CHAPTER 1

INTRODUCTION: THE MIDDLE OF MOORE'S LAW

The history of technology is a history of unintended consequences, of revolutions that never happened, and of unforeseen disruptions. Take railroads, for instance. In addition to quickly moving things and people around, railroads brought a profound philosophical crisis of timekeeping. Before railroads, clock time followed the sun. "Noon" was when the sun was directly above, and local clock time was approximate. This was accurate enough for travel on horseback or foot, but setting clocks by the sun proved insufficient to synchronize railroad schedules. One town's noon would be a neighboring town's 12:02, and a distant town's 12:36. Trains traveled fast enough that these small changes added up. Arrival times now had to be determined not just by the time to travel between two places, but the local time at the point of departure, which could be based on an inaccurate church clock set with a sundial. The effect was that trains would run at unpredictable times and, with terrifying regularity, crash into each other.

It was not surprising that railroads wanted to have a consistent way to measure time, but what did "consistent" mean? Their attempt to answer this question led to a crisis of timekeeping: Do the railroads dictate when noon is, does the government, or does nature? What does it mean to have the same time in different places? Do people in cities need a different timekeeping method than farmers? The engineers making small steam engines in the early nineteenth century could not possibly have predicted that by the end of the century their invention would lead to a revolution in commerce, politics, geography, philosophy and just about all human endeavors.[1]

We can compare the last twenty years of computer and networking technology to the earliest days of steam power. Once, giant steam engines ran textile mills and pumped water between canal locks. Miniaturized and made more efficient, steam engines became more dispersed throughout industrial countries powering trains, machines in workplaces, and even personal carriages. As computers shrink, they too are getting integrated into more places and contexts than ever before.

We are at the beginning of the era of computation and data communication embedded in, and distributed through, our entire environment. Going far beyond how we now define "computers," the vision of ubiquitous computing

[1] See Chapter 2 of O'Malley (1990) for a detailed history of the effect of railroads on timekeeping in America.

Figure 1-1
The adidas_1 shoe, with embedded microcontroller and control buttons. (Courtesy Adidas)

(see Sidebar: The Many Names of Ubicomp) is of information processing and networking as key components in the design of everyday objects (Figure 1-1) using built-in computation and communication to make familiar tools and environments do their jobs better. It is the underlying (if unstated) principle guiding the development of toys that talk back, clothes that react to the environment, rooms that change shape depending on what their occupants are doing, electromechanical prosthetics that automatically manage chronic diseases and enhance people's capabilities beyond what is biologically possible, hand tools that dynamically adapt to their user, and (of course) many new ways for people to be bad to each other.[2]

The rest of this chapter discusses why the idea of ubiquitous computing is important now, and why user experience design is key to creating successful ubiquitous computing (ubicomp) devices and environments.

Sidebar: The Many Names of Ubicomp

There are many different terms applied to what I am calling ubiquitous computing (or ubicomp for short). Each term came from a different social and historical context. Although not designed to be complementary, each built on the definitions of those that came before (if only to help the group coining the term identify themselves). I consider them to be different aspects of the same phenomenon:

- *Ubiquitous computing* refers to the practice of embedding information processing and network communication into everyday, human environments to continuously provide services, information, and communication.
- *Physical computing* describes how people interact with computing through physical objects, rather than in an online environment or on monolithic, general purpose computers.
- *Pervasive computing* refers to the prevalence of this new mode of digital technology.
- *Ambient intelligence* describes how these devices appear to integrate algorithmic reasoning (intelligence) into human-built spaces so that it becomes part of the atmosphere (ambiance) of the environment.

[2]This book will not discuss military ubiquitous computing, although it is certainly a major focus of development. The implication of computers embedded into weapons and surveillance devices has been discussed for as long as ubicomp (DeLanda, 1991), if not longer.

- *The Internet of Things* suggests a world in which digitally identifiable physical objects relate to each other in a way that is analogous to how purely digital information is organized on the Internet (specifically, the Web).

Of course, applying such retroactive continuity (a term the comic book industry uses to describe the pretense of order grafted onto a disorderly existing narrative) attempts to add structure to something that never had one. In the end, I believe that all of these terms actually reference the same general idea. I prefer to use ubiquitous computing since it is the oldest.

1.1 THE HIDDEN MIDDLE OF MOORE'S LAW

To understand why ubiquitous computing is particularly relevant today, it is valuable to look closely at an unexpected corollary of Moore's Law. As new information processing technology gets more powerful, older technology gets cheaper without becoming any less powerful.

First articulated by Intel Corporation founder Gordon Moore, today Moore's Law is usually paraphrased as a prediction that processor transistor densities will double every two years. This graph (Figure 1-2) is traditionally used to demonstrate how powerful the newest computers have become. As a visualization of the density of transistors that can be put on a single integrated circuit, it represents the way semiconductor manufacturers distill a complex industry into a single trend. The graph also illustrates a growing industry's internal narrative of progress without revealing how that progress is going to happen.

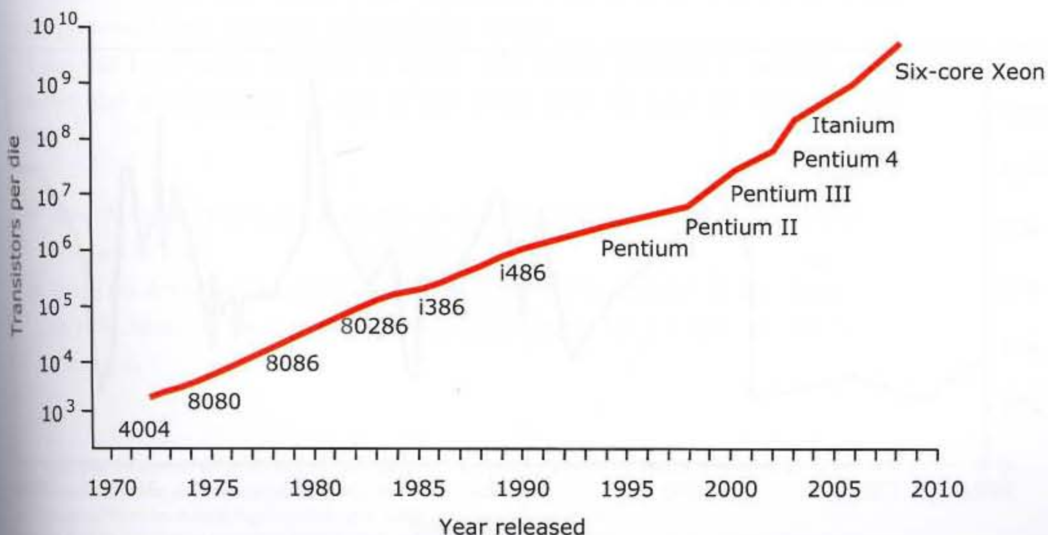

Figure 1-2
Moore's Law. (Based on Moore, 2003)

Moore's insight was dubbed a law, like a law of nature, but it does not actually describe the physical properties of semiconductors. Instead, it describes the number of transistors Gordon Moore believed would have to be put on a CPU for a semiconductor manufacturer to maintain a healthy profit margin given the industry trends he had observed in the previous five years. In other words, Moore's 1965 analysis, which is what his law is based on, was not a utopian vision of the limits of technology. Instead, the paper (Moore, 1965) described a pragmatic model of factors affecting profitability in semiconductor manufacturing. Moore's conclusion that "by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip" is a prediction about how to compete in the semiconductor market. It is more of a business plan and a challenge to his colleagues than a scientific result.

Fortunately for Moore, his model fit the behavior of the semiconductor industry so well that it was adopted as an actual development strategy by most of the other companies in the industry. Intel, which he co-founded soon after writing that article, followed his projection almost as if it was a genuine law of nature and prospered.

The economics of this industry-wide strategic decision holds the key to the emergence of ubiquitous computing. During the Information Revolution of the 1980s, 1990s, and 2000s, most attention was given to the upper right corner of Moore's graph, the one that represents the greatest computer power. However, there was a secondary effect: as processors became more powerful, the cost of older technology fell.

As the processing power increased exponentially, the price of new CPUs remained (fairly) stable (Figure 1-3); and cost of older technology dropped at (roughly) the same rate as the power of new processors rose (Figure 1-4). Since

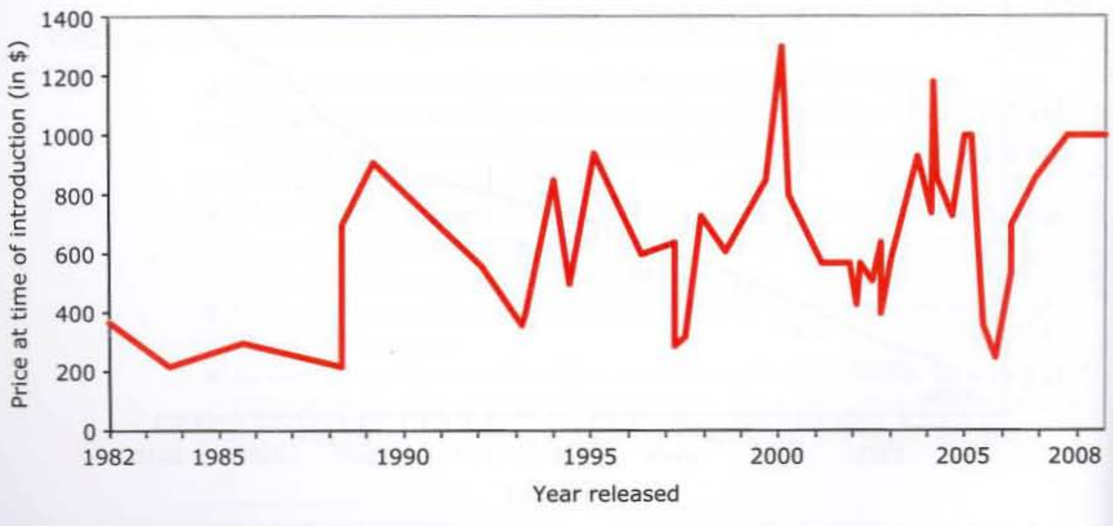

Figure 1-3

CPU Prices 1982–2009. (Data source: Ken Polsson, processortimeline.info)

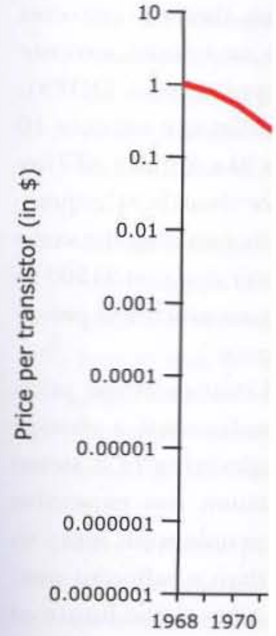

Figure 1-4

Per transistor cost of CPUs,

new technology gets n
just as quickly. Althou
to process informatio
really powerful[3] but n

Take the Intel i486
between the pre-Inter
1990s:

- It ran Microsoft Wi
  released in 1990.
- It was the dominant
  in 1993. Most early
  computer.

[3]This assertion is somewhat of
a manufacturing and pricing s
not fire up a Pentium II-makin
though, is that once Intel con
still make the functional equi
proportionally less expensive t
chips would likely be physicall

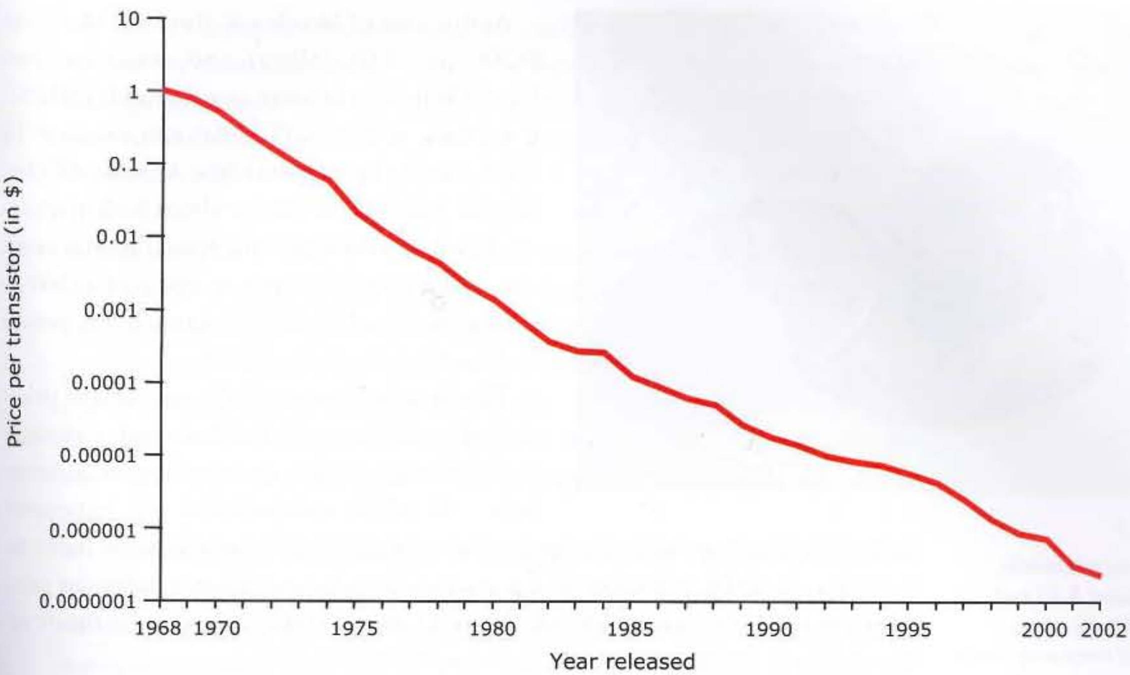

Figure 1-4

Per transistor cost of CPUs, 1968–2002. (Based on: Moore, 2003)

new technology gets more powerful very quickly, old technology drops in price just as quickly. Although old technology gets cheaper, it loses none of its ability to process information. Thus, older information processing technology is still really powerful[3] but now it is (almost) dirt cheap.

Take the Intel i486, released in 1989. The i486 represents a turning point between the pre-Internet PC age of the 1980s and the Internet boom of the 1990s:

- It ran Microsoft Windows 3.0, the first commercially successful version of Windows, released in 1990.
- It was the dominant processor when the Mosaic browser catalyzed the Web boom in 1993. Most early Web users probably saw the Web for the first time on a 486 computer.

[3]This assertion is somewhat of an oversimplification. Semiconductor manufacturing is complex from both a manufacturing and pricing standpoint. For example, once Intel moved on to Pentium IIIs, they could not fire up a Pentium II-making machine at a whim to make cheap Pentium IIs. What is broadly true, though, is that once Intel converted their chip-making factories to Pentium III technology, they could still make the functional equivalent of Pentium IIs, and (for a variety of reasons) making those chips was proportionally less expensive than making Pentium IIIs. In addition, these new Pentium II-equivalent chips would likely be physically smaller and use less power than their predecessors.

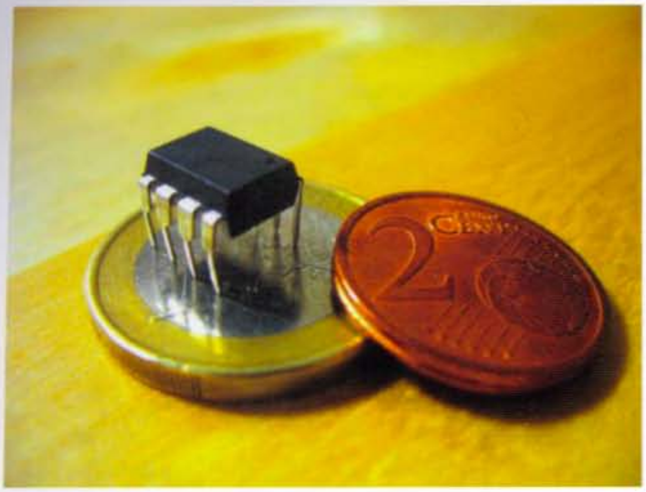

Figure 1-5

ATTiny microcontroller sells for about $.50 and has roughly the same amount of computing power as an Intel i486, which initially sold for the 2010 equivalent of $1500. (Photo © Uwe Hermann, licensed under Creative Commons Attribution — Share Alike 2.0, found on Flickr)

At the time of its release, the Intel i486 cost $1500 (in 2010 dollars) and could execute 16 million instructions per second (MIPS). If we look at 2010 CPUs that can execute 16 MIPS, we find processors like Atmel's ATTiny (Figure 1-5), which sells for about $.50 in quantity. In other words, broadly speaking, the same amount of processing power that cost $1500 in 1989 now costs $.50 and uses much less power and requires much less space.

This is a fundamental change in the price of computation — as fundamental a change as the change in the engineering of a steam boiler. In 1989, computation was expensive and was treated as such: computers were precious and people were lucky to own one. In 2010, it has become a commodity, cheaper than a ballpoint pen. Thus, in the forgotten middle of Moore's Law charts lies a key to the future of the design of all the world's devices: ubiquitous computing.

1.2 UBIQUITOUS COMPUTING

Like many other prescient observations and innovations (Hiltzik, 2000), the researchers at Xerox PARC identified in the 1980s that technology was part of accomplishing social action (Suchman, 1987) and that personal computers were "too complex and hard to use, too demanding of attention, too isolating from other people and activities, and too dominating" (Weiser et al., 1999). They coined the term "ubiquitous computing" to describe their program to develop a range of specialized networked information processing devices to address these issues.[4]

Xerox PARC's then Chief Technology Officer, Mark Weiser, described these ideas in his 1991 *Scientific American* article, "The Computer for the 21st Century." In this article he contrasts the potential of ubicomp technology to portable computers and virtual reality, which was then the state-of-the art in popular computer thought:

[4]It is interesting to hypothesize how apparent the implications of Moore's trend were to Moore's mid-60s contemporaries, especially in terms of how cheaper, denser electronics would affect the size, shape, and use of computers—and who first thought of having multiple computers distributed in the environment. Accompanying Moore's original 1965 *Electronics* magazine article is a cartoon by Grant Compton that shows a salesman hawking a handheld computer alongside stands for "notions" and "cosmetics," with well-dressed men and women crowding around him. The cartoon jokes that if Moore's plan is followed, eventually computers will be as small, as common, and sold in the same way as universally consumed personal items. It exaggerates the implications of Moore's article for humor, but perhaps it was funny because it acknowledged a hope that was shared by Compton's readers for a future where computers were treated like commodities.

> The idea of integrating computers seamlessly into the world at large runs counter to a number of present-day trends. "Ubiquitous computing" in this context does not just mean computers that can be carried to the beach, jungle or airport. Even the most powerful notebook computer, with access to a worldwide information network, still focuses attention on a single box.
>
> [...]
>
> Perhaps most diametrically opposed to our vision is the notion of "virtual reality," which attempts to make a world inside the computer. [...] Although it may have its purpose in allowing people to explore realms otherwise inaccessible [...] virtual reality is only a map, not a territory. It excludes desks, offices, other people not wearing goggles and body suits, weather, grass, trees, walks, chance encounters and in general the infinite richness of the universe. Virtual reality focuses an enormous apparatus on simulating the world rather than on invisibly enhancing the world that already exists.
>
> [...]
>
> Most of the computers that participate in embodied virtuality[5] will be invisible in fact as well as in metaphor. Already computers in light switches, thermostats, stereos and ovens help to activate the world. These machines and more will be interconnected in a ubiquitous network.

Whether or not Weiser used the semiconductor industry's price trends in his calculations, his title accurately anticipated the market. The year 1991, when Weiser wrote his article, was still the pre-Web era of the i486. The vision he described of many small powerful computers, in different sizes, working simultaneously for one person (or a small group) was simply unaffordable. The economics of processors to make it commercially viable would not exist until well into the first decade of the twenty-first century (and, sadly, some years after Weiser's premature death in 1999).

I estimate that the era he envisioned began in 2005. Technologies typically emerge piecemeal at different times, so 2005 is an arbitrary date.[6] But in 2005, Apple put out the first iPod Shuffle, Adidas launched the adidas_1 shoe (Figure 1-1), and iRobot launched the Roomba Discovery robotic vacuum cleaner. None of those products looked like a traditional computer. Moreover, the Shuffle and Discovery were second-generation products, which implies that the first generation's success justified additional investment, and the adidas_1 was deeply embedded in a traditionally non-technological activity (running).

Also, by 2005, a range of industry factors made possible the efficient development of products that roughly fit Weiser's vision of ubiquitous computing. No

[5] Weiser plays on the "virtual reality" term to mean "the process of drawing computers out of their electronic shells."

[6] Ambient Devices' Ambient Orb, for example, came out in 2002.

longer did the elements — software, hardware, and networks — have to be integrated from scratch, often painfully, as they had been throughout the 1990s. Starting around 2000, several factors pointed to an emergence of ubicomp as a commercial agenda:

- CPU technology prices had fallen to the point that information processing had gotten powerful and inexpensive.
- The Internet had become familiar with clear social and commercial benefits outside of the scientific and engineering community.
- A number of standard communication and data exchange protocols had been developed and refined through widespread deployment.
- Digital telephony was firmly established, and many people were carrying lightweight, network-connected computers in the form of mobile phones.
- Wireless communication had become common, standardized, and successful with millions of access points deployed throughout the world.
- Designers spent the first dotcom boom developing a wide range of interactive products and were experienced with interaction design for networked services.

Thus, in 2000 the information processing technology, the networks, and, most important, technological familiarity among designers, developers, and businesspeople were all available. By 2005, the fruit of their efforts could be seen in stores and, after nearly two decades of anticipation, the era of ubiquitous computing had begun.

1.3 THE NEED FOR DESIGN

The ubicomp vision may have existed twenty years ago, but throughout the 1990s the complexity of the technology overshadowed nearly all consideration of user experience. The design of embedded systems (as small specific-purpose computers were typically called) was the concern of electrical engineers in R&D departments and universities rather than interaction designers in startups and product groups. Just getting the pieces to interoperate was a kind of victory, never mind whether the resulting product was usable or enjoyable.

The lack of precedent for devices that combined computers with everyday objects meant that the user experience design for each new object had to start from scratch. Nearly every product represented a new class of devices, rather than an incremental evolution to an existing known device. The final nail in the coffin of 1990s ubicomp was (unexpectedly) the Web: by the middle of the decade it was a known quantity with known benefits and (presumed) revenue models. There were few incentives for designers, companies, and entrepreneurs to risk jumping into another new set of technologies that needed to be first understood, then explained to a consumer market.

Thus, the potential within the technology was relatively unrealized in the mainstream. However, something else was happening at the edges, outside of the main consumer electronics and personal computer worlds. Toy designers, appliance manufacturers, car designers, and industrial designers realized that the products they were creating could incorporate information processing technology more deeply. These groups already used computer technology, but did not necessarily consider themselves in the same business as computer manufacturers.

Today, the market is changing again and the incentives are shifting. The success of Web services on mobile phones demonstrates that networked products stretch beyond a laptop browser. Intelligent, connected toys show that objects with little processing power can exhibit interesting behaviors with just a little networking. The prices for powerful CPUs have fallen below a threshold where incorporating them becomes a competitively viable business decision. The concept of designing a single general purpose "computation" device is fading progressively into the same historical background as having a single steam engine to power a whole factory. As it fades, the design challenges grow clearer.

Right now is the time to create a practice of ubiquitous computing user experience design. The technology is ready. Consumers are ready. Manufacturers are ready. The world is ready. Now it is up to designers to define what that practice will mean.

And what of the railroads and time? Time zones, a ubiquitous technology we have come to take for granted, were invented in the 1860s, standardized by the railroads in the 1880s, and hotly debated until the 1918 Standard Time Act made them US law (O'Malley, 1990). Once trains ran on schedule, they could save countless lives, create enormous fortunes, displace native peoples, pollute the air, and transform the world. Ubiquitous computing is poised to be the next such transformational technology.