Quantifying Magic In Ubicomp Systems Scenarios

John Barton and Jeff Pierce

IBM Almaden Research Center

Abstract. A ubiquitous computing (ubicomp) system that successfully realizes an impractical scenario is technically interesting but functionally useless. Therefore, an evaluation of a ubicomp system should consider both the system itself and the broader scenario that motivates and shapes it. Our vocabulary for describing and evaluating the suitability of scenarios is, however, very limited. We believe that a useful first step toward improving that vocabulary is to assemble a set of guidelines for considering the type and amount of "magic" in scenario

Introduction Research papers on ubiquitous computing (ubicomp) systems and applications often use "scenarios": vignettes describing users and uses of new technology. These scenarios illustrate and motivate the improvements offered by a proposed technology. Moreover, scenarios provide both a list of requirements and an effectiveness test. Many ubicomp papers, then, provide some evidence that their implemented system fulfills the scenario requirements by a convincing demonstration or through an evaluation with the putative users. However, evaluating a research system against its own scenario makes the assumption that the proposed scenario is an appropriate and effective choice for driving technological progress.

Unfortunately, we do not have any practical basis for evaluating the scenarios themselves. We instead create subjective benchmarks for our systems and then work to demonstrate their success against those (potentially unrealistic) goals. While we do not wish to abandon the fun, inspiring, and futuristic scenarios that are often central to the ubicomp approach, we believe that ubicomp systems research would improve if we had more objective metrics for success. We hope the Ubisys workshop discussions can lead to new objective criteria for evaluating ubicomp scenarios.

As a first step, we propose that the Ubisys Workshop create a working set of objective and practical guidelines for the evaluation of Ubicomp scenarios as a critical tool for the development and evaluation of Ubicomp systems and applications research. We believe that these guidelines need to "quantify magic": they should provide a recipe for describing the practicality of a proposed future technology and its use. The recipe should include the economic, social, legal, and ethical boundaries on the technology, as well as the hardware, software, and services required by it. By discussing various practical limits and devising and sharing ways of estimating their impact, the Ubisys workshop can improve both

the reporting of ubicomp research and its practice. Concretely we propose to use discussions in the workshop to create a paper outlining a proposed discipline for evaluating ubicomp scenarios and solutions.

Types of Magic In Ubiquitous Scenarios Most scenarios, because they are trying to motivate a future use of technology, make assumptions about how people, society, or technology might change. Some of these changes are discussed explicitly, often as part of the scenario introduction. Others are left unsaid: to readers of the scenario these changes appear as if created by "magic". Any changes from the status quo that the authors are not directly providing with their new technology assumes some type of magic.

Elucidating the types of magic in ubicomp scenarios allows us to concretely consider the practicality, cost, and timeframe of the proposed usage. While these types are open to debate, we propose an initial set to spur discussion:

- Computational: a scenario might require computing a known NP-hard algorithm (e.g. optimal scheduling of a shared service or resource).
- Algorithmic: a scenario might require the application of algorithms that do not yet exist or that are impossible, e.g. recognition of complex activities from low-level sensor data.
- Financial: a scenario might require the expenditure of money by individuals, corporations, or governments that is unlikely;
- Technological: scenarios often require changes in technology (e.g. scale, cost, power) that are not the focus of the described research.
- Biological: a scenario might require the use of perceptual and cognitive resources beyond those currently possessed by people.
- Sociological: a scenario might assume that individuals, groups, or institutions will change current behaviors (e.g. toward trust, altruism, selflessness, etc.).
- Logistical: scenarios almost never describe how a proposed technology will be maintained, and yet those costs currently consume a significant proportion of lifetime system costs.
- Legal: scenarios may presume uses of technology or information that are legally difficult or impossible.
- Infallability: scenarios tend to ignore the possibility of technological failures, either as a result of design mistakes or as a result of actions by (potentially malicious) individuals.

Some Magic has Quantifiable Limits While enumerating and discussing the dimensions of magic in a scenario is a step toward evaluating its practicality and usefulness, the ability to quantify the amount of magic along each dimension would add significantly more value and would allow researchers to more effectively compare and contrast different scenarios.

There is some evidence that at least some types of magic are quantifiable. For example, the field of computer science has developed a number of techniques for determining a algorithm's computability and computational complexity. Over time, the field has learned how to categorize whether a particular problem is P or

NP, as well as to calculate upper bounds and provide estimates of running times. Indeed, publication of an algorithm typically requires an analysis of complexity to demonstrate value.

As another example, consider new computerized hardware. Given a rough idea of the design, a bill of materials, and an estimate of the cost of assembly, researchers can generate a ballpark estimate of the device cost. Then, using trends in market prices and, in some cases, longer term trends in technology advances, they can estimate future costs of the devices (with increasing uncertainty further in the future). As long as the device is somewhat similar to existing devices, they can estimate the time when early adopters would buy the device if they found it useful or interesting. Today such analyses are essential before any business commits significant resources to new devices.

Toward Quantifying Other Types of Magic Unfortunately, while we have mechanisms for describing the amount of some types of magic in a scenario, we still need to explore methods of quantifying the other types of magic and for combining the analyses of different types. Fortunately we have two significant advantages over our examples from algorithms and hardware engineering: we do not need to be nearly as thorough or as accurate in our analyses and we may be able to develop and share on-line tools to reuse them. Since today's systems are presented with almost no analysis, any improvements will be advances. As we get better at analysis we can formulate ways of reusing it – as was done in computational complexity – or develop tools for accelerating it – as has been done with computer aided design tools for hardware engineering.

To give a sense of how a lightweight evaluation might appear in a publication, consider describing a system for future cell phones to connect to future computerized televisions. A paper introducting that system might include a section "Context of Our Contribution" starting with:

> Using Wonderland estimation[23] with storage and wireless predictions from Kubrick et al[24], we extrapolate 100k early adopters by late 2008 carrying 20GB in cellphones with Internet radio connection >10MBs. This coincides with the expected roll out of Internet connected televisions using the CES2014 interconnection standard[25]. Venn's analysis of early adopters predicts a 10% overlap of these population. Our service could be deployed on the Internet for low incremental cost and rapid upgrade. Significant issues of privacy and security[26] are not addressed here, limiting the scope of our contribution to controlled environments like homes and businesses.

Once a few papers or, even better, online tools appeared, citations of this work by ubicomp practitioners would encourage addition analytical concepts and tools.

We wish to emphasize that we do not believe that the presence of magic in ubicomp scenarios is good or bad in and of itself. Instead of removing magic, we propose that researchers should categorize and quantify the types and amounts of magic that they incorporate into scenarios and reflect upon the likely impacts.

As a primary consumer of Ubicomp scenarios, we believe that the Ubisys community is well-positioned to assemble an initial set of guidelines for describing and evaluating the use of magic in scenarios.