Disconnecting the application from the interaction model

Ing-Marie Jonsson, Neil Scott, Judy Jackson
Project Archimedes, CSLI
Stanford University
{ingmarie,ngscott,jackson}@csli.stanford.edu

Abstract: Disconnecting the applications from the interaction model is based on two principles; to enable users to interact with applications according to their needs, abilities and preferences; and to provide access to a standard version of an application instead of relying on various implementations that are adapted to specific terminals and/or interaction models. It should be possible to access an application regardless of choice of terminal, from TVs to telephones. And also to interact with the application using all possible combinations of input and output modalities.

Based on the assumption that *information is neither accessible or inaccessible - it is the form in which it is presented that makes it one way or the other,* the Archimedes project main focus is to provide access to computers for disabled users. Over time we have seen that numerous technologies and solutions developed for disabled users become conveniences for everyone else. This led us to explore how well suited the initial Archimedes approach would be to implement smart offices and smart houses.

1. Background

The Archimedes Project at CSLI Stanford University seeks to promote equal access to information for individuals with disabilities. Equal access to information requires the ability to communicate rapidly and efficiently with humans and computers. This includes a wide variety of activities, such as: everyday conversations, committee meetings, instructions on the job, requests for information, access to databases, use of application programs, and use of computer-controlled technology such as VCRs, fax machines, and microwave ovens.

Many of the accomplishments in removing barriers to communication have dealt with physical access to the sources of information: building ramps, lowering of ATM machines and telephones, installing elevators, and so forth. Many obstacles remain for people once they want to interact with some of these services, however. A person may have difficulties to use a keyboard, a mouse or dial a telephone. A person may have no speech, unintelligible speech, or use another language. The goals of the Archimedes Project are to develop and transfer to the marketplace productive and efficient ways for individuals with disabilities to communicate with computers, computer-driven devices, and the internet by choosing from a variety of input and output devices and productivity tools that can be transferred across all the computers and computer technologies that the individual wants to use.

This work has naturally flowed into providing supporting environments for aging people. And functions that start out as necessities for disabled and aging people become conveniences for everyone else. And this forms the starting point for us to broaden the project to cover a wider variety of communicating and interactive devices as well as a broader range of users.

2. Overview

Our goal is to create a solution that enables any user to access any computer, information processing device, communication device, tool or appliance according to their preferences. In this we recognize that some devices have various levels of access incorporated in their design. And that we address the need to provide external access for devices that are intrinsically inaccessible or that provide only limited access for some potential users.

We believe that peoples needs, abilities and preferences should define the interface. Today both applications and information, are often presented in a form not suitable or accessible to all users and/or terminals. The chosen interaction model can be to complicated, as with the VCR, or the information is only presented in one format - for instance graphically, which excludes blind users and terminals without a screen.

2.1 Total Access System

The infrastructure and enabling technology we are using is the Total Access System (TAS) [1,2,3]. The TAS separates the needs of the user from the interfacing requirements of the target device. TAS consists of three components, a personal accessor, an access port and a network that connects an accessor to an access port. A typical target system, be it a computer or window shutters, has specialised input and output capabilities. This target system is connected to the TAS via an access port that provides standardized access to all user input and output. Users who cannot, or prefer not to use the input and output capabilities provided by the target system, can use an accessor that connects to the access port. This will give them the ability to fully control the application on the target system using an interaction model of their choice.

Access port: Devices are connected to the TAS via an access port, which provides a low level interface to the applications on that device. The access port is specialised with respect to the connected device, in the case of a computer, the access port emulates keyboard and mouse events, in the case of window shutters, the access port emulates the events that will open and close the shutters.

Accessor: The accessor is a personal device chosen and designed to provide the most appropriate input and output strategies for satisfying individual needs. An accessor can be a any device, a workstation, a laptop, a Palmpilot, or a phone. And it has all the hardware and all the software for the accessibility devices that a particular user needs built into it.

Network: The network provides communication between an accessor and a target device.

2.2 Extensions to TAS

We extended the network to handle communication between all networked devices, that is communication between target devices and communication between accessors. This enabled us to connect any accessor to any target device, and thereby application, on the network. And furthermore also to combine functions from different target devices.

Access port: The access port is extended to contain a description of the device to which it is connected. This information is used by the resource manager of the network to present the device both to other devices and to users.

Accessor: User preferences - the user profile - and software to protect communication and information is added to the accessor. The result being that the accessor no longer has to

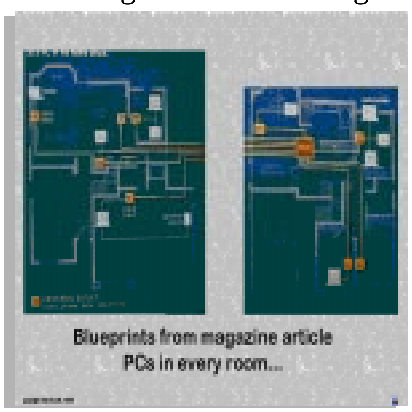
contain all hardware and software that implements the accessibility devices. A description of the preferred interaction model and software to communicate via the TAP is adequate.
Network: The network is extended with network management to provide and maintain a plug-and-play-network of devices. And a resource manager is added to handle available devices and /or applications.

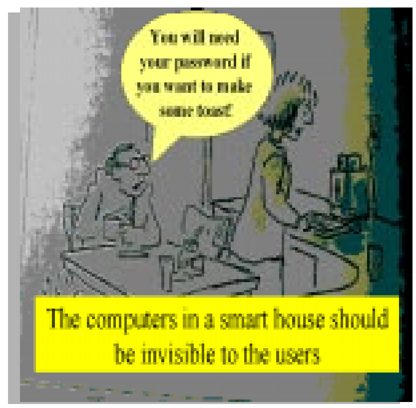
3. Tools and Applications

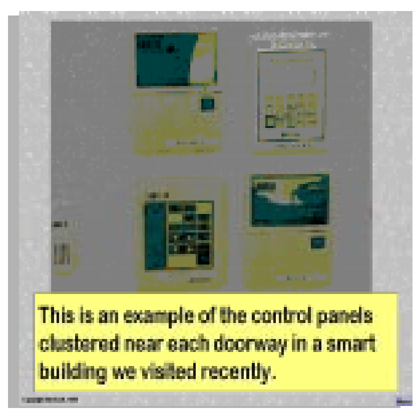
The extended functionality of the TAS enabled us to work with application domains, such as office and home automation. We also found that Multimodal Interaction, where input modalities are mixed and integrated to form commands, was made possible by the ability to communicate between devices.

3.1 Home and Office Automation

Many appliances in both the home and office are computer controlled and there are commercial solutions and initiatives that provides components for automation [4]. There are numerous alternatives suggested for home area networks and control panels for smart offices and homes, and most of them designed without thought of users needs and abilities.







Most of the time both the networks and the control panels seems to be designed by engineers for engineers, with no focus on usability aspects.

The architecture we propose for the smart house and the smart office, consists of three separate parts, an Infrastructure, (smart) Appliances and Human Centered Interfaces.

We mainly focus on interface issues for the smart house and the smart office, but to build prototypes and evaluate our approach we also look at the infrastructure. Here we have identified requirements for network mechanisms necessary to implement smart functions and interfaces that are easy to upgrade and configure. And we also found a need for standardized protocols to communicate between devices.

3.1.1 Human Centered Interfaces

It is essential to design the smart house with an interface that can be adapted to different needs abilities and preferences. Each house or office will have different inhabitants, where culture, education and tradition are but a few areas that will influence the behaviour of the house or office. Also, inhabitants change with time, employees change rooms, families move, kids grow up and parents get old. All of these factors play an important role when determining the architecture for interfaces in a smart house or smart office.

Appliances are manufactured by different companies and they are designed to perform their intended task perfectly using a “standard” interface, there is no focus on interfaces that provides various interaction models. There will be a wide variety of different networks offered for home and office automation, none of which are likely to offer mechanisms to configure interfaces to suit different users.

These are some of the reasons that we believe that TAS, with access ports and personal accessors, offers the right level for interfacing to smart appliances and home/office networks. This will enable a focus on interfaces adapted to particular users needs:

- It is easy to change the interaction modality by changing the input and output devices used by the accessor. And it is possible to change the behaviour of the house/office by combining appliances and functions in new ways.
- It is easy to move or change a function from a portable interface to an interface that is attached to a strategic location in the house or office.
- It is equally easy to change and upgrade devices and functions in the house or in the office given that the home-area and office-area networks implements the plug-and-play mechanisms and resource managers added in the extended TAS.

3.2 Multimodal Communication

We communicate naturally by mixing modalities, and given the mechanisms to do so in the TAS we wanted to explore the benefits to control an application in the same manner. Ongoing research with multimodal interaction models show that people prefer to use multiple modalities when interacting with an application. Changing between modalities when the system fails to understand. When using speech input combined with some other modality, a switch from speech the other modality occurred when expressing deictic information, navigating menu systems and also when entering data that is not visualized on the screen, such as entering keywords to search engines.

Here we looked at a related projects at SRI [5] and OGI [6], to understand the mechanisms of a multimodal interface and to understand where the difficulties involved in interpreting and combining multiple input streams are. And furthermore also to understand how users mix modalities, what makes them switch modalities, and how taking advantage of linguistic convergence can guide human spoken input during human-computer interaction.

TAS supports using multiple modalities, merging input from various devices, combining the input into commands that are forwarded to the target application. Multimodal control of applications seems to be preferred, but user studies are needed to determine the applicability to smart houses and offices.

4. Design, Results and Open issues

In TAS the clean separation of the user input and output from the target device provides the base for disconnecting an application from the interaction model. This design enables access to the communication between user and application on a level where:

- No changes to the application is required. There is no need to implement special versions for different devices and input/output modalities.
- No impact on performance of the target device. There is no need to add software that interfere with the performance or cause conflicts on the target device
- Existing access tools can be used. Input/output devices such as, special keyboards, special screen displays, speech recognisers and synthesizers, pen tablets and eye trackers can be connected as-is to control the application.

The communication between devices and access ports is implemented as simple protocol. This protocol enables us to implement TAS on top of any network, such as RF, fiber, twisted pair, and power lines.

The properties described above and the protocol, provides a sound base for our approach, but there are many more design issues that needs to be addressed. Separation of functionality into hardware and software and where to define interfaces? Separation into open parts and proprietary parts, to define collaborative standards? How to combine multiple input/output modalities, general or application based solution? How to set and comply with user’s expectations, general or modality based solution?

In these design issues we are dependant upon the status and maturity of input/output technology, such as speech recognisers/synthesizers, haptic devices, eye-trackers, cameras, microphones etc. This will in many cases limit the functionality, performance and/or perception of the solution

Previously, TAS has mainly been used to implement speech accessors to general platforms, this has enabled professional programmers and writers to return to work after leaving because of repetitive strain injuries (RSI). There are prototypes and ongoing work with the deaf-to-hearing communicator, a haptic display, and to use head-trackers and eye-trackers as pointing devices. Leveraging on results from these projects, we are now moving into areas of concurrent multimodal interaction and collaborating devices [9].

5. Conclusions and Discussion

We have repeatedly found that we need to interface applications on a low enough level that the access-function will survive between upgrades of the application. TAS, emulating keyboard and mouse events to control applications on a computer seems to be on the right level. We believe that this assumption will be hold for functions offered in the smart house and the smart office as well.

We are not developing networks in the project, but the lack of open solutions for home and office are networks has forced us to implement the infrastructure for the smart house and smart office. By doing this we have identified a need for certain network mechanisms and a need for an open standard for communicating collaborating appliances.

Furthermore, it has become evident that the success of the smart house and the smart office is resting on the acceptance of the users. We need to learn how to build trust and confidence in the idea of a smart house and smart office [8], and we need to focus on users needs, abilities and preferences to implement the concept of a smart house and smart office.

Adhering to users needs, abilities and preferences, we also need to explore more ways of interacting with applications. We need to move away from the focus on the keyboard, the mouse, and a screen as (almost) the only way to interact with computer controlled applications. This means exploring, speech, eye tracking, gestures, tactile input/output as alternatives interactions using keyboard and mouse.

Visiting conferences like CHI, it is apparent that the focus of the computer-human interaction community is still on the screen, the keyboard and the mouse. Contributions presenting speech, gesture and tactile interaction models were clearly outnumbered.

References

[1] Neil Scott, The Universal Access System, Presentation at the American Voice Input and Output Society Conference, Atlanta September 1991

[2] Neil Scott, Using the Total Access System to Access the World Wide Web, Sixth International WWW Conference, 1997

[3] Neil Scott, Universal Speech Access, Proceedings of Speech Tech/ Voice Systems Worldwide, 1992

[4] Home Automation, a WWW collection, *http://www.len.moscow.id.us/ha.html*

[5] Adam Cheyer and Luc Julia, Multimodal Maps: an Agent-based Approach, Proc. of the International Conference on Cooperative Multimodal Communication (CMC95), May 1995. *http://www.ai.sri.com/~cheyer/mmap.html*

[6] Sharon Oviatt, User Centered Modeling for Spoken Language and Multimodal Interfaces, IEEE Multimedia, winter1996, vol. 3, no. 4, 26-35. *http://www.cse.ogi.edu/CHCC/Papers/sharonPaper/Ieee/Ieee.html*

[7] D. Ross, Personal freedom: a wearable interactive universal access device. Proc. of RESNA96 Annual Conference, June 1996.

[8] Michael Coen, SodaBot: A Software Agent Environment and Construction System. Proc. of 1994 CIKM Workshop on Intelligent Information Agents, Sept. 1994. *ftp://ftp.ai.mit.edu/pub/users/mhcoen/aitr-1493.ps*

[9] A. Mane, S. Boyce, D. Karis, N. Yankelovich, Designing the User Interface for Speech Recognition Applications, Results from CHI96 Workshop. SIGCHI bulletin Vol. 28, No 4, Oct. 1996.