Cognitive Representation of Common Ground in User Interfaces

Derek Brock and J. Gregory Trafton*

U.S. Naval Research Laboratory, Washington, DC, USA

Abstract. We argue that Clark's theory of participant representations of common ground in joint activities between people is relevant to the design of human-computer interaction. Features of common ground can be shown to exist in current interaction models suggesting that computers as sophisticated information processors can be legitimately construed as participants in the interactive process. Software, though, is rarely designed to explain itself or to demonstrate a knowledgeable awareness of the user's concerns. We describe an effort to provide a principled basis for this capacity in the form of an embedded cognitive simulation representing an application's task and user related common ground. **Keywords:** *common ground, cognitive modeling, joint activities, task model tracing*

Modern user interfaces continue to have fundamental shortcomings, many of which are directly related to the limits of their abilities to communicate well with users. For instance, users are rarely given the direct means to find out what a program knows about its domain of activity or, more importantly, what it has done in the course of a series of interactions that may be unclear or confusing. As another example, applications rarely devote functionality to the job of anticipating users' concerns or recognizing what they are trying to do. User interfaces are seldom designed to present such information, much less to keep track of it. One consequence is that users are often on their own when it comes to puzzling out a program's various features and abilities. In complex, feature-rich systems, this can be a barrier to successful or timely task performance.

Communicative shortcomings in user interfaces are easily identified but are notoriously difficult to solve in a comprehensive sense. One reason for this may be the enormous complexity of such problems and another may be the general lack of a coherent theoretical framework. In his 1996 book *Using language,* Herbert Clark (1996b) advances an insightful proposal about the nature of communication between people. Specifically, he argues that "language use is really a form of *joint action*" in which participants – people – act with intentions and in coordination with each other to accomplish goals that are part of their broader ends in *joint activities*. Clark describes joint activities as a basic category that encompasses all participatory circumstances in which conventional language plays a role. More to the point, he notes, "If we take language use to include such communicative acts as eye gaze, iconic gestures, pointing, smiles, and head nods – and we must – then all joint activities rely on language use." To coordinate joint activities, some form of signaling is required. For Clark, language in its linguistic sense is simply one of many possible signaling systems, some highly organized and others spontaneously improvised.

Clark argues that all joint activities advance through the accumulation of common ground – the knowledge, beliefs, and suppositions that participants believe they share about an activity.

* This work was conducted at the U.S. Naval Research Laboratory with sponsorship from the Office of Naval Research.

Common ground has been widely studied in discourse, so Clark's broader claim that it is an essential underpinning of all joint activities follows naturally from his argument that language use is a form of joint action. At any moment, an individual's common ground can be thought of as being made up of three parts, all three of which have important empirical characteristics. For instance, in many joint activities, a telling aspect of *initial common ground* is each participant's knowledge of relevant conventions – standard ways of doing things and a sense of what is expected. In their representation of the *current state of a joint activity* people often depend greatly upon "external representations" – features of and in the immediate physical environment that are taken to be germane to the activity. And, as people keep track of the *public events so far* in a joint activity, they do so mostly in the form of annotated records and characterize events colloquially in terms of their significance.

The basic thesis of this paper is that much of Clark's theory of the nature of joint activities is applicable to the design of human-computer interaction (see also Clark, 1996a). Models of communication between humans have often been proposed as a basis for models of human-computer interaction (for a survey, see Pérez-Quiñones, 1996), but little or no work has explicitly studied the idea of using common ground as both a way to interpret the demands of the interactive process and as an analytical basis for its design. Computers, though, as sophisticated information and display processors, inherently use programmatic representations of meaning and process for the coordination of interactions and presentation of information. These representations are functionally elements of common ground for the purpose of participatory activities with people. Accordingly, we find that human-computer interaction can be usefully viewed as a form of joint activity – albeit with its own language (signaling system) and affordances (cf. Clark's use of the term "availability") – well suited to the application of communicative principles identified in Clark' theory.

In joint activities, each participant's purpose in keeping track of common ground is to know enough of what the other participants know to jointly succeed at coordinating the activity itself. In other words, a participant in a joint activity presumes to model both the activity itself and the understanding maintained by the activity's other participants. Between a user and a computer, the same ideas apply. Common ground's full potential is achieved only to the extent that the computer can successfully simulate two skills that are ultimately cognitive in nature: the skill of keeping track of both the activity itself and the user's understanding of the activity, and the skill of making use of this representation to coordinate its participation in the joint activity.

The effort we describe here focuses cognitively on the accumulation of common ground in joint activities between users and computers. Our working system is an application embedded with a cognitively modeled representation of its domain-related common ground. The application is a complex, non-trivial, resource allocation task in a probabilistic military setting. The embedded representation of common ground is modeled in ACT-R (Anderson and Lebiére, 1998). Our working strategy has not been to computationally model common ground in a conversational sense but to work at the higher, more schematic level of a task analysis. By this, we simply mean we have modeled information about the task (the joint activity) that the application should be aware of and that may be of use to the user.

Our application domain is very roughly that of a military mission planning tool. Many of the factors the user must consider in planning a mission are interdependent, and to further complicate matters there are several probabilistic risks of failure in carrying out a mission. The application user interface utilizes a standard point-and-click paradigm and is composed of several dialog-box style windows in which the user can review and select destinations, equip and allocate tanks, and subsequently evaluate the success or failure of a mission.

Our cognitive modeling effort utilizes a model tracing paradigm (Anderson, 1987) we call "task model tracing." The emergence of unified theories of cognition (Newell, 1990) in cognitive science has given researchers an invaluable tool for exploring the ramifications of simulated cognition in applied settings such as user interfaces. ACT-R's adaptive processing strengths with regard to memory and learning (Anderson and Lebiére, 1998) have proven to be a good fit with our line of inquiry. In task model tracing, as the application responds to user input, it also drives the cognitive model, effectively "tracing" each interaction. As the model runs, it reasons independently about changes in the state of the system and implications of the user's input.

Applications are rarely if ever designed to explain themselves. By cognitively modeling an application's task domain, we can provide a requisite basis for implementing this capacity. Our approach has been to represent knowledge gleaned from an analysis of portions of our task. ACT-R provides us with a principled means for accounting for salience in a non-deterministic world as well as a proven framework for the job of representing a task analysis in the user's terms (Anderson, 1987). These features constitute a form of user modeling. The system provides the user with a direct means for consulting the application about the task (confirming the common ground) by augmenting the user interface with an additional dialog-box style window in which the user can request a "situation analysis" with a button click. This produces an annotated report of the current context and an advisory list of situationally relevant information and actions the user may wish to take next. The system is also able to carry out any of the actions it proposes at the user's discretion. When an action requires the system to make a choice for the user, ACT-R's theory of memory retrieval correctly identifies the most salient choice based on the composition of both its recency and number of mentions. The model also makes note of when it or the user does an action and demonstrates sensitivity to the meaning and occurrence of previous interactions though its choice of terms.

How then does our work demonstrate an advance in the joint activity of human-computer interaction? Our model instantiates a cognitively-based, participatory representation of each of the three parts of common ground for the application at any moment. Its initial common ground – the underlying representations of the task and the user's likely concerns – inform the application's accumulation of common ground during the task. And through its presentations, the system's representation of the current state of the activity serves to reinforce the user's own accumulating representation with its advisory knowledge and selectively annotated record of the task's public events so far.

References

Anderson, J. R., and Lebiére, C. (1998). *The Atomic Components of Thought.* Lawrence Erlbaum.

Anderson, J. R. (1987). Production system, learning, and tutoring. In *Production System Models of Learning and Development*. MIT Press.

Clark, H. H. (1996a). Arranging to do things with others. In *Conference Companion of the Conference on Human Factors in Computing Systems—CHI'96*. Association for Computing Machinery.

Clark, H. H. (1996b). *Using Language.* Cambridge University Press.

Newell, A. (1990). *Unified Theories of Cognition*. Harvard University Press.

Pérez-Quiñones, M. A. (1996). *Conversational Collaboration in User-Initiated Interruption and Cancellation Requests*. Ph.D. Dissertation, The George Washington University.