, presenting and communicating-remotely. They infer these states from low-level data (like sound volume, movement in surveillance images, key and mouse use) using layered Hidden Markov Models.

Like Oliver et al., Liao et al. and Patterson et al. attempt to infer higher level behaviors from low-level data (*Liao et al.* [2001]; *Patterson et al.* [2003]). In this case, their data is GPS coordinates. From this data they successfully infer modes of transportation (e.g., walking or riding the bus) as well as destinations (e.g., heading home or heading to work) and their durations. Their models are dynamical Bayesian networks.

3.6 Summary

Applying our definitions of activity and context in a computational model is a distinguishing feature of this dissertation. Traditionally, theories and models of activity and context have been applied descriptively by social scientists who identify the elements of activities and context from qualitative observation (e.g., *Chaiklin and Lave* [1993]; *González and Mark* [2004]; *Mark et al.* [2005]; *Engeström et al.* [1999b]; *Engeström* [1999, 2000]; *Nardi* [1996]; *Suchman et al.* [1999]). Instead, we attempt to automatically infer first-order statistical patterns in observations of human behavior that correspond to ACUs. More specifically, our computational model is a generative probabilistic model of human behavior that infers which markers belong together in different ACUs. This generative computational model automatically detects the markers that map to, compose, and are indicative of a person's ACUs.

Related computational models are either not applied to the domain of activity and context, or they attempt to define a-priori what markers are important for a specific set of activities and contexts. Our computational model differs from these related approaches by modeling emergent (not a-priori specified) activities and contexts in a single, generic framework.

Chapter 4

CAAD: the Context-Aware Activity Display

In this chapter we describe the end-user application driven by our computational model of activity and context. The application is a personal awareness display. The content of the display is the set of ACU representations automatically discovered by our computational model. We refer to this application as the Context-Aware Activity Display (or CAAD). CAAD supports users in two ways: First, it enables more efficient identification and access of relevant information. Second, it provides a feedback mechanism for users to reflect on the activities and contexts they are involved in through their computer.

4.1 Introduction

Recent HCI research has shown strong interest in tools for information workers (people whose primary work function is to create, share, and analyze information). As many studies have shown, information work is characterized by multiple ongoing, often disjoint, activities (*Bellotti et al.* [2004]; *Czerwinski et al.* [2004]; *González and Mark* [2005]). It follows that many problems that arise in the day-to-day lives of information workers relate to activity management (at both the intra- or inter-activity levels). Applications in traditional computing environments provide poor support for information workers because they are unresponsive to the dynamic nature of activity management (*Fogarty et al.* [2005b, a]; *Morteo et al.* [2004]). Recent attempts to improve support for activity management have focused on information management through representations of ongoing activities (e.g., *Bardram et al.* [2006]; *Kaptelinin* [2003]; *Moran* [2005]; *Morteo et al.* [2004]; *Muller et al.* [2004]; *Robertson et al.* [2000]; *Smith et al.* [2003]; *Stumpf et al.* [2005]; *Voida et al.* [2002]). However, most of these systems suffer from one or both of the following issues: (1) they require too much overhead on the part of the user (e.g., expecting users to generate and maintain representations of their activities); (2) they lack contextual awareness of the user's activities (e.g., what information is relevant). In this chapter, we present a novel approach to personal activity management that addresses these issues by automating the creation and maintenance of activity (more specifically, ACU) representations.

Our approach to personal activity management is implemented in a system we call CAAD (Context-Aware Activity Display). CAAD minimizes user overhead by automatically gathering marker occurrence information and building ACU representations that correspond to the user's ongoing activities. As noted in the previous chapter, the marker information gathered by the system primarily consists of computer interaction events: e.g., the use of a file, the browsing of a web page, or the execution of an application. Once per day, CAAD performs an offline model update that builds ACU representations from the user's actual work-flow.

Most importantly, the awareness display in CAAD leverages its context-awareness to support the user in two ways. First, it makes real-time (every 30 seconds) predictions on what information (documents, web pages, folders, etc.) is most relevant to the user. These predictions are an online calculation that leverages the offline calculation of the user's ACU representations. The higher the predicted relevance of an information item, the more prominently it is displayed (Figure 4.1). By increasing visual prominence, CAAD makes relevant information easier to access. Second, the awareness display in CAAD presents information in groups that explicitly reveal the various ACU representations that have been automatically inferred from the observed behavior of the user. These ACU representations are not always perfectly aligned to the user's perception of his/her activities; however, they do have enough meaningful overlap to be useful to users as they manage their ongoing activities. Specifically, by displaying ACU information, CAAD provides a mechanism for users to become more reflective about the organization of their work behavior. Users can also edit the ACU representations to better align CAAD's model of their work activities with their own.

4.2 Motivation

In this section we introduce two motivating scenarios. These scenarios are inspired by dialogue from informal interviews with information workers, conducted during our literature review and prior to the building of CAAD. These scenarios both ground what we mean by "support computer-based information workers" and highlight CAAD's actual functionality.

1. Information access scenario. Kate is a knowledge worker writing the related work section of a project report. While writing she realizes that she needs to cite a paper she read at the start of the project, a few weeks earlier. However, she cannot remember the title or author of the paper, or when she read the paper.

2. Work awareness scenario. Paul is a project manager coordinating multiple projects. To strengthen each project, he has been actively mediating dialogue between

members of different project teams. To assess which mediation strategies are most effective, Paul conducts brief interviews with members of each project team. However, he would prefer a lighter-weight and more accurate measure of inter-team impact.

In both of these scenarios, the overhead and lack of context issues are key concerns. In the information access scenario, the overhead of having to perform an ill-specified search will likely result in Kate loosing the mental context and flow of writing. Moreover, the success of this search depends on Kate's ability to translate the context of her past usage of the document into concrete query terms or search constraints. If Kate had been using CAAD since the start of her project, then all of the documents that are contextually relevant to her project report would be prominently displayed (Figure 4.1). Through the display, Kate could list the relevant documents in the ACU representation(s) associated with her project; likely recognizing the one she was looking for from its title.

In the work awareness scenario, Paul currently relies on interviews for gathering data. These interviews create overhead for Paul and for the project members who must suspend their normal work flow to be interviewed. With the appropriate context data around each project, updated as the projects progress and evolve, Paul could rely on simple similarity measures to detect whether projects are influencing one another or not. For example, if each project member was using CAAD, Paul could measure the number of relevant documents that were shared between projects before he intervened as well as after. If his mediation was effective, the number of shared documents should have increased.

These scenarios indicate the type of support we expect CAAD to provide: it should maintain contextual awareness of the user's various activities and contexts and use this awareness to (1) minimize the amount of overhead in accessing relevant information and (2) track and reveal the state of the user's various ACUs.

4.3 CAAD System Description

CAAD's design objective is to support computer-based information workers with a minimum amount of interaction overhead. To meet this objective, CAAD coordinates three components: (1) a logging component that automatically captures computer interaction events (i.e., activity and context markers), (2) a computational model component that performs offline and online calculations that detect and track the user's evolving activities by maintaining dynamic ACU representations, and (3) an awareness display that presents the user's ACU representations in a direct manipulation user interface.

4.3.1 Logging Component

The logging component gathers evidence of activity and context marker occurrences on and through the computer. This evidence consists of interaction events like file access and modification, email transmission, application use and state, and web browsing activity. Many related systems capture similar activity and context markers (e.g., *Kaptelinin* [2003]; *Stumpf et al.* [2005]; *Voida et al.* [2002]). Most of these systems target Microsoft Windows users and rely on various hooks into the user input stream or the Component Object Model (COM) interface to capture marker occurrence events. They are "push" architectures – events are pushed at the logging system. Alternatively, the logging component in CAAD uses a "pull" architecture. It periodically checks for relevant marker occurrences.

The decision to pull information stems from our interest in specific types of marker occurrences. All of the marker occurrences that are logged can be described as "using X" where X can be applications, files, folders, web pages, and email addresses. If our logger was only receiving push events like "file Y was opened" and "file Y was closed", it would need to maintain state variables to determine "using Y". These state variables would be highly sensitive to missed events, requiring potentially sophisticated back-up mechanisms.

Files that are logged can be of many origins and types. Specifically, files on local and network drives, web pages, email attachments, email subject lines (as short text documents), and email body texts are logged. Additionally, file use produces two marker occurrence events: one with the file path name and another with an md5 (message digest algorithm 5) hash of the file contents. By creating two events, we can track changes on a single file (same file path name, different md5 hashes) as well as file moves (different file path name, same md5 hash). Email-related marker occurrence events are restricted to outgoing email because incoming email has limited temporal correlation to work-flow at the second or minute time-scales. (We originally included incoming email, but the resulting ACU representations were often inaccurate due to "out-of-context" messages.)

Application-based marker occurrences are logged using several redundant pieces of information. First, the logger tracks active windows. This is necessary because many computer users leave multiple applications and windows open, even if they are not being used. The list of active windows generates a list of active applications. We also employ a time-out after 30 seconds – if an application has not been the active window in the last 30 seconds it ceases to generate use markers (i.e., is no longer logged). This list is cross-checked against the list of running applications. Applications that are both active and still running are logged. Finally, if the active window has not changed in five minutes and the user has not moved the mouse or pressed any keys, the logger assumes the user is taking a break and nothing is logged.

As described earlier, ACU representations are sets of relevant markers which correspond to various forms of information in CAAD: files, folders, web pages, applications, and email addresses. To accurately infer ACU representations, CAAD must know when, and for how long, each marker occurred – i.e., when, and for how long, each information item was accessed/used. In line with this, the logging component polls once every 2 seconds, depending on CPU load. Marker occurrence events at shorter time scales are missed. However, preliminary experiments did not reveal any events that would justify polling at a higher rate. The logging component itself averages to about 2% CPU load on a 2 GHz Pentium 4 with 1GB of RAM (spiking during the calculation of file hashes).

4.3.2 Computational Model

The computational model in CAAD detects ACU representations in logs of computer interaction events. Chapter 3 describes this model in detail; including a discussion of how it is applied in the personal computing environment. We will not repeat this discussion here. Rather, we focus on the distinction between offline and online model updates.

Recall that the computational model stores the user's ACU representations in a non-negative matrix, $[Y_{kj}]$. Each ACU representation corresponds to a row of this matrix, $\vec{Y}_k$.

ACU representations are calculated offline – once per day – with the most recent 6 weeks of logged marker observations. On average, with 6 weeks of data, the algorithm takes between 20 and 45 minutes to run an offline update on a 2 GHz Pentium 4 with 1GB of RAM. CAAD is currently configured to perform this update in the middle of the night. However, it is reasonable to run the update during a lunch break or a meeting if overnight updates are not feasible.

An important characteristic of the offline calculation is that it tracks the evolution of ACU representations. This is important because activities and contexts change and evolve, requiring any activity management system to handle these changes.

CAAD's model of the user's ACU representations remain fixed between daily updates. Between these daily updates, the fixed ACU representations are used to make online, real-time, predictions on what the user is doing – and hence what information is relevant to him or her. The online updates basically calculate the likelihood of each ACU occurring (or being worked on) during the last 30 seconds. Online updates require minimal memory and computational resources – they are negligible in comparison to the requirements of the logging component.

4.3.3 Awareness Display

The awareness display in CAAD presents the user's ACU representations, which are detected and maintained by the other components of CAAD. The display is dynamically configured according to online, real-time predictions on which ACUs are currently relevant to the user. Predictions are made every 30 seconds. The only component of CAAD that the user can directly interact with is the activity display.

The display supports users in two ways. First, it acts as a portal through which users can access information relevant to their various activities and contexts. Second, it provides a mechanism for users to reflect on the organization of their work behavior. The display provides this support in a context-aware way by leveraging the real-time predictions of what is relevant to the user. The predicted relevance of an information item determines its size in the display – the most relevant information items are the largest elements in the display and hence easiest to access.

The information in the display, as discussed above, represents the markers of the user's activities and contexts, as organized by their work behavior. By providing this information, the activity display enables user reflection on how their day-to-day

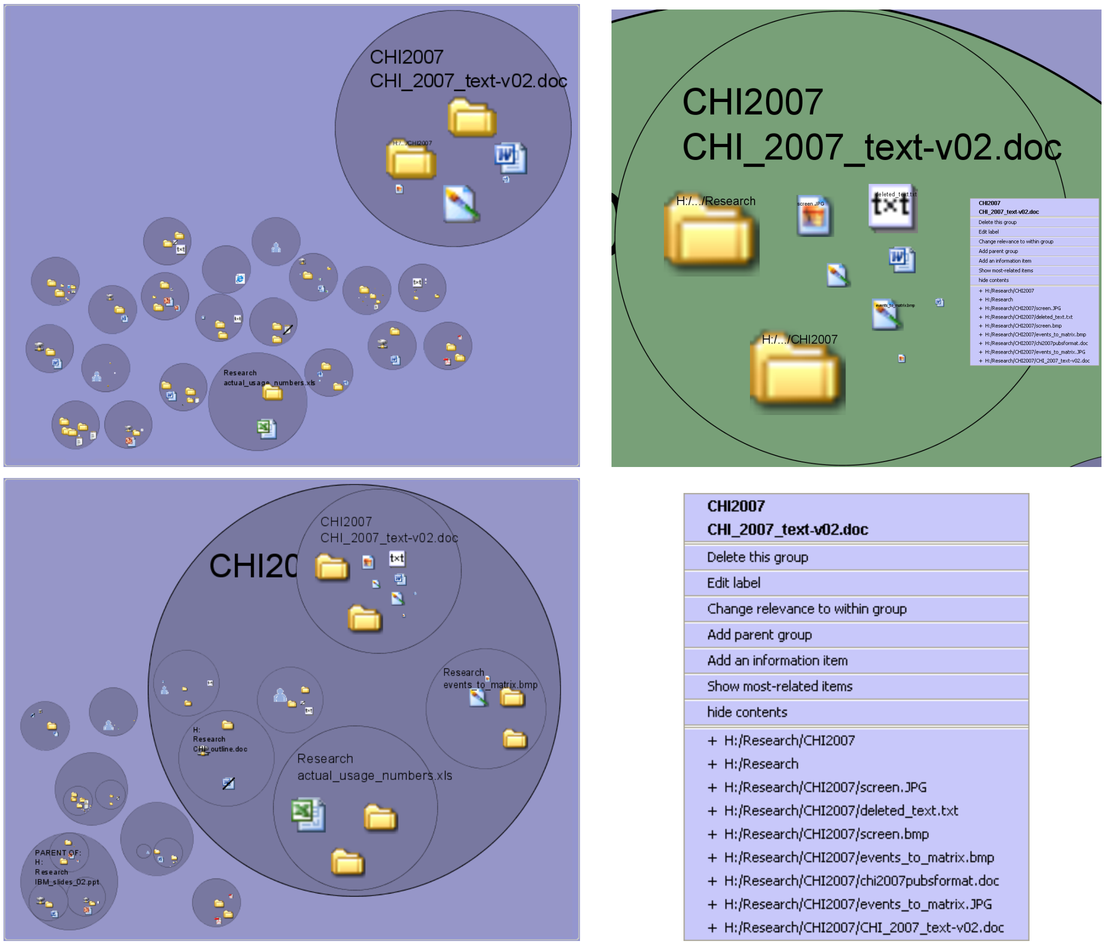

Figure 4.1. Four screen shots of the activity display. The top-left image was automatically generated by CAAD. The bottom-left image illustrates a user-edited display. The top-right image shows one ACU representation in more detail and the bottom-right image shows the context menu, with content list, for the ACU representation in the top-right image.

work routines are organized. This reflection could be superficial (e.g., answering "What have I been working on lately?") or more profound (e.g., realizing that "I'm pretty distracted at work. I'm always looking at sports news web pages."). Although the activity display can support these types of reflection, the user must make some inferences to do so – for example observing that there are sports news web pages in every ACU representation in the display and interpreting what this means about how his/her work behavior is organized.

Figure 4.1 shows two, full-screen captures. The top-left image is from a default configured (i.e., no edits have been performed) display. The bottom-left image is from a user-edited display. Each ACU is represented as a circle in the display. Each ACU representation (circle in the display) contains icons representing relevant files, folders, web pages and people. We refer to these icons, and the things they represent, as information items. Parent circles, which contain the ACUs, are shown in the bottom-left image in Figure 4.1. Currently, parent circles are not automatically generated; users must manually create them.

By default, the display only shows information items and ACU representations. Again, the ACU representations are updated once per day during the offline calculation. In the display, only the most relevant information items are shown. The threshold that determines which information items to show is dependent on the most relevant information item for that ACU representation. Any items with a relevance weight greater than 5% of the weight of the most relevant information item are shown (the sorted relevance weights for each ACU representation fall-off steeply around the 10% point). In addition to this threshold-based scheme, another threshold scheme is used to decide which ACU representations from the computational model to display. Specifically, ACUs that occurred in less than 1% of the time segments logged during the last 6 weeks are hidden. Generally, between 20% and 30% of a user's ACU representations are hidden.

In addition to the visual representations, every element in the display has a textual label. For information items, the labels are an abbreviation of the path, URL, or filename. To create ACU labels, the top five labels from all the contained information items, according to relevance weight, are concatenated.

Orthogonal to the content of the awareness display are the ways in which users can interact with this content. The awareness display supports the following user interactions:

- navigation through the display using mouse clicks,
- accessing information items by double-clicking,
- listing the contents of any ACU representation or parent,
- adding/removing information items, ACU representations, and parents,
- changing the labels information items and ACU representations, and
- changing the relevance of information items within ACUs, or ACUs within a parent.

Structural edits (e.g., relevance changes, additions, removals) performed by the user modify the computational model of the ACU (i.e., alter the value of $[Y_{kj}]$). These edits can be performed either through context menus or via direct manipulation – by dragging items into and out of ACUs.

4.3.4 Discussion of CAAD System

A constant concern with CAAD, and generally any logging system, is privacy. Although all of the marker occurrence events that CAAD logs are already collected by Windows or other applications, there might be some additional risk related to their centralized aggregation. We currently address this concern by (1) storing the data logs in a single location, which the user of the system can easily access and delete; and (2) performing all of the necessary calculations on the data locally. For future applications of our computational model that require some sharing of detected ACUs, we could re-factor the model update computation so that calculations can be encrypted and run over a centralized network of computers (e.g., *Duan and Canny* [2004]).

A practical concern with the current logging architecture is that it can require significant computational resources if, for example, there are many files whose contents need to be hashed.

Finally, the display in CAAD must balance the natural tension between showing updated ACU representations and retaining enough visual cues between updates so that the user can utilize recognition rather than search and recall. We handle this balance by letting the system generate a new layout after each offline update, and then modifying this new layout using the previous layout details – specifically, labels from the previous layout take precedence and any new information items in an ACU representation (relative to the previous day) are shuffled spatially so that previously included information items can be shown in their original positions. The differences between the new and previous layouts can also be highlighted using a special color scheme (which the user can toggle on and off from the context menu).

Finally, the axes of the display are semantically ordered. The x-axis corresponds to time – newer ACUs are further to the right. The y-axis corresponds to the total amount of time each ACU occurred (i.e., the number of time segments in the marker logs) – more time moves the ACU representation higher in the display.

4.4 Related Work

In this section we consider systems that share our overall goal of supporting computer-based information workers. We group these systems into two categories: those that depend entirely on manual input by the user and those that function in a semi-automatic way, requiring only guiding input by the user. To the best of our knowledge, our work is the first fully automatic activity management system.

Systems falling into the manual category include Unified Activity Manager (*Moran*

[2005]), Activity Explorer (*Muller et al.* [2004]), SphereJuggler (*Morteo et al.* [2004]), Activity Based Computing (*Bardram et al.* [2006]), GroupBar (*Smith et al.* [2003]), Rooms (*Henderson and Card* [1986]), and TaskGallery (*Robertson et al.* [2000]). Because these systems require direct input from the user, they often only capture the text (as opposed to context) of their users' work practices. To overcome this, many of these systems rely on generic templates to pre-populate activity (or task or project) representations (with the usual difficulties of finding representations that are not too generic). CAAD differs from these systems by automatically generating its ACU representations from logs of low-level, interaction events.

Semi-automatic systems include TaskTracer (*Stumpf et al.* [2005]), UMEA (*Kaptelinin* [2003]), and Kimura (*Hansen et al.* [2001]; *MacIntyre et al.* [2001]; *Voida et al.* [2002]). The primary user input in these systems is a ground-truth label of what activity is being worked on at any given time. Generally, these systems rely on users to specify activity labels in real-time. However, people often have trouble labeling and delimiting new activities and, more importantly, often forget to declare activity switches as they happen (partially because activity boundaries can be ambiguous and partially because they can be multiplexed). CAAD overcomes these issues by automatically generating ACU representations and by allowing these representations to evolve and interleave.

In addition, CAAD supports user editing of within-ACU and between-ACU structures. This editing functionality gives users the ability to both correct the ACUs detected by CAAD and to organize the ACUs into more meaningful arrangements.

In the next chapter we discuss the user studies we performed on CAAD that verify its utility for end-users.

Chapter 5

Evaluation: Assessing Utility

In this chapter we describe the two user studies we conducted on CAAD. The first study was designed to assess two things: (1) the utility of CAAD and (2) the strengths and weaknesses of our computational model. Using quantitative measures of perceived usefulness, we show that CAAD provides some utility to its users. To understand the strengths and weaknesses of our computational model, through the (indirect) lens of CAAD, we also discuss qualitative findings from semi-structured interviews with study participants.

In the second user study, we focused on the ways in which representations of one's past activities and contexts, in addition to CAAD, are or are not useful. We relied primarily on semi-structured interviews to elicit evidence on this topic. A one week deployment of CAAD, with each participant, was used as an interview probe and to ground the discussion about representations of activities and contexts and their utility.

5.1 First User Study: the Utility of CAAD

5.1.1 Background

The central question of this dissertation calls for a demonstration of the utility of a context-aware application that automatically tracks and models a user's activity and context. In CAAD, this utility takes the specific form of helping computer-based information workers (1) minimize the amount of overhead required to access relevant information, and (2) improve their awareness of the various activities and contexts that compose their day-to-day work flow and behavior.

The study started with 10 participants (7 graduate students, 3 undergraduate students) and was conducted in the actual work settings of each participant. All of the participants were working on single-monitor, desktop computers prior to the study. For the study they were provided with a second monitor, used primarily, but not exclusively, for the display component of CAAD. Usage was not strictly enforced and participants often placed other application windows over the display.

The key characteristic for all participants was their involvement in some form of computer-based information work. To assess this, we administered an initial questionnaire that asked participants about the amount of time they spent on the computer and what tasks or projects they worked on during this time. Example projects found in this questionnaire include: preparing lecture notes, managing a small research team, conducting studies, processing study data, searching for and reading related papers, and developing software. The vital concern that the initial questionnaire was designed to assess is whether the participants spent a significant amount of time working on the computer. (We defined "significant" as at least 60% of their scheduled work time.) If they did not, then detecting their computer-based ACUs would be of limited value in supporting their work. Based on the questionnaire results, we were confident that the participants in the study could be supported by CAAD.

During the study we collected data from three sources: questionnaire results measuring perceived usefulness and ease-of-use, logs of actual usage events with CAAD's display, and semi-structured interviews. Before describing the results we discuss the method details for each of these data sources.

5.1.2 Study Methods

Perceived Usefulness and Ease-of-use Questionnaire

The questionnaire we used measures perceived usefulness and perceived ease-of-use (*Davis* [1989]). Table 5.1 contains the actual questions we used. All questions were scored on a 7-point Likert scale, ranging from -3 to +3. Scores greater than zero indicate that the participant found the system useful or easy to use. We modified the standard questionnaire (prior to administering it) by removing the lowest correlating question for ease-of-use. This question targeted system flexibility, which is not always correlated with ease-of-use or with overall system usage (*Davis* [1989]).

Eight of the original ten participants completed the questionnaire after the first week of the study. One of the remaining two participants worked heavily in a development environment from which CAAD could not log interaction events. The other remaining participant worked in multiple locations, making a single work log practically infeasible to gather. Additionally, five participants took the questionnaire a second time (the other three participants were unavailable for subsequent interviewing). The second application of the questionnaire was designed to assess two things: the novelty effects in the first application of the questionnaire and the effects of longer-term use of CAAD.

Table 5.1. Questionnaire results. Questions were scored on a 7-point Likert scale ranging from -3 to +3. Mean and standard deviation response values are reported for each question and for each overall response average.

Question	Mean	S.D.	
Using the activity display while I work would enable me to accomplish tasks more quickly.	1.63	0.52	Perceived usefulness
Using the activity display would improve my work performance.	0.63	0.92	Perceived usefulness
Using the activity display would make it easier to do my job.	1.25	0.71	Perceived usefulness
Using the activity display would enhance my effectiveness at work.	0.88	0.83	Perceived usefulness
Using the activity display while I work would increase my productivity.	0.88	0.83	Perceived usefulness
I would find the activity display useful in my work.	1.00	0.93	Perceived usefulness
Average perceived usefulness score:	**1.04**	**0.69**	Perceived usefulness
I find it easy to get the activity display to do what I want it to do.	1.13	0.99	Perceived ease-of-use
My interaction with the activity display is clear and understandable.	1.25	1.04	Perceived ease-of-use
Learning to operate the activity display was/is easy for me.	1.25	1.83	Perceived ease-of-use
It was/would be easy for me to become skillful at using the activity display.	2.00	0.82	Perceived ease-of-use
I find the activity display easy to use.	1.50	0.93	Perceived ease-of-use
Average perceived ease-of-use score:	**1.42**	**0.92**	Perceived ease-of-use

Actual Usage of the Awareness Display

To assess how well CAAD met its design goal of supporting access to information we logged usage events with the display. These logs include every information access as well as every edit or modification (Figure 5.1 shows the events that were logged). We collected these logs for 7 of the original 10 participants during the first week of the study (system compatibility issues resulted in minimal exposure to CAAD for two participants and a third participant could view the display but not interact with it during the first week of the study).

	Usage events per participant							
access event								61
delete info. item								51
add info. item								18
label change								12
merge ACUs								10
add parent								8
remove ACU								7
add ACU								2
change relevance								1
	25	18	58	23	7	19	20	Totals

Figure 5.1. Visualization highlighting actual usage patterns. The seven middle columns correspond to participants. Shading linearly scales with the percentage of that type of event, per participant – the darker the shade, the higher the percentage. Column totals are the number of events generated per participant. Row totals are the number of each specific event type generated by all participants.

Interviews

In addition to validating the questionnaire results, semi-structured interviews were used to understand the strengths and weaknesses of the computational model (which creates the ACU representations in CAAD). We interviewed 8 of the original 10 participants (excluding the two with minimal exposure to CAAD due to system compatibility issues). We were particularly interested in examples of ACU representations that were or were not accurate in the participants' opinions. Through these examples, we assessed how well the computational model was able to automatically detect accurate ACU representations.

5.1.3 Results

Perceived Usefulness and Ease-of-use Questionnaire

Again, all questions were scored on a 7 point Likert scale, ranging from -3 to +3. Both the average perceived usefulness responses (T-stat = 4.2785, df = 7, p $\leq$ 0.0037) and the average perceived ease-of-use responses (T-stat = 4.36, df = 7, p $\leq$ 0.0033) were statistically significant in the positive direction (zero is a neutral response). This means study participants found the activity display both useful and easy to use. Specific question results are shown in Table 5.1.

Of the five participants that took the questionnaire a second time, three took it after the second week of the study, one after the third week of the study, and one after the fourth week of the study. The averaged score differences were positive for both perceived usefulness and ease-of-use for each participant (indicating an increase in perceived usefulness and perceived ease-of-use) with the exception of one participant on ease-of-use (this participant felt like the semantic zooming, which hides labels when they are too small, made the display hard to quickly read). The overall increase in questionnaire scores indicates that: (1) the initial questionnaire results were not artificially inflated due to novelty effects, and (2) that CAAD was able to effectively track the evolution of people's ACUs over multiple weeks.

Actual Usage of the Awareness Display

Actual usage results are presented in Figure 5.1. Aggregating the seven usage logs, we found that 36% of the usage events were information access events. The majority of edit events corresponded to the deletion of an information item from an ACU representation. On average, users generated 8.1 events per day (the logs covered at most three days worth of interaction for each participant). However, most participants deviated from this average significantly. The least active participant only generated 7 total events while the most active generated 58. The variance in these usage results, coupled with the questionnaire results, provides evidence of CAAD's ability to support information workers with different working styles.

We also calculated a derived metric of CAAD's utility from the usage logs. If we count the number of elements displayed by CAAD and divide by the number of structural edits that users made, we get an estimate of the value CAAD adds relative to a manual task-management system (where users have to insert and organize information items from scratch). The validity of this metric is conditional on the perceived usefulness of CAAD - basically we assume users performed as many edits as were required to develop the opinion that CAAD is useful. Likely, however, users performed more edits than this, making this metric a conservative measure of CAAD's utility.

Values greater than 1.0 for this metric indicate benefit to the user. For the seven participants that we have usage data for, these values ranged from 1.5 to 68.0 (mean 18.4, standard deviation 25.2, not statistically significant above 1.0). The fact that all of these values are greater than 1.0 indicates that CAAD provides an overhead reduction relative to manual task management systems.

An alternatively way to measure the utility is to take the difference instead of the ratio. Whereas the ratio metric can be thought of as a relative measure of utility (relative to how many edits each user made), the difference based metric captures more of an absolute measure of utility (how many edits did the user save). Values for the difference based metric ranged from 6 to 177 (mean 77.3, standard deviation 61.4) and were statistically significant above zero: T-stat = 3.3291, df = 6, $p \leq 0.016$. Again, the fact that all of these values are greater than zero indicates that CAAD provides an overhead reduction relative to manual task management systems.

Interviews

To start the interview, we asked participants to describe their overall impressions of CAAD. The comments we received were evenly split between positive assessments (e.g., "it makes reasonable prediction on what things belong together") and more negative, but balanced, criticisms (e.g. "groups [ACUs] are mostly correct sometimes they have a few additional things that don't belong"). We then asked about specific interactions as evidence for the participants' overall assessments. In particular, we made sure to gather the participant's reasons for negative and/or critical comments.

In pursuing the evidence that participants had for their critical comments, we found that they were always the result of contradictions between the participants' mental models of their own work and the ACU representations found by the computational model component of CAAD. Note, however, that the quantitative results presented above indicate that CAAD was useful in spite of these contradictions – user's were still able to access relevant information items through the display and to gain insights into the way their work behavior is structured. Additionally, we do not believe, nor are we trying to imply, that the ACU representations in CAAD, or the user's mental models, are "right". They are merely two perspectives on work practice, whose difference is a potent source for reflection.

The contradictions we uncovered fit into four common types. The most frequent contradiction concerned ACUs with information items that "did not belong" - e.g., web pages related to sports or news grouped with work documents. About 40% of ACUs across the eight participants we interviewed highlighted this contradiction. The notion of belonging was determined by the participants' mental models of their work - e.g., knowing a sports score does not help them complete their work. However, they were aware of reading sports and news pages while they worked. These breaks, or rather micro-breaks, generally lasted less than one minute but occurred regularly.

Because these breaks were so short, participants often left the documents and applications relevant to their current activity open. Hence, in terms of temporal correlation, these sports and news web pages seem to belong in the ACUs associated with their work. They are, in a counter-intuitive sense, part of the context of the participant's routine work-flow. With more elaborate content analysis, it might be possible to separate them, but we think for now it is more appropriate to include them. At the very least they may be indicators of context switches and trigger associations between the users' tasks.

Another frequent contradiction related to separate ACU representations that "belong together" or "should be merged" - either into a single ACU or as children of a single parent (we estimate about 25% of ACUs highlight this contradiction). For example, one participant said: "My 'preparing lecture notes activity' really belongs with my 'preparing the midterm activity'." However, with additional questioning, we found that the participant had not worked on these tasks at the same time, nor did they share many information items. Thus, from CAAD's perspective, these activities were practically disjoint. Some participants dealt with this problem by creating parent circles and grouping ACU representations together (Figure 4.1 illustrates this).

A third contradiction, voiced by three participants, was that "there were too many groups [ACUs] in the display". One participant even stated that no matter what he was working on, he did not want to see more than 4 or 5 ACUs in the display. However, while discussing the ACU representations in the display, only two of seventeen in this participant's display were not readily identifiable by him, and these two were identifiable after some extended recall of minor past activities. In other words, this participant knew he worked on more than 5 things but did not want to be shown these things by CAAD. The other two participants who made this comment had similar, although less extreme, sentiments.

Finally, the fourth contradiction concerned ACU representations that were "missing relevant documents" - we estimate about 23% of ACUs were subject to this comment. There seem to be two causes for this type of contradiction. First, CAAD does miss events for some document types, like rich web-based media. However, this did not apply to any of the eight participants we interviewed. The second reason, as with the ACU merging contradiction discussed above, is that participants had some abstract connection in their mental model of their work that was not part of their day-to-day work routine, and hence was invisible to CAAD. With sufficient follow-up questioning, it was clear that participants had never used the "missing" document while working with the other information items in the ACU - i.e., there was no temporal coordination between their uses. So the computational model in CAAD would treat the "missing" document as disjoint, in terms of ACU representations, from the other information items.

5.1.4 Discussion

Based on the quantitative results, people found CAAD useful and easy to use. Using a metric derived from the amount of ACU structure automatically detected by CAAD and the number of edits people performed on these structures, we showed that CAAD provides clear overhead reduction relative to manual activity/task management approaches. We also suspect that CAAD requires less overhead than semi-automatic methods (e.g. *Stumpf et al.* [2005]), although we have not specifically studied this difference.

During the semi-structured interviews, we elicited 4 types of contradictions between CAAD's ACU representations and participants' mental models of their activities and contexts. Each of these contradictions highlights some strengths and weaknesses of our computation model.

The first contradiction, that some ACUs in CAAD contained information items that did not belong, is a consequence of our computational model's dependency on the temporal coordination of observed activity and context markers. If a person regularly accesses sports web sites as a micro-break during their "work" activity, the regularity of these accesses leads to a significant temporal correlation between markers of work and markers of the person's sports interests. While temporal coordination is clearly a reasonable assumption for activities and contexts, it should be augmented with some awareness of temporal boundaries (across which coordination is known not to happen). These temporal boundaries are often referred to as "interruptions" (*González and Mark* [2004]) in a work setting. Our computational model of activity and context could be improved by explicitly including temporal boundaries.

The second and fourth contradictions, that some ACUs should be merged and that some ACUs are missing relevant information items, highlight another weakness of the temporal coordination assumption: that not all relevant markers, i.e., information items, must be present every time an activity is worked on or a context is evoked. Projects often have stages, with each stage having a (slightly) different set of coordinated, observable markers. Our computational model could be improved by including higher-order structures that model how ACU representations work together in larger activities and contexts.

The remaining contradiction, that there were too many ACUs in the display, highlights an important factor in the perceived utility of CAAD as well as in our computational model of activities and contexts. Generally speaking, people are not interested in seeing a representation of every activity and context they are involved in. Sometimes, this is because certain activities and contexts are embarrassing – some of our participants commented that they did not appreciate seeing how often they took breaks when they "should be working". Alternatively, some activities and contexts are inconsequential to people's sense of identity and purpose; and so they do not care to see evidence of them. In terms of our computational model, improvements could be made by identifying ‘inconsequential’ or ‘embarrassing’ ACUs. However, we think the best place for this type of identification is in the logging component – in filtering which markers the computational model sees.

Finally, and in prelude to the second user study, we note the comment of one participant: “[CAAD] definitely shows relevant things. But they are not always useful.” This participant was referring to his lecture preparation for the course he was teaching. While preparing new slides, CAAD would show him slides from previous lectures he had already written. Although related, these old slides were not really useful in preparing the new slides. CAAD, in its current implementation, can only support ongoing activities that re-use the same (or a similar) sets of information resources. CAAD is not able to track an ACU if the relevant documents and people (i.e., markers) change too quickly. Furthermore, CAAD, and really our computational model, cannot project future requirements of activities into current ACU representations. However, as we discuss in the next user study, many self-maintained representations of people’s activities blend records of the past with expectations for the future. We believe this feature, or lack of feature, impacts the utility of CAAD for its users.

5.2 Second User Study: the Utility of ACU Representations

5.2.1 Background

In the first user study, we demonstrated that CAAD can provide utility for its users by automatically building and maintaining representations of ACUs. Interestingly, this utility exists in spite of evidence that CAAD's ACU representations often differ from users' mental models of their various (computer-mediated) activities.

In this study, we conducted interviews to better understand the utility of any form of ACU representation – whether they be in CAAD or not. More specifically, we know that the utility of CAAD is connected to its support for accessing relevant information and improving people's awareness of their work flow. However, it is unclear whether these forms of utility transfer to other activity and context representations in people's lives. Using CAAD as an interview probe, we asked participants to describe other representations of their activities and contexts, and how they both maintain and derive utility from these representations.

For this study, we recruited 12 participants. They were all graduate students at a major university. Two of the participants were studying humanities; the rest were in engineering or related disciplines. Participants ran the CAAD software on their primary work computer (laptop or desktop) for at least ten days before the interview.

5.2.2 Study Methods

As in the first user study, the interviews in this study were semi-structured. The interviews followed the following rough outline: (1) discuss the ACU representations in CAAD; (2) identify other people or tools that maintain representations of the participant's activities (and contexts), with particular attention to differences between these representations and the participant's own mental model; and (3) identify strategies for maintaining and using the representations revealed in (2).

5.2.3 Results

In terms of the differences between CAAD's ACU representations and participants' mental models, we reconfirmed the same four types that we found in the first user study. They also occurred in roughly the same proportions: with 33% of ACUs containing non-relevant items, 31% of ACUs needing to be merged or grouped together with other ACUs, 3 of 12 participants desiring fewer ACUs in the display overall, and 21% of ACUs missing relevant information items.

Interestingly, none of the participants believed that CAAD could ever develop ACU representations that perfectly matched their mental models. The primary reason for this, cited by every participant, was the fact that all of their work/research activities had some non-computer-mediated component. The most common non-computer component of their activities was reading physical papers or books. Other frequently mentioned non-computer components were: face-to-face meetings (with their advisers and other colleagues), thinking, writing (on paper), and organizing physical materials.

Of course, our participants were also engaged in non-work activities as well. Many of these activities, at least the ones that have some computer-mediation, involved web-searches for information (e.g. looking up nearby hikes or plane tickets), playing games, and/or listening to music. But, as with the work activities, these activities also had significant non-computer-mediated components – like physically traveling to various locations, talking to people on the phone or in person, and engaging in some kind of physical behavior (like camping).

An interesting aspect of these activities is the range of strategies that participants described for minimizing the differences between their mental models of these activities and CAAD's ACU representations. All of the strategies discussed during the interviews involve selectively weighting or filtering certain types of information items

– e.g. for some work activities, web pages are not of central importance, and should be weighted less in the calculation for finding ACUs. Similarly, in activities where the participant is authoring a paper, more weight should be placed on the type of documents being authored. Yet another strategy, offered by three of the participants, was to highly weight the folder structure on the computer – because it highly resembles their mental model of their activities. However, three other participants, when asked about this particular strategy, said it would not work for them because they do not keep their folders up to date – the range of styles in managing folders, and information, is known to be large (*Boardman and Sasse* [2004]).

Furthermore, the range of activities that people are involved in, both individually and collectively, necessitates a conflicting set of strategies for improving the match between ACUs in CAAD and users' mental models. For example, for the afternoon hiking activity, one could improve CAAD's ACU representation by increasing the weight of web pages. However, for a research project, one would need to increase the weight of tex and Word files. A compromise could be reached. However, it would likely continue to generate differences between the user's mental model of their activities and CAADs ACU representations of them. Moreover, the optimal trade-off point between the weight of web pages and tex/Word files is likely different for different people – depending on the number of activities serviced by each strategy, and the subjective importance placed on these activities.

After focusing on CAAD and its ACU representations, the interviews shifted focus to other things, not necessarily human, that maintain some representation of the participants' activities. The most common non-human agent that maintained a representation of the participants' activities was a planner/calendar or to-do list. The key difference between planners and to-do lists relative to CAAD is that they often focus on the future: including upcoming events and tasks/actions that have not yet been completed. Past behavior is almost never recorded after the fact, and past to-do items and planner entries are rarely consulted (and often discarded). Interestingly, these planners and to-do lists were as likely to be in electronic form as not.

Another common representation of the participants' activities was maintained in the folder hierarchy of the participants' computers. The most diligent organizers would edit this representation on a weekly basis; the least diligent once every few months. What is striking, however, about the representation in the computer folder hierarchy is the amount of "potentially" relevant, but unused, information items. Many participants collected documents that might become useful in the future, but that they historically were not using. The set of documents they did use were often located in multiple sub-folders, representing different types of tasks like "coding" and "writing", within a higher-level project folder. Interestingly, project hierarchies often had similar structures – i.e., exhibited genre-level similarities.

In terms of people who maintain representations of the participants' activities, the most commonly mentioned were research collaborators. However, the representations that collaborators would maintain about the participants' activities were limited to certain shared parts of a select few activities. No participant had a person in their life that was maintaining some representation of all the activities (and contexts) they were involved in – or even trying to maintain such a representation. Furthermore, as with the planners and to-do lists, activity representations held be collaborators often focused on future plans – e.g., next steps to be completed, submission deadlines, budgets, etc.

For both non-human and human generated representations of the participants' activities (and contexts), the key method for reconciling differences relative to the participants' mental models was regular feedback/communication. More specifically, the accuracy of the representations seemed to depend primarily on how often the participants edited the representations (for the non-human ones), or engaged in open communication about what has been done and what needs to be done (with other people). A necessary condition for regular feedback and communication is ease of access – if it is cumbersome for participants to edit a representation of their activities, or to find regular meeting times with their colleagues, then the likelihood of the representation becoming inaccurate increased. For some representations, this was not a problem – e.g., one participant told us that his friends and family only received the "high-level elevator pitch" of his research because they did not need any more detail than that. For other representations, like a critical to-do list or a key collaborator's understanding of one's activities, inaccuracies can be costly.

In terms of use, the primary utility in all of these representations was their ability to guide future activity. Sometimes, this guidance was fairly detailed, as in listing out the time, location, participants, and topic of a meeting. The more likely situation, however, was that the representation served as a loose guide for the person's activity – e.g., stating in high-level, general terms what behaviors needed to, or should, be performed.

5.2.4 Discussion

The differences between CAAD's representations of its users' activities and contexts and their own representations (in their mental models) are analogous to the differences that arise between any two representations of activity and context built from different perspectives. Different perspectives are, necessarily, employed by different people, or tools, or even the same person separated in time (*Korzybski* [1933]). These differences in perspectives impact group work (*Hill et al.* [2001]; *Hinds and Baily* [2003]; *Hinds and Mortensen* [2005]; *Mohammed et al.* [2000]; *Mohammed and Dumville* [2001]), and can be understood and conflicting mental models of the group's activity and the context around it (*Carroll et al.* [2006]). The primary strategy for dealing with these differences is regular and frequent communication and feedback. We think part of CAAD's ability to provide utility for its users hinges on the ease with which the content of the display, the ACU representations, can be edited.

More important, though, is the ability (or lack of ability in CAAD's case) to provide guidance for future activities. All of the other activity and context representations that our study participants maintained and employed in their lives supported this functionality. And, as the work flow of a graduate student can be characterized by the coupled problems of identifying and then executing "the next step" in their research project(s), this functionality had significant utility for our study participants. Other research has verified that this functionality is important in general information work settings outside of academia (*González and Mark* [2005]).

It follows that the real utility of automatically detecting and maintaining ACU representations might not be in an activity awareness display. However, we note that building CAAD was a useful and necessary process to assess the strengths and weaknesses of our computational model. In the Future Work chapter that follows, we discuss other applications of our computational model.

Chapter 6

Conclusion and Future Work

There are three primary contributions in this dissertation. The first contribution is a survey of existing theories of activity and context, both within and outside of Computer Science. This survey justifies and grounds our definitions of activity[1] and context[2]; and it reveals the key characteristics that a model of activity and context must address: observability (of activity and context markers), inseparabilitiy (it is a matter of perspective what qualifies as activity versus context), dynamism (activity and context are not static phenomena), multiplexity (multiple activities and contexts can occur simultaneously), and ontological richness (activities and contexts exhibit inter-structure relationships that generate genre-like classification systems).

The second contribution of this dissertation is the translation of our notions of activity and context into a computational model. The key strength of our model is its ability to dynamically determine which context markers are relevant and how these markers group together into representations of activities and contexts. However, as discussed above, our model does not address the ontological richness of activities

[1] We define activity as the intentional behavior that produces both context markers and the socially shared, i.e., cultural, mapping between context markers and appropriate activities.

[2] We define context as a set of markers (perceivable differences in the world) that enables a person's assessment of which activities, and which interpretations of these activities, are appropriate.

and contexts. It does not look for structural similarity between representations of activities and contexts. This omission simplifies our computational model. We briefly discuss how to add ontological richness into our model in the future work section below.

The third contribution of this dissertation is the validation of an automatic approach to Context-Aware Computing. Specifically, by building and studying the Context-Aware Activity Display (CAAD), we show that automatically generating representations of activities and contexts provides some utility to users. However, as discussed in the previous chapter, the utility of CAAD has two limitations: First, the activity and context representations generated by CAAD often differ from users' mental models. These differences are unavoidable and not necessarily negative – they provide fertile grounds for self-reflection. The second limitation is that CAAD provides little to no support for future activities (or future actions for an existing activity) – which is a key characteristic of many activity representations that people employ. Based on these findings, we think other applications of our activity and context model (i.e., other than an awareness display) might provide significantly more utility to end users. We discuss some potential applications below.

Finally, we want to emphasize that this dissertation is not intended to provide the final statement on any outstanding open problems, or to end any lines of research. Rather, the validation provided in this dissertation should be understood as a clear step in a new direction of activity and context based technology design – towards context-aware applications that rely on open-ended, dynamic, and automatically generated models. Our ultimate objective is to inspire more research attention on technology that can effectively handle and respond to the nature of human activity and its relationship to context.

6.1 Future Work

In remainder of this chapter we briefly discuss three directions for extending the work described above. These three directions represent three important trends in the research areas of activity and context modeling and system design.

6.1.1 Richer Models of Activity and Context

As described in chapter 2, one of the key characteristics of our conceptual model of activity and context was ontological richness – the existence of higher order patterns within and between activities and contexts that correspond to genres and roles. In translating our conceptual model of activity and context to a computation one, we chose to leave much of this ontological richness out of our model. However, as argued by the many theories of activity and context, and evidenced in our user studies, such higher order patterns are not only ubiquitous, they are often key components in people's mental models of their lives. Hence, a system that can accurately track and leverage (possibly through simple visualization) genre and role patterns that are exhibited in a person's behavior could be more semantically meaningful than the activity-context unit representations currently displayed by our system.

As noted earlier, there are a number of past and current projects that we could draw on to inform the addition of genre and role patterns in our computational model of activity and context (e.g., *Bardram* [1998]; *Chaiklin and Lave* [1993]; *Ducheneaut* [2003]; *Engeström* [2000]; *Fjeld et al.* [2002]; *Goffman* [1959, 1967, 1974]; *Hill et al.* [2001]; *Hutchins* [1995]; *Hutchins and Klausen* [1996]; *Latour* [1986, 1995]; *Lave and Wenger* [1991]; *Law* [1992]; *Millen et al.* [2005]; *Mukherjee et al.* [2003]; *Muller et al.* [2004]; *Palen and Dourish* [2003]; *Shaw and Gaines* [1999]; *Wasserman and Faust* [1994]; *Wells* [2002]; *Yoshioka et al.* [2000]; *Yoshioka and Herman* [2000]; *Yoshioka* *et al.* [2001]). Simple hybrid systems that combine the main pattern mining components of CAAD with some of the pattern mining algorithms used for genres and roles would be a good place to start.

6.1.2 Moving Beyond Individuals

The CAAD system is currently designed for individual users. The activity-context unit representations that are discovered by CAAD are based completely on individual user behavior. However, higher order structures like genres and roles will likely require pattern mining across individuals – searching for regularities that are visible and important for entire social groups. To mine for these group patterns, one needs to coordinate the logged behavior sensing data from multiple individuals – e.g., one needs to figure out how a file opening event for one person corresponds (or does not correspond) to a file authoring event by another person. While such coordination is highly visible in some domains (e.g., *Hutchins* [1995]; *Hutchins and Klausen* [1996]), it may be conflicted, volatile, and/or nearly invisible in others (e.g., *Miettinen* [1999]).

To move in this direction, one will likely need to select a specific domain – e.g., information workers in a modern office setting. Then, conduct detailed qualitative field work to reveal the nature of inter-personal coordination in that domain (e.g., *Beyer and Holtzblatt* [1998]; *González and Mark* [2004]; *Spinuzzi* [2003]). Once the mechanisms of coordination are visible and understood, one could develop pattern mining algorithms to detect group behavior patterns.

6.1.3 Applications of Activity and Context Modeling

There are a number of possible applications that could leverage the activity and context modeling infrastructure of CAAD. Importantly, the probabilistic nature of the

ACU modeling in CAAD creates certain restrictions on which applications are likely to function effectively (i.e., provide utility to their users). Specifically, we require applications that handle uncertainty in the structure of the ACUs. We simply list some potential applications along with related work references:

- context/activity sensitive information retrieval (*Budzik et al.* [2001]; *Hearst* [1999]; *Hong and Landay* [2001a]; *Honkaranta* [2003]; *Lieberman* [1995]; *Proctor et al.* [1998]; *Rhodes* [2000]) – even for old information (*Dumais et al.* [2003])

- activity based (and hence collaborative work-group based) data access, sharing, and control (*Dourish and Bellotti* [1992]; *Dourish et al.* [1999]; *Hong and Landay* [2001b]; *Millen et al.* [2005]; *Muller et al.* [2004]; *Palen and Dourish* [2003]; *Spiteri and Bates* [1998]; *Trevor et al.* [2002]; *Truong et al.* [2001]) – even for access to people in explicit or tacit workgroups (*Begole et al.* [2002]; *Bull et al.* [2001]; *Donath* [1994]; *Fisher* [2000]; *Fogarty et al.* [2005b]; *Isaacs et al.* [1996]; *Nardi et al.* [2000]; *Oliver et al.* [2002])

- context/activity sensitive task management (*Bellotti et al.* [2004]; *Cadiz et al.* [2002]; *Christensen and Bardram* [2002]; *Flores et al.* [1988]; *Gwizdka* [2002]; *Hansen et al.* [2001]; *MacIntyre et al.* [2001]; *Mynatt* [1999]; *Nardi et al.* [2000]; *Voida et al.* [2002])

- context/activity sensitive monitoring systems (*Cadiz et al.* [2002]; *Erickson and Laff* [2001]; *Fisher* [2000]; *Gross et al.* [2003]; *Hardless and Nulden* [1999]; *Maglio et al.* [2001]; *Pedersen* [1998]; *Sawhney et al.* [2001]; *Truong et al.* [2001])

- context/activity sensitive metadata management (*Davis et al.* [2004]; *Honkaranta* [2003])

- OS enhancements that support task/action execution in context like more effective caching of related/relevant information and the pre-starting of commonly used applications/tasks (*Crowley et al.* [2002]; *Franklin et al.* [2002]; *Horvitz and Paek* [1999]; *Intille* [2002]; *Petrelli et al.* [1999]; *Terry and Mynatt* [2002]) and even plan generation (*Küpper and Kobsa* [1999, 2002]; *Torre* [2001]) and notification (*McCrickard et al.* [2003]; *Tarasewich et al.* [2003]; *van Dantzich et al.* [2002])

- and better error recognition and recovery, possibly through more intelligent help systems (*Fischer* [2001]).

Bibliography

Abowd, G., E. Mynatt, and T. Rodden, The human experience, in *IEEE Pervasive Computing*, vol. 1(1), pp. 48 – 57, IEEE Press, 2002.

Ackerman, M., The intellectual challenge of CSCW: the gap between social requirements and technical feasibility, in *Human-Computer Interaction Journal*, vol. 15, pp. 179 – 203, 2000.

Ackerman, M., T. Darrell, and D. Weitzner, Privacy in Context, in *Human-Computer Interaction: Special issue on context-aware computing*, vol. 16, pp. 167 – 176, Taylor and Francis Group, 2001.

Addelson, K., Action, scientific knowledge, and the location of the scientist, in *Mind, Culture, and Activity*, vol. 2(1), pp. 33 – 38, University of California Press, 1995.

Agar, M., We Have Met the Other and Were All Nonlinear: Ethnography as a Nonlinear Dynamic System, in *Complexity*, vol. 10(2), pp. 16 – 24, Wiley Periodicals, 2004.

Akman, V., Rethinking context as a social construct, in *Journal of Pragmatics*, vol. 32(6), pp. 743 – 759, 2000.

Allan, J., C. Wade, and A. Bolivar, Retrieval and novelty detection at the sentence level, in *Proceedings of the International Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 314 – 321, ACM Press, 2003.

Allwood, J., An activity-based approach to pragmatics, in *Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics*, edited by H. Bunt and B. Black, pp. 47 – 80, John Benjamins, 2000.

Atkinson, J., and J. Heritage (Eds.), *Structures of Social Action: studies in conversation analysis*, Cambridge University Press, 1984.

Bakhtin, M. M., *Speech genres and other late essays*, University of Texas Press, Slavic Series, 1979.

Bardram, J., Collaboration, Coordination, and Computer Support: an activity theoretical approach to the design of Computer Supported Cooperative Work, Ph.D. thesis, University of Aarhus, 1998.

Bardram, J., J. Bunde-Pedersen, and M. Soegaard, Support for activity-based computing in a personal computing operating system, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 211 – 220, ACM Press, 2006.

Bateson, G. (Ed.), *Steps to an Ecology of Mind*, University of Chicago Press, 1972.

Begole, J., J. Tang, R. Smith, and N. Yankelovich, Work rhythms: analyzing visualizations of awareness histories of distributed groups, in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 334 – 343, ACM Press, 2002.

Bellotti, V., and W. K. Edwards, Intelligibility and Accountability: Human Considerations in Context-Aware Systems, in *Human-Computer Interaction: Special issue on context-aware computing*, vol. 16, pp. 193 – 212, Taylor and Francis Group, 2001.

Bellotti, V., B. Dalal, N. Good, P. Flynn, D. Bobrow, and N. Ducheneaut, What a to-do: studies of task management towards the design of a personal task list manager, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 735 – 742, ACM Press, 2004.

Bertelsen, O. W., and S. Bodker, Activity Theory, in *HCI models, theories, and frameworks: toward a multidisciplinary science*, edited by J. M. Carroll, pp. 291 – 324, Morgan Kaufmann, 2003.

Beyer, H., and K. Holtzblatt, *Contextual Design: Defining Customer-Centered Systems*, Academic Press, 1998.

Bianchi-Berthouze, N., L. Berthouze, and T. Kato, Understanding subjectivity: an interactionist view, in *Proceedings of the International Conference on User Modeling*, edited by J. Kay, pp. 77 – 86, Springer-Verlag, 1999.

Billsus, D., and M. Pazzani, A hyrbid user model for news story classification, in *Proceedings of the International Conference on User Modeling*, edited by J. Kay, pp. 99 – 108, Springer-Verlag, 1999.

Billsus, D., and M. Pazzani, User modeling for adaptive news access, in *User Modeling and User-Adapted Interaction*, vol. 10, pp. 147 – 180, Kluwer Academic Publishers, 2000.

Blackler, F., Knowledge and the theory of organizations: organizations as activity systems and the reframing of management, in *Journal of Management Studies*, vol. 30(6), pp. 863 – 884, Blackwell Publishing, 1993.

Blackler, F., N. Crump, and S. McDonald, Managing Experts and competing through innovation: an activity theoretical analysis, in *Organization*, vol. 6(1), pp. 5 – 31, Sage Publications, 1999.

Blackler, F., N. Crump, and S. McDonald, Organizing processes in complex activity networks, in *Organization*, vol. 7(2), pp. 277 – 300, Sage Publications, 2000.

Blei, D., A. Ng, and M. Jordan, Latent dirichlet allocation, in *Journal of Machine Learning Research*, vol. 3, pp. 993 – 1022, MIT Press, 2003.

Blomberg, J., M. Burrell, and G. Guest, An ethnographic approach to design, in *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, edited by J. Jacko and A. Sears, pp. 964 – 986, Lawrence Erlbaum Associates, 2003.

Boardman, R., and M. A. Sasse, 'Stuff goes into the computer and doesn't come out': a cross-tool study of personal information management, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 583 – 590, ACM Press, 2004.

Boer, N.-I., P. van Baalen, and K. Kumar, An activity theory approach for studying the situatedness of knowledge sharing, in *Proceedings of the Annual Hawaii International Conference on System Sciences (HICSS)*, IEEE Press, 2002.

Breure, L., *Development of the genre concept*, (available online), 2001.

Brown, J. S., and P. Duguid, *Organizational learning and communities-of-practice: toward a unified view of working, learning, and innovation*, Available electronically: http://www2.parc.com/ops/members/brown/papers/orglearning.html, 1991.

Brown, J. S., and P. Duguid, *The Social Life of Documents*, Available electronically: http://www2.parc.com/ops/members/brown/papers/sociallife.html, 1995.

Bruner, J. S., *Toward a Theory of Instruction*, Harvard University Press, 1966.

Budzik, J., K. Hammond, and L. Birnbaum, Information access in context, in *Knowledge Based Systems*, vol. 14, pp. 37 – 53, 2001.

Bueno, D., and A. David, METIORE: a personalized information retrieval system, in *Proceedings of the International Conference on User Modeling*, edited by M. Bauer, P. Gmytrasiewicz, and J. Vassileva, pp. 168 – 177, Springer-Verlag, 2001.

Bull, S., J. Greer, G. McCalla, L. Kettel, and J. Bowes, User modeling in I-help: what, why, when, and how, in *Proceedings of the International Conference on User Modeling*, edited by M. Bauer, P. Gmytrasiewicz, and J. Vassileva, pp. 117 – 126, Springer-Verlag, 2001.

Bunt, A., and C. Conati, Probabilistic student modeling to improve exploratory behaviour, in *User Modeling and User-Adapted Interaction*, vol. 13, pp. 269 – 309, Kluwer Academic Publishers, 2003.

Buntine, W., and A. Jakulin, Discrete principal component analysis, in *Proceedings of the Subspace, Latent Structure and Feature Selection Techniques: Statistical and Optimisation perspectives Workshop*, 2005.

Burke, K., *A Rhetoric of Motives*, University of California Press, 1950.

Cadiz, J. J., G. Venolia, G. Jancke, and A. Gupta, Designing and deploying an information awareness interface, in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 314 – 323, ACM Press, 2002.

Canny, J., GAP: a factor model for discrete data, in *Proceedings of the International Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 122 – 129, ACM Press, 2004.

Card, S., A. Newell, and T. Moran, *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Inc., 1983.

Carroll, J. M., M. B. Rosson, G. Convertino, and C. H. Ganoe, Awareness and teamwork in computer-supported collaborations, in *Interacting with Computers*, vol. 18(1), pp. 21 – 46, Elsevier, 2006.

Chaiklin, S., and J. Lave (Eds.), *Understanding Practice: Perspectives on Activity and Context*, Cambridge University Press, 1993.

Cheng, H., Knowledgescapes: a probabilistic model for mining tacit knowledge for information retrieval, 2000.

Cheyne, J. A., and D. Tarulli, Dialogue, Difference and Voice in the Zone of Proximal Development, in *Theory and Psychology*, pp. 5 – 28, 1999.

Christensen, H., and J. Bardram, Supporting human activities – exploring activity-centered computing, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 107 – 116, 2002.

Clancey, W., Simulating activities: relating motive, deliberation, and attentive coordination, in *Cognitive Systems Research*, vol. 3, edited by T. Ziemke, pp. 471 – 499, Elsevier, 2002.

Convertino, G., T. P. Moran, and B. A. Smith, Activity patterns: managing, capturing, and reusing collaboration knowledge, in *Proceedings of the CSCW workshop on Awareness in Activity-Centric Groupware Design*, 2006.

Crowley, J., J. Coutaz, G. Rey, and P. Reignier, Perceptual components for Context Aware Computing, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 117 – 134, Springer-Verlag, 2002.

Czerwinski, M., E. Horvitz, and S. Wilhite, A diary study of task switching and interruptions, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 175 – 182, ACM Press, 2004.

Davis, F., Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, in *MIS Quarterly*, vol. 13(3), pp. 319 – 340, Regents of the University of Minnesota for the Management Information Systems Research Center, 1989.

Davis, M., S. King, N. Good, and R. Sarvas, From context to content: leveraging context to infer media metadata, in *Proceedings of the International Conference on Multimedia*, pp. 188 – 195, ACM Press, 2004.

Davydov, V. V., V. P. Zinchenko, and N. F. Talyzina, The problem of activity in the works of A. N. Leont'ev, in *Soviet Psychology*, vol. 21(4), edited by M. Cole, pp. 31 – 42, M. E. Sharpe, 1983.

Delin, J., J. Bateman, and P. Allen, A model of genre in document layout, in *Information Design Journal*, vol. 11(1), pp. 54 – 66, John Benjamins Publishing Company, 2000.

Dewey, J., *Human nature and conduct*, Dover Publications, Inc., 2002.

Dey, A., Providing architectural support for building context-aware applications, Ph.D. thesis, Georgia Tech, 2000.

Dey, A., Understanding and using context, in *Proceedings of the International Conference on Ubiquitous Computing*, vol. 5, pp. 4 – 7, Springer-Verlag, 2001.

Dey, A., D. Salber, and G. Abowd, A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications, in *Human-Computer Interaction Journal*, vol. 16, pp. 97 – 166, Springer-Verlag, 2001.

Doane, S., and Y. Sohn, ADAPT: a predictive cognitive model of user visual attention and action planning, in *User Modeling and User-Adapted Interaction*, vol. 10, pp. 1 – 45, Kluwer Academic Publishers, 2000.

Donath, J., Casual collaboration, in *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, pp. 490 – 496, IEEE Press, 1994.

Dourish, P., Seeking a foundation for context-aware computing, in *Human-Computer Interaction: Special issue on context-aware computing*, vol. 16, pp. 229 – 241, Taylor and Francis Group, 2001.

Dourish, P., What we talk about when we talk about context, in *Personal and Ubiquitous Computing*, vol. 8(1), pp. 19 – 30, Springer-Verlag, 2004.

Dourish, P., and V. Bellotti, Awareness and coordination in shared workspaces, in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 107 – 114, ACM Press, 1992.

Dourish, P., and S. Bly, Portholes: Supporting awareness in a distributed work group, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 541 – 547, ACM Press, 1992.

Dourish, P., V. Bellotti, W. Mackay, and C.-Y. Ma, Information and context: lessons from the study of two shared information systems, in *Proceedings of the Conference on Organizational Computing Systems*, pp. 42 – 51, ACM Press, 1993.

Dourish, P., J. Lamping, and T. Rodden, Building bridges: customisation and mutual intelligibility in shared category management, in *Proceedings of the International ACM SIGGROUP Conference*, pp. 11 – 20, ACM Press, 1999.

Duan, Y., and J. Canny, Protecting user data in ubiquitous computing environments: towards trustworthy environments, in *Privacy Enhancing Technologies*, pp. 167 – 185, Springer-Verlag, 2004.

Ducheneaut, N., The reproduction of open source software programming communities, Ph.D. thesis, University of California, Berkeley, 2003.

Dumais, S., E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, and D. Robbins, Stuff I've seen: a system for personal information retrieval and re-use, in *Proceedings of the International Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 72 – 79, ACM Press, 2003.

Eagle, N., Machine Perception and Learning of Complex Social Systems, Ph.D. thesis, Massachusetts Institute of Technology, Boston, MA, 2005.

Eliassi-rad, T., and J. Shavlik, A system for building intelligent agents that learn to retrieve and extract information, in *User Modeling and User-Adapted Interaction*, vol. 13, pp. 35 – 88, Kluwer Academic Publishers, 2003.

El'konin, D., Toward the problem of stages in the mental development of the child, in *Soviet Psychology*, vol. 10, edited by M. Cole, pp. 225 – 251, M. E. Sharpe, 1972.

Engeström, Y., *Learning by Expanding: An Activity - Theoretical Approach to Developmental Research*, Available electronically: http://lchc.ucsd.edu/MCA/Paper/Engestrom/expanding/toc.htm, 1987.

Engeström, Y., Expansive visibilization of work: an activity-theoretical perspective, in *Computer Supported Cooperative Work*, vol. 8, pp. 63 – 93, Kluwer Academic Publishers, 1999.

Engeström, Y., Activity theory as a framework for analyzing and redesigning work, in *Ergonomics*, vol. 43(7), pp. 960 – 974, Taylor and Francis, 2000.

Engeström, Y., and F. Blackler, On the life of the object, in *Organization*, vol. 12(3), pp. 307 – 220, Sage Publications, 2005.

Engeström, Y., R. Engeström, and T. Vähäaho, When the center does not hold: the importance of knotworking, in *Activity Theory and Social Practice*, pp. 345 – 374, Aarhus University Press, 1999a.

Engeström, Y., R. Miettinen, and R.-L. Punamäki (Eds.), *Perspectives on activity theory*, Cambridge University Press, 1999b.

Erickson, T., Making sense of computer-mediated communication (CMC): conversations as genres, CMC systems as genre ecologies, in *Proceedings of the Annual Hawaii International Conference on System Sciences (HICSS)*, IEEE Press, 2000.

Erickson, T., and M. Laff, The design of the 'Babble' timeline: a social proxy for visualizing group activity over time, in *CHI extended abstracts*, pp. 329 – 330, ACM Press, 2001.

Faaborg, A., and H. Lieberman, The goal-oriented web browser, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 751 – 760, ACM Press, 2006.

Feld, S., The focused organization of social ties, in *The American Journal of Sociology*, vol. 86(5), pp. 1015 – 1035, University of Chicago Press, 1981.

Fischer, G., User modeling in human-computer interaction, in *User Modeling and User-Adapted Interaction*, pp. 65 – 86, Kluwer Academic Publishers, 2001.

Fisher, D., Implementation of web places, 2000.

Fjeld, M., K. Lauche, M. Bichsel, F. Voorhorst, H. Krueger, and M. Rauterberg, Physical and virtual tools: activity theory applied to the design of groupware, in *Special Issue of Computer Supported Cooperative Work on Activity Theory and Design*, vol. 11(1-2), edited by B. Nardi and D. Redmiles, pp. 153 – 180, Kluwer Academic Publishers, 2002.

Fleming, M., and R. Cohen, A user modeling approach to determining system initiative in mixed-initiative AI systems, in *Proceedings of the International Conference on User Modeling*, edited by M. Bauer, P. Gmytrasiewicz, and J. Vassileva, pp. 54 – 63, Springer-Verlag, 2001.

Flores, F., M. Graves, B. Hartfield, and T. Winograd, Computer systems and the design of organizational interaction, in *ACM Transactions on Office Information Systems*, vol. 6(2), pp. 153 – 172, ACM Press, 1988.

Fogarty, J., S. Hudson, C. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. Lee, and J. Yang, Predicting human interruptibility with sensors, in *ACM Transactions on Computer-Human Interaction*, vol. 12(1), pp. 119 – 146, ACM Press, 2005a.

Fogarty, J., A. Ko, H. H. Aung, E. Golden, K. Tang, and S. Hudson, Examining task engagement in sensor-based statistical models of human interruptibility, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 331 – 340, ACM Press, 2005b.

Franklin, D., J. Budzik, and K. Hammond, Plan-based interfaces: keeping track of user tasks and acting to cooperate, in *Proceedings of the International Conference on Intelligent User Interfaces*, pp. 79 – 86, ACM Press, 2002.

Freedman, A., and P. Medway (Eds.), *Genre and the new rhetoric*, Taylor and Francis Group, 1994.

Gal'perin, P., On the notion of internalization, in *Soviet Psychology*, vol. 5(3), edited by M. Cole, pp. 28 – 33, M. E. Sharpe, 1967.

Garfinkel, H., and H. Sacks, On formal structures of practical actions, in *Theoretical Sociology: perspectives and developments*, edited by J. McKinney and E. Tiryakian, pp. 337 – 266, Appleton-Century-Crofts, 1970.

Garrison, J., An introduction to Dewey's theory of functional "Trans-action": an alternative paradigm for activity theory, in *Mind, Culture, and Activity*, vol. 8(4), pp. 275 – 296, University of California Press, 2001.

Giddens, A., *The Constitution of Society: Outline of the Theory of Structuration*, University of California Press, 1984.

Goffman, E., *The presentation of self in everyday life*, Doubleday, 1959.

Goffman, E., *Interaction Ritual: essays on face-to-face behavior*, Doubleday, 1967.

Goffman, E., *Frame Analysis: an essay on the organization of experience*, Northeastern University Press, 1974.

González, V., and G. Mark, "Constant, constant, multi-tasking craziness": managing multiple working spheres, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 113 – 120, ACM Press, 2004.

González, V., and G. Mark, Managing currents of work: multi-tasking among multiple collaborations, in *Proceedings of the 8th European Conference of Computer-supported Cooperative Work*, Kluwer Academic Publishers, 2005.

Goodwin, C., and A. Duranti, Rethinking context: an introduction, in *Rethinking context: language as an interactive phenomenon*, vol. Chpt. 1, Cambridge University Press, 1992.

Greenberg, S., Peepholes: Low cost awareness of one's community, in *CHI Conference Companion*, pp. 206 – 207, ACM Press, 1996.

Greenberg, S., Context as a Dynamic Construct, in *Human-Computer Interaction: Special issue on context-aware computing*, vol. 16, pp. 257 – 268, Taylor and Francis Group, 2001.

Gross, T., W. Wirsam, and W. Graether, AwarenessMaps: visualizing awareness in shared workspaces, in *CHI extended abstracts*, pp. 784 – 785, ACM Press, 2003.

Grudin, J., Desituating Action: Digital Representation of Context, in *Human-Computer Interaction: Special issue on context-aware computing*, vol. 16, pp. 269 – 286, Taylor and Francis Group, 2001.

Gwizdka, J., TaskView: design and evaluation of a task-based email interface, in *Proceedings of the Conference of the Centre for Advanced Studies on Collaborative Research*, pp. 136 – 145, IBM Press, 2002.

Hagtvet, B., On the dialogical basis of meaning: inquiries into Ragnar Rommetveit's writings on language, thought, and communication, in *Mind, Culture, and Activity*, vol. 10(3), pp. 186 – 204, University of California Press, 2003.

Halverson, C., Activity theory and distributed cognition: or what does CSCW need to do with theories?, in *Special Issue of Computer Supported Cooperative Work on Activity Theory and Design*, vol. 11(1-2), edited by B. Nardi and D. Redmiles, pp. 243 – 267, Kluwer Academic Publishers, 2002.

Hansen, K., B. MacIntyre, E. Mynatt, J. Tullio, and S. Voida, Hypermedia in the kimura system: Using spatial, temporal, and navigational relationships to support multitasking and background awareness, in *Poser Presentation at the ACM Conference on Hypertext and Hypermedia*, 2001.

Hardless, C., and U. Nulden, Visualizing learning activities to support tutors, in *CHI extended abstracts*, pp. 312 – 313, ACM Press, 1999.

Hearst, M., Multi-paragraph segementation of expository text, in *Proceedings of the Association for Computational Linguistics*, 1994.

Hearst, M., *User interfaces and Visualization*, chap. 10, pp. 257 – 324, Addison-Wesley Longman Publishing Co., Inc., 1999.

Henderson, D. A., and S. Card, Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface, *ACM Trans. Graph.*, *5(3)*, 211 – 243, 1986.

Hill, A., S. Song, A. Dong, and A. Agogino, Identifying shared understanding in design teams using document analysis, in *Proceedings of the International Conference on Design Theory and Methodology*, 2001.

Hinds, P. J., and D. E. Baily, Out of sight, out of sync: understanding conflict in distributed teams, in *Organization Science*, vol. 14(6), pp. 615 – 632, Informs, 2003.

Hinds, P. J., and M. Mortensen, Understanding conflict in geographically distributed teams: the moderating effects of shared identity, shared context, and spontaneous communication, in *Organization Science*, vol. 16(3), pp. 290 – 307, Informs, 2005.

Hollan, J., E. Hutchins, and D. Kirsh, Distributed cognition: toward a new foundation for human-computer interaction research, in *ACM Transactions on Computer-Human Interaction*, vol. 7(2), pp. 174 – 196, ACM Press, 2000.

Hong, J., and J. Landay, A Context/Communication Information Agent, in *Personal and Ubiquitous Computing*, vol. 5(1), pp. 78 – 81, Springer-Verlag, 2001a.

Hong, J., and J. Landay, An infrastructure approach to context-aware computing, in *Human-Computer Interaction Journal*, vol. 16(2-3), 2001b.

Honkaranta, A., Evaluating the 'genre lens' for analyzing requirements for content assembly, in *CAiSE/IFIP8.1 EMMSAD workshop*, 2003.

Horvitz, E., and T. Paek, A computational architecture for conversation, in *Proceedings of the International Conference on User Modeling*, edited by J. Kay, pp. 201 – 210, Springer-Verlag, 1999.

Horvitz, E., J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, The Lumière project: bayesian user modeling for inferring the goals and needs of software users, in *Proceedings of 14th Conference on Uncertainty in Artificial Intelligence*, pp. 256 – 265, Morgan Kaufmann Publishers, 1998.

Hoyer, P. O., Non-negative matrix factorization with sparseness constraints, in *Journal of Machine Learning Research*, vol. 5, pp. 1457 – 1469, ACM Press, 2004.

Huang, E., and E. Mynatt, Tailoring public displays for small, co-located groups, in *Companion Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 2002.

Huang, E., J. Tullio, T. Costa, and J. McCarthy, Promoting awareness of work activities through peripheral displays, in *Proceedings of CHI*, ACM Press, 2002.

Hudson, S., B. John, K. Knudsen, and M. Byrne, A tool for creating predictive performance models from user interface demonstrations, in *Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 93 – 102, ACM Press, 1999.

Hutchins, E., *Cognition in the wild*, MIT Press, 1995.

Hutchins, E., and T. Klausen, Distributed cognition in an airline cockpit, in *Cognition and Communication at Work*, edited by Y. Engeström and D. Middleton, Cambridge University Press, 1996.

Huỳnh, T., M. Fritz, and B. Schiele, Discovery of Activity Patterns using Topics Models, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 10 – 19, ACM Press, 2008.

Intille, S., Designing a home of the future, in *IEEE Pervasive Computing*, vol. 1(2), pp. 76 – 82, IEEE Press, 2002.

Isaacs, E., J. Tang, and T. Morris, Piazza: A desktop environment supporting impromptu and planned interactions, in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 315 – 324, ACM Press, 1996.

James, W., *Pragmatism, a new name for some old ways of thinking, and The Meaning of Truth, a sequel to Pragmatism*, Harvard University Press, 1975.

John, B., and D. Kieras, Using GOMS for user interface design and evaluation: which technique?, in *ACM Transactions on Computer-Human Interaction*, vol. 3(4), pp. 287 – 319, ACM Press, 1996.

John, B., A. Vera, M. Matessa, M. Freed, and R. Remington, Automating CPM-GOMS, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 147 – 154, ACM Press, 2002.

Kaptelinin, V., UMEA: translating interaction histories into project contexts, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 353 – 360, ACM Press, 2003.

Karjalainen, A., and A. Salminen, Bridging the gap between hard and soft information genres, in *Proceedings of IRMA*, Idea Group Publishing, 2000.

Karlgren, J., and D. Cutting, Recognizing text genres with simple metrics using discriminant analysis, in *Computational Linguistics*, pp. 1071 – 1075, 1994.

Kobsa, A., Generic user modeling systems, in *User Modeling and User-Adapted Interaction*, edited by M. Bauer, P. Gmytrasiewicz, and J. Vassileva, pp. 49 – 63, Kluwer Academic Publishers, 2001.

Korzybski, A., *Science and Sanity: an Introduction to Non-Aristotelian Systems and General Semantics*, 1933.

Kumaran, G., and J. Allan, Text classification and named entities for new event detection, in *Proceedings of the International Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 297 – 304, ACM Press, 2004.

Küpper, D., and A. Kobsa, User-tailored plan generation, in *Proceedings of the International Conference on User Modeling*, edited by J. Kay, pp. 45 – 54, Springer-Verlag, 1999.

Küpper, D., and A. Kobsa, Generating and presenting user-tailored plans, in *Proceedings of the International Conference on Intelligent User Interfaces*, pp. 198 – 199, ACM Press, 2002.

Kuutti, K., and L. Bannon, Searching for unity among diversity: exploring the "interface" concept, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 263 – 268, ACM Press, 1993.

Lakoff, G., and M. Johnson, *Metaphors we live by*, University of Chicago Press, 1980.

Landauer, T., P. Foltz, and D. Laham, Introduction to Latent Semantic Analysis, in *Discourse Processes*, vol. 25, pp. 259 – 284, 1998.

Latour, B., Visualization and cognition: thinking with eyes and hands, in *Knowledge and Society: Studies in the Sociology of Culture Past and Present*, vol. 6, pp. 1 – 40, JAI Press, Inc., 1986.

Latour, B., Mixing humans and nonhumans together: the sociology of a door-closer, in *Ecologies of Knowledge: Work and Politics in Science and Technology*, edited by S. Star, pp. 257 – 277, SUNY Press, 1995.

Latour, B., *Pandora's Hope: essays on the reality of science studies*, Harvard University Press, 1999.

Lave, J., and E. Wenger, *Situated learning: legitimate peripheral participation*, Cambridge University Press, 1991.

Law, J., The heterogeneity of texts, in *Mapping the dynamics of science and technology: sociology of science in the real world*, edited by M. Callon, J. Law, and A. Rip, pp. 67 – 83, Macmillan Press, 1986a.

Law, J., On the methods of long distance control: vessels, navigation, and the Portuguese route to India, in *Power, Action and Belief: a new sociology of knowledge?*, edited by J. Law, pp. 234 – 263, Routledge and Kegan Paul Books Ltd, 1986b.

Law, J., Notes on the theory of the actor-network: ordering, strategy and heterogeneity, in *Systems Practice*, vol. 5(4), pp. 379 – 393, 1992.

Lee, D., and S. Seung, Algorithms for non-negative matrix factorization, in *Proceedings of NIPS Conference*, pp. 556 – 562, 2001.

Lemon, O., and A. Gruenstein, Multithreaded context for robust conversational interfaces: context-sensitive speech recognition and interpretation of corrective fragments, in *ACM Transactions on Computer-Human Interaction*, vol. 11(3), pp. 241 – 267, ACM Press, 2004.

Leontiev, A. N., *Activity and consciousness*, Progress Publishers, 1977.

Leontiev, A. N., *Activity, consciousness, and personality*, Available electronically: http://lchc.ucsd.edu/MCA/Paper/leontev/, 1978.

Leontiev, A. N., *Problems of the Development of the Mind*, Progress Publishers, 1981.

Levine, D. N. (Ed.), *Georg Simmel: On Individuality and Social Forms*, University of Chicago Press, 1971.

Li, T., S. Ma, and M. Ogihara, Document clustering via adaptive subspace iteration, in *Proceedings of the International Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 218 – 225, ACM Press, 2004.

Liao, L., D. Fox, and H. Kautz, Learning and inferring transportation routines, in *Proceedings of the conference on Artificial Intelligence AAAI*, pp. 348 – 353, AAAI Press, 2001.

Lieberman, H., Letizia: an agent that assists web browsing, in *International Joint Conference on Artificial Intelligence*, Cambridge University Press, 1995.

Linton, F., D. Joy, and H.-P. Schaefer, Building user and expert models by long-term observation of application usage, in *Proceedings of the International Conference on User Modeling*, edited by J. Kay, pp. 129 – 138, Springer-Verlag, 1999.

Liu, S., F. Liu, C. Yu, and W. Meng, An effective approach to document retrieval via utilizing WordNet and recognizing phrases, in *Proceedings of the International Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 266 – 272, ACM Press, 2004.

Luria, A., *Cognitive Development: Its Cultural and Social Foundations*, Harvard University Press, 1976.

MacIntyre, B., E. Mynatt, S. Voida, K. Hansen, J. Tullio, and G. Corso, Support for multitasking and background awareness using interactive peripheral displays, in *Proceedings of the ACM symposium on User Interface Software and Technology*, pp. 41 – 50, ACM Press, 2001.

Maglio, P., C. Campbell, R. Barrett, and T. Selker, An architecture for developing attentive information systems, in *Knowledge-Based Systems*, vol. 14, pp. 103 – 110, Elsevier Science, 2001.

Magnini, B., and C. Strapparava, Improving user modeling with content-based techniques, in *Proceedings of the International Conference on User Modeling*, edited by M. Bauer, P. Gmytrasiewicz, and J. Vassileva, pp. 74 – 83, Springer-Verlag, 2001.

Mark, G., V. González, and J. Harris, No task left behind?: examining the nature of fragmented work, in *Proceeding of the SIGCHI Conference*, pp. 321 – 330, ACM Press, 2005.

Maturana, H., Autopoiesis, structural coupling and cognition.

McCrickard, D. S., C. M. Chewar, J. Somervell, and A. Ndiwalana, A model for notification systems evaluation – assessing user goals for multitasking activity, in *ACM Transactions on Computer-Human Interaction*, vol. 10(4), pp. 312 – 338, ACM Press, 2003.

Miettinen, R., The riddle of things: activity theory and actor-network theory as approaches to studying innovations, in *Mind, Culture, and Activity*, vol. 6(3), pp. 170 – 195, University of California Press, 1999.

Miettinen, R., Artifact mediation in Dewey and in cultural-historical activity theory, in *Mind, Culture, and Activity*, vol. 8(4), pp. 297 – 308, University of California Press, 2001.

Millen, D., M. Muller, W. Geyer, E. Wilcox, and B. Brownholtz, Patterns of media use in an activity-centric collaborative environment, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 879 – 888, ACM Press, 2005.

Mohammed, S., and B. C. Dumville, Team mental models in a team knowledge framework: expanding theory and measurement acorss disciplinary boundaries, in *Journal of Organizational Behavior*, vol. 22, pp. 89 – 106, John Wiley and Sons, 2001.

Mohammed, S., R. Klimoski, and J. Rentsch, The measurement of team mental models: we have no shared schema, in *Organizational Research Methods*, vol. 3(2), pp. 123 – 165, Sage Publications, 2000.

Moran, T., Unified Activity Management: Explicitly Representing Activity in Work-Support Systems, in *Proceedings of the ECSCW Workshop on Activity: From a Theoretical to a Computational Construct*, 2005.

Moran, T., and P. Dourish, Introduction to special issue, in *Human-Computer Interaction: Special issue on context-aware computing*, vol. 16, pp. 87 – 95, Taylor and Francis Group, 2001.

Morteo, R., V. González, J. Favela, and G. Mark, Sphere Juggler: Fast Context Retrieval in Support of Working Spheres, in *Proceedings of the Mexican International Conference in Computer Science*, pp. 361 – 367, IEEE Press, 2004.

Mukherjee, R., N. Sajja, and S. Sen, A movie recommendation system – an application of voting theory in user modeling, in *User Modeling and User-Adapted Interaction*, vol. 13, pp. 5 – 33, Kluwer Academic Publishers, 2003.

Muller, M., W. Geyer, B. Brownholtz, E. Wilcox, and D. Millen, One-hundred days in an activity-centric collaboration environment based on shared objects, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 375 – 382, ACM Press, 2004.

Mynatt, E., The writing on the wall, in *Proceedings of the Conference on Human-Computer Interaction (INTERACT)*, pp. 196 – 204, IOS Press, 1999.

Nardi, B. (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction*, MIT Press, 1996.

Nardi, B., and D. Redmiles (Eds.), vol. 11(1-2), Kluwer Academic Publishers, 2002.

Nardi, B., S. Whittaker, and E. Bradner, Interaction and outeraction: instant messaging in action, in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 79 – 88, ACM Press, 2000.

Nardi, B., S. Whittaker, and H. Schwarz, NetWORKers and their activity in intensional networks, in *Special Issue of Computer Supported Cooperative Work on Activity Theory and Design*, vol. 11(1-2), edited by B. Nardi and D. Redmiles, Kluwer Academic Publishers, 2002.

Narendra, N. C., and T. P. Moran, Improving activity awareness in artifact-centric business processes, in *Proceedings of the CSCW workshop on Awareness in Activity-Centric Groupware Design*, 2006.

Oliver, N., E. Horvitz, and A. Garg, Layered representations for human activity recognition, in *Proceedings of IEEE International Conference on Multimodal Interfaces*, IEEE Press, 2002.

Palen, L., and P. Dourish, Unpacking "privacy" for a networked world, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 129 – 136, ACM Press, 2003.

Paolillo, J., The virtual speech community: social network and language variation on IRC, in *Journal of Computer-Mediated Communication*, vol. 4(4), 1999.

Patterson, D., L. Liao, D. Fox, and H. Kautz, Inferring high-level behavior from low-level sensors, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 73 – 89, Springer-Verlag, 2003.

Pedersen, E., People presence or room activity supporting peripheral awareness over distance, in *CHI extended abstracts*, pp. 283 – 284, ACM Press, 1998.

Pedersen, E., and T. Sokoler, AROMA: abstract representation of presence supporting mutual awareness, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 51 – 58, ACM Press, 1997.

Perkowitz, M., M. Philipose, K. Fishkin, and D. Patterson, Mining models of human activities from the web, in *Proceedings of the World Wide Web Conference*, ACM Press, 2004.

Petrelli, D., A. D. Angeli, and G. Convertino, A user-centered approach to user modeling, in *Proceedings of the International Conference on User Modeling*, edited by J. Kay, pp. 255 – 264, Springer-Verlag, 1999.

Pianta, E., L. Bentivogli, and C. Girardi, MultiWordNet: developing an aligned multilingual database, in *Proceedings of the International Conference on Global WordNet*, 2002.

Proctor, R., A. Goldenberg, E. Davenport, and A. Mckinlay, Genres in Support of Collaborative Information Retrieval in the Virtual Library, in *Interacting with Computers*, vol. 10(2), pp. 157 – 175, 1998.

Rattenbury, T., D. Nafus, and K. Anderson, Plastic: A Metaphor for Integrated Technologies, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 232 – 241, ACM Press, 2008.

Reiss, S., Multifaceted nature of intrinsic motivation: the theory of 16 basic desires, in *Review of General Psychology*, vol. 8(3), pp. 179 – 193, Education Publishing Foundation, 2004.

Rhodes, B., Just-In-Time Information Retrieval, Ph.D. thesis, Massachusetts Institute of Technology, 2000.

Robertson, G., M. van Dantzich, D. Robbins, M. Czerwinski, K. Hinckley, K. Risden, D. Thiel, and V. Gorokhovsky, The Task Gallery: a 3D window manager, in *Proceedings of the ACM Conference on Human Factors in Computing Systems (SIGCHI)*, pp. 494 – 501, ACM Press, 2000.

Russell, D., Rethinking genre in school and society: an activity theory analysis, in *Written Communication*, vol. 14(4), pp. 504 – 554, ACM Press, 1997.

Ryder, M., What is actor-network theory?

Sawhney, N., S. Wheeler, and C. Schmandt, Aware Community Portals: Shared Information Appliances for Transitional Spaces, in *Personal and Ubiquitous Computing*, vol. 5(1), pp. 66 – 70, Springer-Verlag, 2001.

Scribner, S., Vygotsky's uses of history, in *Culture, Communication, and Cognition: Vygotskian Perspectives*, edited by J. Wertsch, pp. 119 – 145, Cambridge University Press, 1985.

Scribner, S., and M. Cole (Eds.), *The Psychology of Literacy*, Harvard University Press, 1981.

Searle, J., *Speech acts: an essay in the philosophy of language*, Cambridge University Press, 1969.

Shaw, M., and B. Gaines, Modeling the social practices of users in internet communities, in *Proceedings of the International Conference on User Modeling*, edited by J. Kay, pp. 77 – 86, Springer-Verlag, 1999.

Simon, H., *The Sciences of the Artificial*, MIT Press, 1996.

Smith, G., P. Baudisch, G. Robertson, M. Czerwinski, B. Meyers, D. Robbins, and D. Andrews, GroupBar: The TaskBar Evolved, in *Proceedings of OZCHI*, 2003.

Soeffner, H.-G., The art of experienced analysis: Anselm Strauss and his theory of action, in *Mind, Culture, and Activity*, vol. 2(1), pp. 29 – 32, University of California Press, 1995.

Spinuzzi, C., Genre ecologies: an open-system approach to understanding and constructing documentation, in *ACM Journal of Computer Documentation*, vol. 24(3), pp. 169 – 181, ACM Press, 2000.

Spinuzzi, C., Software development as mediated activity: applying three analytical frameworks for studying compound mediation, in *Proceedings of the International Conference on Computer Documentation*, pp. 58 – 67, ACM Press, 2001.

Spinuzzi, C., Modeling genre ecologies, in *Proceedings of SIGDOC*, pp. 200 – 207, ACM Press, 2002.

Spinuzzi, C., *Tracing genres through organizations: a sociocultural approach to information design*, MIT Press, 2003.

Spiteri, M., and J. Bates, Supporting storage and retrieval of computer and human activity, in *Proceedings of the European workshop on Support for Composing Distributed Applications*, pp. 268 – 275, ACM Press, 1998.

Stalder, F., Actor-network-theory and communication networks: toward convergence, 1997.

Star, S., Listening for connections, in *Mind, Culture, and Activity*, vol. 2(1), pp. 12 – 17, University of California Press, 1995.

Star, S., *Working together: symbolic interactionism, activity theory, and information systems*, chap. 13, pp. 296 – 318, Cambridge University Press, 1997.

Strauss, A., Research issues and an interactionist theory of action, in *Mind, Culture, and Activity*, vol. 2(1), pp. 18 – 28, University of California Press, 1995.

Stumpf, S., X. Bao, A. Dragunov, T. Dietterich, J. Herlocker, K. Johnsrude, L. Li, and J. Shen, The TaskTracer system, in *Proceedings of the conference on Artificial Intelligence AAAI: Intelligent Systems Demonstrations*, AAAI Press, 2005.

Suchman, L., *Plans and situated actions: the problem of human-machine communication*, Cambridge University Press, 1987.

Suchman, L., J. Blomberg, J. Orr, and R. Trigg, Reconstructing technologies as social practice, in *American Behavioral Scientist*, vol. 43(3), pp. 392 – 408, Sage Publications, 1999.

Svanaes, D., Context-Aware Technology: A Phenomenological Perspective, in *Human-Computer Interaction: Special issue on context-aware computing*, vol. 16, pp. 379 – 400, Taylor and Francis Group, 2001.

Tarasewich, P., C. Campbell, T. Xia, and M. Dideles, Evaluation of visual notification cues for ubiquitous computing, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 349 – 366, Springer-Verlag, 2003.

Teo, G., Getting the right information to the right person, in *Proceedings of the International Conference on User Modeling*, edited by M. Bauer, P. Gmytrasiewicz, and J. Vassileva, pp. 257 – 259, Springer-Verlag, 2001.

Terry, M., and E. Mynatt, Side views: persistent, on-demand previews for open-ended tasks, in *Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 71 – 80, ACM Press, 2002.

Torre, I., Goals, tasks and application domains as the guidelines for defining a framework for user modeling, in *Proceedings of the International Conference on User Modeling*, edited by M. Bauer, P. Gmytrasiewicz, and J. Vassileva, pp. 260 – 262, Springer-Verlag, 2001.

Trevor, J., D. Hilbert, and B. Schilit, Issues in personalizing shared ubiquitous devices, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 56 – 72, Springer-Verlag, 2002.

Truong, K., G. Abowd, and J. Brotherton, Who, what, when, where, how: design issues of capture and access applications, in *Proceedings of the International Conference on Ubiquitous Computing*, pp. 209 – 224, Springer-Verlag, 2001.

Turoff, M., S. Hiltz, M. Bieber, J. Fjermestad, and A. Rana, Collaborative discourse structures in computer mediated group communications, in *Journal of Computer-Mediated Communication*, vol. 4(4), 1999.

Tversky, B., J. B. Morrison, and J. Zacks, On bodies and events, in *The imitative mind: Development, evolution and brain bases*, edited by A. Meltzoff and W. Prinz, pp. 221 – 232, Cambridge University Press, 2002.

van Dantzich, M., D. Robbins, E. Horvitz, and M. Czerwinski, Scope: providing awareness of multiple notifications at a glance, in *Proceedings of the ACM Conference on Advanced Visual Interfaces*, ACM Press, 2002.

Voida, S., E. Mynatt, B. MacIntyre, and G. Corso, Integrating virtual and physical context to support knowledge workers, in *IEEE Pervasive Computing*, vol. 1(3), pp. 73 – 79, IEEE Press, 2002.

Vygotsky, L., Play and it's role in the mental development of the child, in *Soviet Psychology*, vol. 5(3), edited by M. Cole, pp. 6 – 18, M. E. Sharpe, 1967.

Vygotsky, L., *Mind in Society: the development of higher psychological processes*, Harvard University Press, 1978.

Vygotsky, L., *Thought and Language*, MIT Press, 1985.

Wasserman, S., and K. Faust, *Social Network Analysis: methods and applications*, Cambridge University Press, 1994.

Webb, G., M. Pazzani, and D. Billsus, Machine learning for user modeling, in *User Modeling and User-Adapted Interaction*, vol. 11, pp. 19 – 29, Kluwer Academic Publishers, 2001.

Wells, G., The role of dialogue in activity theory, in *Mind, Culture, and Activity*, vol. 9(1), pp. 43 – 66, University of California Press, 2002.

Wenger, E., Communities of practice and social learning systems, in *Organization*, vol. 7(2), pp. 225 – 246, Sage Publications, 2000.

Wertsch, J. (Ed.), *The concept of activity in soviet psychology*, M. E. Sharpe, Inc., 1979.

Wertsch, J., The semiotic mediation of mental life: L. S. Vygotsky and M. M. Bakhtin, in *Semiotic mediation: sociocultural and psychological perspectives*, edited by E. Mertz and R. Parmentier, pp. 49 – 71, Academic Press, Inc., 1985a.

Wertsch, J., *Vygotsky and the social formation of mind*, Harvard University Press, 1985b.

Wertsch, J., *Voices of the mind: a sociocultural approach to mediated action*, Harvard University Press, 1991.

Wertsch, J., *Mind as action*, Oxford University Press, 1998.

Wiener, N., *Cybernetics or Control and Communication in the Animal and the Machine*, The Technology Press, 1948.

Winograd, T., Architectures for context, in *Human-Computer Interaction Journal: special issue on Context-Aware Computing*, vol. 16, pp. 401 – 419, 2001.

Winograd, T., and F. Flores, *Understanding computers and cognition: a new foundation for design*, Addison-Wesley, 1986.

Xu, W., X. Liu, and Y. Gong, Document clustering based on non-negative matrix factorization, in *Proceedings of the International Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 267 – 273, ACM Press, 2003.

Yoshioka, T., and G. Herman, Coordinating information using genres, 2000.

Yoshioka, T., J. Yates, and W. Orlikowski, Community-based interpretive schemes: exploring the use of cyber meetings within a global organization, 2000.

Yoshioka, T., G. Herman, J. Yates, and W. Orlikowski, Genre taxonomy: a knowledge repository of communicative actions, in *ACM Transactions on Information Systems*, vol. 19(4), pp. 431 – 456, ACM Press, 2001.

Zager, D., Collaboration as an activity: coordinating with pseudo-collective objects, in *Special Issue of Computer Supported Cooperative Work on Activity Theory and Design*, vol. 11(1-2), edited by B. Nardi and D. Redmiles, Kluwer Academic Publishers, 2002.

Zinchenko, V., Vygotsky's ideas about units for the analysis of mind, in *Culture, Communication, and Cognition: Vygotskian Perspectives*, edited by J. Wertsch, pp. 94 – 118, Cambridge University Press, 1985.

Zukerman, I., and D. Albrecht, Predictive statistical models for user modeling, in *User Modeling and User-Adapted Interaction*, vol. 11, pp. 5 – 18, Kluwer Academic Publishers, 2001.

Zukerman, I., and D. Litman, Natural language processing and user modeling: synergies and limitations, in *User Modeling and User-Adapted Interaction*, vol. 11, pp. 129 – 158, Kluwer Academic Publishers, 2001.