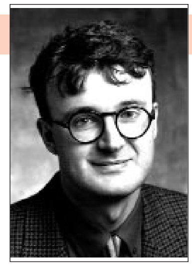


By Giovanni Flammia
gflammia@alum.mit.edu

The wireless Internet today and tomorrow

In this installment, I review the current technology for accessing the Web using wireless handheld devices. There is a lot of excitement in the news about the wireless Internet. If you just read newspapers and magazines, it would look as if we were all able to read and send news and e-mail from little handheld devices in our pockets. In reality, the current technology is limited, but fortunately, everything is about to change in the wireless world.

The situation is similar to the early days of the Web, when Web-based graphical user interfaces were about to replace simple text-based tools such as telnet and Gopher. The following scenario give an idea of what browsing the Web on a wireless device feels like today. Imagine logging into your university computer with your old DOS computer, using a black and white text window of at most five lines by 30 characters with a 9.6 Kbits/sec analog modem. And imagine that, instead of using HTML, Web pages from different sites were written using different markup languages. Times have changed dramatically since those early days, and we'll soon see the same sea change in current wireless technology.

Current technology

Right now, you can't use your wireless telephone everywhere you go. Most networks don't cover the entire United States, and they have blind spots outside metropolitan areas. In addition, US standards differ from the ones used in Europe and in Japan. Wireless telephones in the US are connected using two standards: Time Division Multiple Access and Code Division Multiple Access, which enable connections between 9.6 and 14.4 Kbits/s. In the rest of the world, however, the Global System for Mobile communications (GSM) standard enables connections up to 19.2 Kbits/s.[1]

If you design a Web page today to be browsed by a wireless telephone, you must write your pages in a language other than HTML. In Europe and the US, there is a specific protocol designed to connect devices to the wireless network called the Wireless Access Protocol, which specifies a markup language for Web pages called the Wireless Markup Language (WML).[2] Meanwhile, in Japan, NTT DoCoMo offers a very popular service called I-mode, which is based on a subset of HTML called CHTML (the C stands for compact).

Currently, you can use the same wireless telephone handset for either voice or text communication, but you can't do both forms of communication simultaneously. In today's telephones, speech and text are like oil and water—they don't seem to mix too well. If you browse the Internet using WML, you cannot speak commands and it cannot read to you aloud a Web site's content—for example, if you talk to someone about a recipe, you can't go and fetch the recipe text from a Web site. If you use voice input and output to browse the Internet with a speech recognition portal such as Tellme.com, you quickly find out how limited the service is (it understands short phrases and keywords).

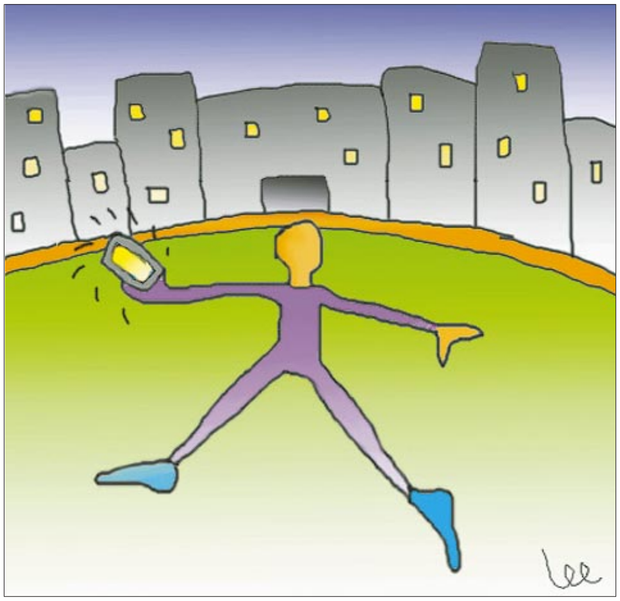
In spite of all these limitations, online and offline Internet browsing with handheld devices is wildly popular. In Japan and Europe, millions of teenagers use handheld devices and Short Messaging System services to send each other messages. In the US, millions of people "hotsynch" and "beam" each other's Palm devices with Web content from portals such as AvantGo and Vindigo. AvantGo offers, among many other things, headline news from the *Wall Street Journal* and cocktail recipes. Vindigo offers Zagat listings for many cities that you can search by street address directly on your PDA. One of my friends never forgets to hotsynch her Palm to the Internet every day before breakfast. She "reads" it on her way to work, tapping away with the stylus and using the scroll button.

Tomorrow's technology

The network, the devices connected to it, and the user interface are set to change in the near future. In spite of all the changes, the same applications will be popular—messaging, news, entertainment, and location-based services—and new applications will emerge.

Wireless networks will be everywhere, and they will all differ, but that won't be a problem anymore. Within small distances (such as 10 meters), devices will connect at 1Mbit/s using the Bluetooth standard. Whereas today's Palm devices beam each other, Bluetooth will let them beam to printers, banking machines, vending machines, and airport ticket counters. Outdoors, devices will connect at 56 Kbits/s using packet-based protocols such as the General Packet Radio Service or Enhanced Data Rates for GSM Evolution. Eventually, new protocols will allow connection speeds greater than 100 Kbits/s.[1]

As network protocols evolve, you might not need to buy new devices to benefit from them. Handheld devices will be able to cope with different network protocols and different conditions because the signal processing will mostly occur in software with general-purpose hardware, analog-to-digital cards, and antennas. Technologies such as Software Radio and SpectrumWare will let devices reconfigure and upgrade the network connections without the need to



change the underlying hardware.[3]

The handheld device's user interface will combine voice, text, and graphics—the different media will complement each other. Take for example the MiPad, the next-generation PDA under development at Microsoft.[4] It combines a small color screen with familiar icons (e-mail, to do list, address book, and calendar), hyperlinks, and forms with stylus input and speech input and output. Instead of having to type long text messages on a small touchscreen, you can tap on the entry form and simply say the words—the device will display the words you speak. Tapping on specific areas of the screen triggers the speech-recognition engine, which narrows down the recognizable list of words and improves the language understanding rate. Conversely, some long text messages can be spoken rather than displayed to you. For example, as you point to a location on a map, the device will tell you about the closest restaurant, without obfuscating the small map with unnecessary text labels. It performs speech recognition and natural language understanding on dedicated network servers—it has the ability to understand complete sentences in the context of an ongoing dialogue, which improves technology beyond simply recognizing short phrases. For example, within the context of the Oxygen project, the MIT Spoken Language Systems Group is developing systems that engage in a conversation with users in domains such as weather, flight, and tourist information.[5]

From HyperText to HyperLanguage

Small handheld devices will soon be ubiquitous. Although users currently browse the Web mostly offline, they will browse online much more frequently in the future.

When necessary and appropriate, the small display will feature speech input and output. The next version of HTML must combine several features needed to support small display sizes and multimodal user interfaces. Currently, there are several languages available, each one specialized for a particular modality—for example, WML uses the concept of a deck of cards to model interaction over a sequence of very small pages. The VoiceXML standard describes the dialog flow between the user and the server using speech input and output. A markup language for voice communication must model the various steps in a dialog, such as questions, answers, corrections, and restarts. Macromedia's Flash and the SMIL language can model interactions that involve the simultaneous use of text, graphics, audio, animation, and video.

We are about to face changes across the board in wireless Internet access: infrastructure, devices, and user interfaces will all change dramatically. However, if there isn't a common "glue" between devices and servers connected to the wireless Internet, the growth will be slower than expected.

One of HTML's most important contributions has been its ability to provide a glue—a common language for rendering text and graphics content across different operating systems, different file systems, different display devices, and different languages. The fact that all servers and browsers use a common markup language is one factor in the global spread of Internet usage. HTML's successor should be able to keep up with the multiplication of display devices of different types and capabilities. If we can clearly organize the language into classes of features and embrace, rather than isolate, different user interface modalities, we will witness again a rapid growth, similar to the Web's early days. In the next generation of handheld devices, the user will be able to tap anywhere on a screen and speak to the device, and the machine should be able to either use the display or speak back. We could develop Web applications faster and improve consistency across modalities if a single markup language could model the abstract interaction between a group of Web page elements and the user and model how the various user interface elements implement the interaction. This markup language should be able to combine all the features available in all the various languages, including CHTML, VoiceXML, and WML. In a markup language, each page element or tag might have one or more events that should trigger actions programmable with a scripting language such as JavaScript or ECMAScript.

Inevitably, different devices will have different user interfaces. This means that not all the devices can implement elements of the next HTML—each device might implement a subset of the markup tags. Each specific subset of the markup language should define a class of devices and include the list of tags (page elements) and events each tag supports.

If Web application developers know in advance which tags and events are available in each class, they will be able to rapidly develop Web applications targeting specific classes of devices.

References

1. S. Prandini and D. Sims, "Wireless Carrier Technology Roadmap," www.oreillynet.com/pub/a/308 (current Aug. 2000).
2. *The Wireless Markup Language Tutorial*; www.wirelessdevnet.com/training/WAP/WML.html (current Aug. 2000).
3. J. Guttag, "Communications Chameleons," *Scientific American*, Vol. 281, No. 2, Aug. 1999, pp. 56–57; www.sciam.com/1999/0899issue/0899guttag.html (current Aug. 2000).
4. X.D. Huang et al., "MIPAD: A Next Generation PDA Prototype," to appear in *Proc. Int'l Conf. Spoken Language Processing*, 2000.
5. V. Zue, "Talking with Your Computer," *Scientific American*, Vol. 281, No. 2, Aug. 1999, pp. 56–57; www.sciam.com/1999/0899issue/0899zue.html (current Aug. 2000).