Contextual Computer Support for Human Activity

Donald J. Patterson[1], Ken Fishkin[2], Dieter Fox[1], Henry Kautz[1], Mike Perkowitz[2], and Matthai Philipose[2]

[1] University Of Washington,
Department of Computer Science and Engineering,
{djp3,fox,kautz}@cs.washington.edu
http://www.cs.washington.edu
[2] Intel Research
Seattle, WA USA
kenneth.p.fishkin@intel.com,{mikep,matthai}@intel-research.net
http://seattleweb.intel-research.net/

The modern knowledge worker has become very adept at working with desktop computers through familiar user interface devices such as keyboards, mice and screens. This interaction has relied on human adaptation for people to enter into the digital world by learning to type, double-click, etc. This paradigm has been sufficient to assist people in an enormous number of tasks, but at the same time has limited the breadth of tasks in which a computer can assist.

In order for general–purpose computers to extend their assistance further, it is necessary for them to adapt to the non-digital world. Laptops and personal digital assistants (PDAs) are the first wave of this adaptation. While miniaturization and mobilization have enable this adaptation and provided new functionality, these devices generally remain blind to the world around them. As a result they have become intrusive in our daily lives as cell-phones ring in movie theaters, laptops interrupt presentations with instant messages, and PDAs rattle and beep for our attention in meetings.

Intel Research Seattle (IRS) [1] and the Assisted Cognition project [2, 3] at the University of Washington in Seattle are jointly developing a system entitled Guide whose goal is to probabilistically infer human activity. It is designed to be scalable to many thousands of activities, to be robust to noise, and to be adaptive to new behaviors. Such a system is valuable simply for its ability to record diaries of daily activities. This would, for example, greatly assist the monitoring of Activities of Daily Living (ADLs) [4, 5]; a task that currently burdens medical caregivers of the elderly. Eventually such a system would be useful as a basis for providing information with more contextual sensitivity, for giving proactive advice to avoid safety and security errors and for training workers in the best known methods of manufacturing operations.

Our approach relies heavily on Radio Frequency Identification (RFID) tags. RFID tags are small computer chips embedded in a sticker the size of a postage stamp. There is no power supply associated with an RFID tag, but when a special antenna broadcasts an information request, the tag collects the energy from the request and re-broadcasts a uniquely identifying number that is embedded within it. The querying antenna listens for the id numbers of any tags which respond and processes the information. Versions of this technology are familiar from their use in retail anti-shopping-lifting systems. RFID tags are cheap, durable, low-maintenance and have very little noise associated with their localization.

We have prototyped an antenna embedded in a glove which is able to see all tags that are located within three inches of the palm of the wearer. This closely approximates sensing when an object is touched and is compelling because it does not require external infrastructure, has a small form factor and effectively guards user privacy with a single point of information control. Guide uses information obtained from this RFID sensor stream, coupled with models of human activity and a powerful inference engine, to probabilistically reason about likely activities in the environment. Once a sensor stream has been recorded, the system is then able to update its models using unsupervised machine learning techniques developed in the context of transportation activity inference [6].

The system relies on three software components to accomplish this, a data miner, an activity inference engine and a visualizer. The data mining component is critical to achieving scalability of the system. Given the wide range of human activity it would be problematic to require an expert to generate a model of every interesting activity that a person could perform. Instead we model ordered compound activities by mining "how-to" sites [7, 8] and interpret the textual order of the steps as sequential constraints. We augment these sequential models with prior probabilities of object touches given an activity by leveraging features of the Google programming API [9]. To date, we have mined the sequential structure of roughly fifteen thousand activities.

The activity inference engine accepts the models from the data miner, converts them to Dynamic Bayesian Networks [10] and uses particle filters [11, 12] to infer activities given an RFID sensor stream. The visualizer gives insight into the reasoning process by showing how the sensors affect the inference engine's belief state and allow a user to improve and debug the generated models. By using Expectation-Maximization the activity inference engine also has the capability to learn model parameters in an unsupervised manner.

In preliminary experiments we asked 14 individuals to wear the RFID glove and go into a home instrumented with RFID tags to perform ADLs. We asked each subject to perform 12 out of 14 activities in the way that they normally would and record the sequence. Example activities included making a snack, making tea, shaving and applying make-up. Under these circumstances and using hand-made models, the system performs well, correctly recognizing 70% of activities performed.

As we continue to integrate and extend the components of our system, we are actively seeking input about how to best support the proactive software components that we regard as the consumers of our reasoning. Understanding the applications and needs of these components will further drive our research agenda and inform our development efforts.

References

1. Intel Research Seattle: Intel Research Seattle Guide Project Website (2003) http://seattleweb.intel-research.net/projects/guide/.
2. Kautz, H., Fox, D., Etzioni, O., Borriello, G., Arnstein, L.: An Overview of the Assisted Cognition Project (2002) http://www.cs.washington.edu/homes/kautz/ac/acaaai02.pdf.
3. Kautz, H., Arnstein, L., Borriello, G., Etzioni, O., Fox, D.: The Assisted Cognition Project. In: Proceedings of UbiCog '02: First International Workshop on Ubiquitous Computing for Cognitive Aids, Gothenberg, Sweden (2002)
4. Lawton, M.P.: Aging and Performance of Home Tasks. Human Factors **32** (1990) 527–536 NLM:0374660;PMID:2074107.
5. Lawton, M.P.: Environment and Other Determinants of Well-Being in Older People. The Gerontologist **23** (1983) 349–357 NLM:0375327;PMID:6352420.
6. Patterson, D.J., Liao, L., Fox, D., Kautz, H.: Inferring High-Level Behavior from Low-Level Sensors. In: Proceedings of UBICOMP 2003: The Fifth International Conference on Ubiquitous Computing. (2003)
7. : (All Recipes) http://www.allrecipes.com/.
8. : (eHow) http://www.ehow.com/.
9. Google Labs: (Google Web APIs) http://www.google.com/apis/.
10. Murphy, K.: Dynamic Bayesian Networks: Representation, Inference and Learning. PhD thesis, UC Berkeley, Computer Science Division (2002)
11. Fox, D., Thrun, S., Dellaert, F., Burgard, W.: Particle filters for mobile robot localization. [12]
12. Doucet, A., de Freitas, N., Gordon, N., eds.: Sequential Monte Carlo in Practice. Springer-Verlag, New York (2001)