Addressing Fluidity through Mixed Technical-Design Practices

Lucian Leahu
Computer Science Department
Cornell University
Ithaca, NY 14853 USA
lleahu@cs.cornell.edu

UbiComp 2009, Sep 30 – Oct 3, 2009, Orlando, FL, USA
ACM XXX-X-XXXXX-XXX-X/XX/XX.

Abstract
Ubicomp systems commonly rely on sensing and recognition capabilities to understand their context. While increasingly successful in simple environments, they face significant challenges addressing the fluid nature of less constrained, human settings. My thesis examines the construction of an interactive, mobile, sensor-based recognition system designed specifically for unconstrained settings. Based on the technical experiments and design explorations required to build such a system, my research makes three key contributions to ubicomp: 1) examines the context independence of statistical inference methods in unconstrained settings, 2) identifies challenges posed by dynamic settings to sensing and recognition technologies, and 3) proposes mixed technical-design approaches to advance ubicomp beyond simple environments.

Problem Statement
In its relatively short history as a field, ubiquitous computing (ubicomp) has come a long way. Context-aware systems – a major target of ubicomp research efforts – have been quite successful in simple environments. Here, approaches such as sensor fusion and activity recognition have proven to be crucial [1]. One major immediate challenge for advancing into everyday human environments is to extend these

capabilities to account for the full complexity and dynamic nature of these environments.

But in situ observations of human activity suggest that while such efforts may be necessary to advance the current state of the field, they may not be sufficient. For example, representations of context are considered problematic because the aspects of context and the relevance of their interplay only become obvious after the action has occurred, *e.g.*, [7]. Bell and Dourish contend that dominant approaches – such as statistical based recognition – assume a world that is orderly, homogeneous and directly accessible to machine sensing and inference. Instead, they argue, the world in which current systems are deployed is messy and heterogeneous. Therefore, if ubicomp is to succeed in everyday environments, it must directly account for the inherent uncertainty and emergent nature of human contexts [2].

Systems probing such conceptual alternatives have so far steered clear of approaches involving recognition and instead address the complexity of human environments through design. For instance, systems designed for user awareness in the home handle this complexity by leveraging the uncertain, designing for the dwellers' interpretation of open-ended representations (*e.g.*, [6]) derived from manipulations of sensor readings. Such a system typically provides opportunities for the users to bridge the gap between its partial understanding of the home and the situated meanings that may emerge. While technically complex in their own right, these alternatives, then, shift the locus of the proposed solution from the technical details of the system to its interaction design.

If one starts from the critiques that context can never be fully captured in a system [3], it might seem that no context-aware system will ever be adequate. My thesis explores what it takes to build a system for everyday human environments that takes the fluidity of such environments as central and yet leverages technical innovation. Specifically, the premise of this work is that we can build systems better suited for everyday contexts by 1) leveraging advances in sensing technologies and statistical classification, and by 2) addressing through design the disjunction between the fluid nature of everyday environments and the limited understanding available to technology.

Research questions and approach

But is this at all possible? If so, how can we benefit from better sensors and improved machine learning (ML) techniques to provide more information to our systems without running counter to the conviction that human contexts are characterized by contingency? What are the consequences and limitations of such an approach?

My thesis opens up the research space around these questions and offers answers through the construction of an ubicomp system. `Freaky` – an interactive system aiming to support user awareness and reflection on fear in the wild – provides the basis to explore how to adapt existing technical and design strategies to address the emergent nature of emotion, activity and context in the construction of a system incorporating a sensor-driven emotion classifier. The challenge in building Freaky, then, is to design for the emergent, contingent, situated nature of human emotion and use whatever insights can be obtained from mining sensor data through statistical inference.

The originality of my approach and the main contribution of this research lie in demonstrating the careful coordination of technical and design strategies required to construct systems for everyday human environments.

In the following, I present the studies making up my dissertation. I begin by looking at the field of AI's attempts to build systems for everyday environments in order to identify alternative strategies for ubicomp systems. I continue by experimenting with statistical classifiers of emotion from physiological sensors and studying the relationship between physical context and classifiers' performance. I describe the consequences of these findings on the interplay between `Freaky`'s design and technical specifications. I conclude with the expected contributions of my thesis.

Lessons from interactionist AI

The field of AI addressed similar challenges in scaling beyond simple environments to those currently faced by ubicomp. Examining a key episode in the history of AI – the emergence of interactionist AI to circumvent difficulties faced by symbolic AI – yielded strategies for ubicomp systems aimed at everyday environments [5].

The success behind such *technical* approaches was owed to their coupling with innovative system *design* strategies. These strategies explicitly targeted the system-environment-humans interplay, *e.g.,* taking advantage of structural regularities in the physical environment, rather than explicitly handling them in the code, and combining the behavioral regularities in the users mind with the system's capabilities to make up for the technology's lack of full understanding of the world. *In this way, the complexity of human environments was tackled through a carefully choreographed collaboration between technical and design elements.*

One of the technical–design strategies identified is particularly salient for `Freaky` and, more generally, for ubicomp systems using recognition. Complex, formal techniques (*e.g.,* plans, inference engines) can be included in the code without reducing user experience to general, prior specified categories. Instead, shortcomings rooted in the technical approach – precisely the reliance on such strict categories – are dealt with in the design of the system. This is achieved by positioning the system's understanding of the world as limited, yet useful, *i.e.,* to be engaged with by the users, rather than relying exclusively on the correctness of the inference modules.

The implication for my system is that fear recognition need not be perfect, as long as the gap between the machine interpretation of user emotion and user emotion is bridged in the interaction. To this end, the design of the system *must* engage the gap between user emotion and the machine interpretation thereof. The complexity of human emotion is thus addressed in the user-system coupling. In this way, the demands for perfect or close to perfect accuracy – which would make or break a traditional system addressing such complexity only through classification – are significantly reduced. As I show in the following section, physiological models of emotion are far from perfect predictors for real world data. Therefore, exploring mixed design-technical approaches is imperative.

Studying the relationship between physical context and recognition in a complex setting

The goal of this study was to understand what happens when we move into contexts that are more complex from the system's perspective. Typically, emotion recognition from biometric readings has been performed with data collected in highly constrained environments, *e.g.,* laboratories. In such settings, ML classifiers can achieve high accuracy. The assumption underlying such studies is that the recognition is context-independent and once good biometric signals, sensors and features are found, one may simply repeat the procedure in other contexts. However, a study I conducted in an everyday setting showed that people attribute their own physiological variations to a variety of factors [4]. One such factor is physical activity (PA). The question we had to investigate before building `Freaky` was whether models can detect changes in physiology caused by emotions, and not other factors. To investigate the physical context-independence of statistical models of emotion, we collected data in three different settings: #1 experiencing fear in a controlled environment, without PA; #2 experiencing fear outdoors while the person is involved in PA; and #3 different degrees of PA, but no emotion.

Data was collected from 4 participants in each setting, using three sensors: EKG, skin conductance, respiration volume. For training, support vector machines (SVMs) were used. For brevity, I skip the ML details.

A large number of models were trained and tested using different combinations of the data. The choice of parameters and features was informed by previous studies and by optimizing measures such as accuracy, recall and precision. Here, I present only a subset of the results and the implications of the study: the need for mixed technical-design approaches.

The context of interest for `Freaky` is #2: an outdoors setting in which the user is involved in physical activity. Models trained and tested with data from condition #2 show promising results: 90.1% average accuracy (70% baseline, *i.e.,* guessing 'no fear' all the time). Similarly robust results have been obtained for the controlled setting (#1).

However, when testing models on data from different contexts, the picture painted by the classification results is one of serious brittleness. For example, models trained in condition #2 and tested on the PA data (condition #3) misclassify a large proportion of the points as fear: 62% accuracy (70% baseline). Similarly, models trained in condition #1 and tested on data from #3 also have a hard time discriminating between fear and PA: 51% accuracy. Other scenarios show similarly poor results (train on #1, test on #2; train on #2, test on #1).

As a way of providing more guidance for the models, I included #3 data into training sets for the other cases. Surprisingly, the performance degrades even more, *e.g.,* training with #3 and #2 data, testing on #1: the model didn't find *any* fear points, suggesting that the fear patterns are context specific. Examining the gain ratio for the 69 features used, we found that the gain ration for the same attribute varies significantly between #1 and #2 data.

These results show that models may not generalize well to new contexts, demonstrated by the sub baseline performance. However, **highly context specific** models – harnessing `local' statistical regularities in the data – may work for complex settings.

A related aspect that had to be clarified is related to the ambiguity and uncertainty that are central aspects of emotion. Statistical classifiers, in contrast, work by reducing noise in the data, *i.e.,* minimizing error. How are these conflicting orientations playing against each other in the statistical models?

One example is labeling data points as fear and non-fear, which proved to be straightforward in the controlled setting (#1), but somewhat arbitrary in the uncontrolled setting (#2). In the controlled setting it was quite obvious when the person was scared; in contrast, in the outdoors setting, the person's experience was much more fluid. Should 'butterflies in the stomach' be counted as fear? What about 'still feeling a little buzz'? These examples demonstrate that part when moving into less constrained environments, categories that appear to be clear cut in simple settings become much more complex. Thus, part of the uncertainty of real-world models is due to the nature of the phenomena (emotion) and thus ML methods of removing noise can not help with this issue. However, this observation extends beyond emotion recognition. Indeed, most aspects pertaining to everyday human contexts exhibit a similar ambiguity – notably for ubicomp efforts around activity recognition.

These observations coupled with the distinct possibility of context specific models to overfitting the data suggest that mixed approaches must be considered: on the one hand, advances in sensors and ML methods are needed to build better models and, on the other, innovative design closely linked to the technical implementation to account for the ambiguity that's inherent to the data and therefore the models.

Technical-design (re)configurations

This section presents work in progress on the elaborate technical and design coordination required to build `Freaky`. The system is conceptualized as an artificial companion. `Freaky` connects to the user via biometric sensors and is to be placed in a baby carrier.

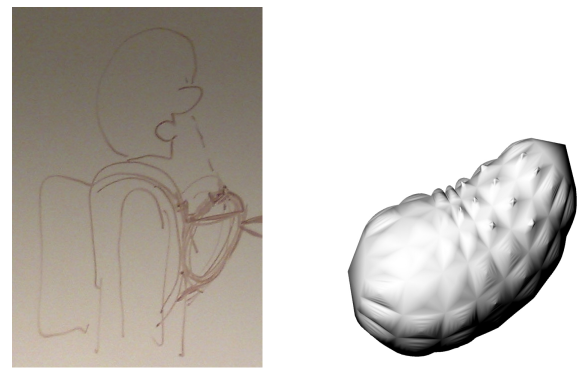

Figure 1. `Freaky` with user (left) and cast surface (right).

The interaction design as well as the artefact's physical design was directly influenced by the system's technical specifications, as well as limitations. Concretely, the limited prediction capabilities in novel contexts called for highly context-specific fear models: training data collected in the deployment setting. Even so, the expected accuracy will be far from perfect as overfitting is likely. To acknowledge this discrepancy between the person's fear and `Freaky`'s understanding of her fear, its deliberately unusual shape and surface gives the user a sense that `Freaky` might have a different, yet hopefully useful, understanding of fear.

This difference between machine emotion and human emotion also guides the interaction design. `Freaky` is endowed through interaction design with its own behaviours, as well as fear: when `Freaky` detects fear it becomes scared, starts shaking and having flash backs (plays sounds recorded from previous scary episodes); in order to return to ‘normal’, `Freaky` requires attention from the user: pressure, light and acceleration sensors give the system an understanding of whether the user is petting it, protecting it, rocking it like a baby. These specific user actions were chosen because they are likely to also calm down the user.

Endowing `Freaky` with its own fear and associated behaviours allows the user to create a narrative about the machine’s fear and understandings of fear as its own, although linked to the user’s physiology. Through the overlaps, as well as differences between machine and user fear, the user has the opportunity to reflect on her own, by continuously being prompted to contrast her assessment of fear with `Freaky`’s.

Given the focus on user reflection and discovery, it is important that `Freaky` fails to detect user fear as seldom as possible. This design decision triggered a change in the way the fear model is to be optimized: minimizing false negatives; in turn, this meant that the number of false positives is likely to increase. To avoid frustrating the user by frequently requiring her attention, I am considering changes in `Freaky`’s personality design: `Freaky` feeds off the user’s fear. If a long time has passed without detecting fear, `Freaky` becomes restless and calms down when it detects fear; it would only require attention when it is hungry *i.e.,* it hasn’t detected fear in a long time. Currently, I am testing these different interaction scenarios and their consequences on users’ experience.

These design scenarios illustrate the subtle and continuous negotiation of design and technical specifications required to account for the fluidity of emotion and context the complexity of interaction in real world settings.

Contribution

Unlike simplified environments such as laboratories or industrial production lines, everyday human environments pose context dependence issues for recognition technologies and exhibit an ambiguity that makes clear cut categorization problematic. My work demonstrates a mixed technical-design practice that couples interaction design ingenuity with technical and computational innovation for systems that fit better in everyday human environments.

References

[1] Abowd, G. D. and Mynatt, E. D. 2000. Charting past, present, and future research in ubiquitous computing. *ACM TOCHI 7, 1* (Mar. 2000), 29-58.

[2] Bell, G. and Dourish, P. 2007. Yesterday’s tomorrows. *PUC 11, 2* (Jan. 2007), 133-143.

[3] Dourish, P. 2004. What we talk about when we talk about context. *PUC 8, 1* (Feb. 2004), 19-30.

[4] Leahu, L., Schwenk, S., and Sengers, P. 2008. Subjective objectivity. In *Proc. DIS '08*, 425-434.

[5] Leahu, L., Sengers, P., and Mateas, M. 2008. Interactionist AI and the promise of ubicomp. In *Proc. UbiComp’08*, 134-143.

[6] Pousman, Z., Romero, M., Smith, A., and Mateas, M. 2008. Living with tableau machine. In *Proc. UbiComp’08*, 370-379.

[7] Salvador, T. and Anderson, K. Practical Considerations of Context for Context Based Systems. In Proc. UbiComp '03, 243-255.