An overview of the iTASK System

Neil G. Scott
Director, Archimedes Project
University of Hawaii
August 2004

Introduction

The Archimedes Project developed the Intelligent Tasking System (iTASK) to make computers more accessible to greater numbers of people. This goal was achieved by making computers behave more like people rather than requiring users to learn interaction strategies dictated by hardware and software designers.

For the most part, computers provide a very limited range of options for interacting with humans: a keyboard for entering text, a mouse for moving a cursor to select objects or enter graphics, and a screen for displaying information. Variations such as speech recognition or printing on paper do not change the fundamental modes of available interaction. When communicating with each other, however, humans use many more modes of interaction such as gestures and body language, usually simultaneously.

While many academic research projects are developing interaction techniques that give computers more human-like qualities, attempts to incorporate these techniques into commercial products have so far been quite limited. Failure to meet the expectations of the human operator is seen as the overriding reason for the limited adoption of these techniques. Whenever a machine appears to operate in a human-like way, users expect it to make intelligent decisions similar to those that would be made by a person. Machines that don't live up to these expectations are quickly rejected.

Attempts to make computers behave in more human-like ways have so far been quite limited due to factors such as the following: (a) it is difficult to design products that can accommodated the breadth of human differences that are likely to be encountered in practical situations, (b) unconstrained application domains become too large and complex for practical systems to handle, (c) the demanding software requirements for current designs seriously impacts the performance of the computer, (d) commercial computing devices are extremely cost sensitive and may be uneconomic if large amounts of additional software are required, (e) deliberately built-in obsolescence depends on user dissatisfaction particularly with the user interface, (f) commercial competition is based on proprietary designs that create artificial differences in an attempt to lock users into a particular manufacturer's application or computing environment, (g) critical human interaction technologies such as speech recognition and gesture recognition are still not working sufficiently well for widespread application.

Overview

The iTASK was developed as a way for incorporating small amounts of human-like intelligence throughout distributed information technology (IT) systems. iTASK

Modules provide a means for connecting a multitude of different devices into collaborative networks, as depicted in Figure 1. There are many different configurations for connecting iTASK modules and these will be described later in this document. For the sake of explaining the relationships between the different devices shown in Figure 1, we will use the local workspace as the frame of reference. The local workspace is an interactive device that provides normal computer input and output capabilities for a user. It can range from simple embedded devices like an LCD display and keypad through to a handheld, tablet, notebook or desktop computer. In some situations it is the primary workstation for the user, in others, it is merely a feedback and control device for interacting with other devices.

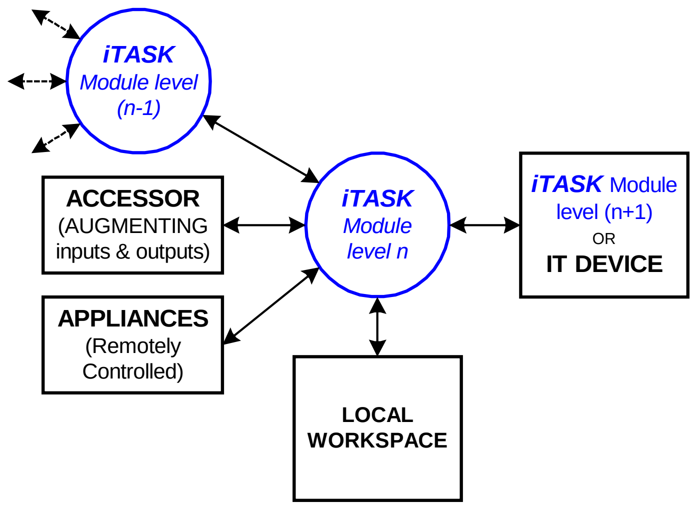

Figure 1. Basic iTASK Configuration

The iTASK module connected to the Local Workspace is said to be at level n with respect to the user. Only one level-n iTASK module can be connected to each local workspace. The level n iTASK module can connect to one device at level (n+1). This may be another iTASK module or it may be an IT device such as a computer, ATM, Info Kiosk, elevator, home appliance, work tool, and so on. Up to 127 level (n-1) devices can be connected to an iTASK module. These may be any combination of iTASK modules, accessors, or appliances.

Accessors can be used to augment the user's ability to interact with the local workspace or with any higher-level devices. Typical accessors provide speech or head tracking inputs for users who are physically disabled or spoken output for users who are blind. The iTASK module allows these augmenting inputs and outputs to be used with any computer or electronic device without requiring any additions or modifications. Figure 2 shows a typical arrangement for using an iTASK module to make a PC (at level n+1) accessible to a person with physical disabilities that limit normal access to the keyboard, mouse and screen. The PC is completely standard and has no additional software. The iTASK module emulates a standard USB keyboard and mouse for inputting commands and data to the PC. The level (n-

1) accessors are selected to provide appropriate alternative input and output options for the user. The level n local workspace provides feedback to the user.

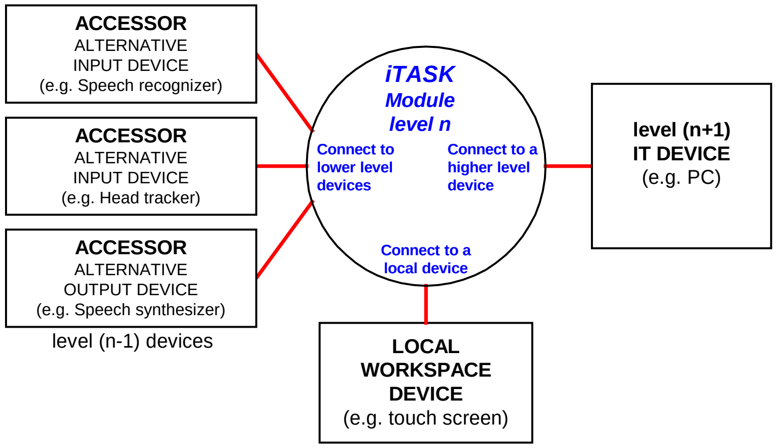

Figure 2. Application of iTASK for Disability Access to an IT Device such as a PC

The major functions performed by an iTASK Module are shown in Figure 3. These functions fall into four main categories: (i) communications between devices and the iTASK module, (ii) networking between groups of iTASK modules, (iii) deriving user intent, and (iv) generating commands and messages. To provide the necessary flexibility in practical situations, the iTASK Module supports a variety of wired and wireless port protocols.

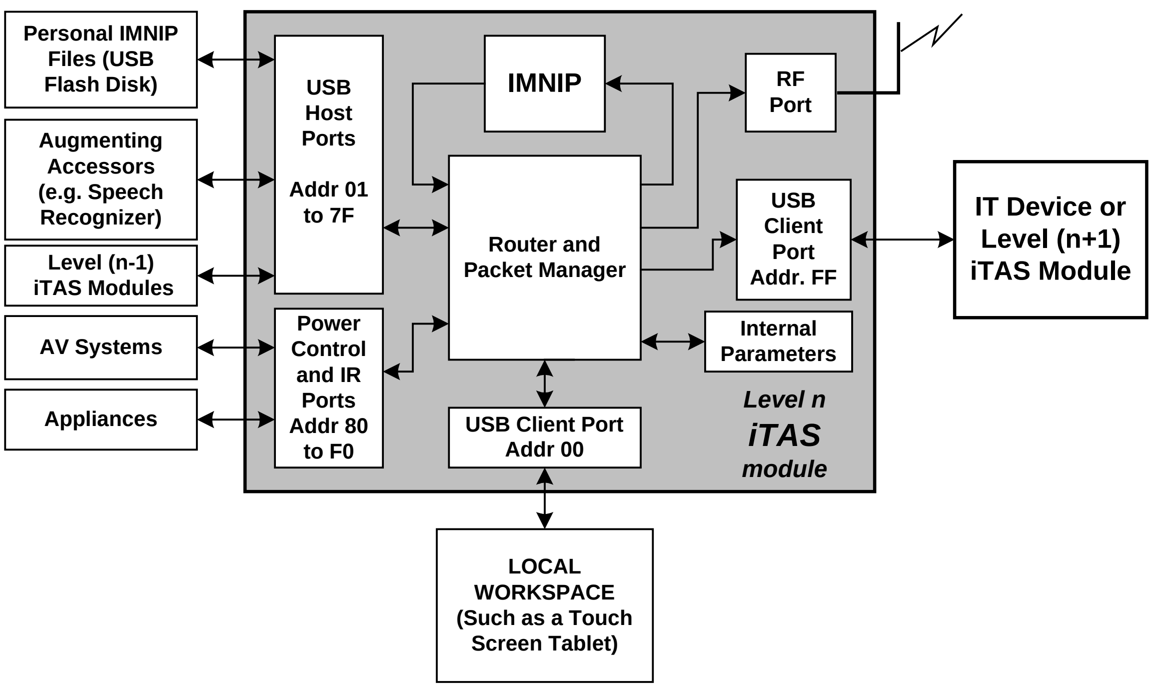

Figure 3. Major components of an iTASK module

Communications

The iTASK Module uses the USB protocol for most of the necessary communications between devices. Several legacy protocols (PS/2, RS232, IR, and X10) are included to simplify interaction with older devices. An 802.15 wireless transceiver is included to enable the system to substitute a wireless link in place of any of the wired connections.

> ***USB Communications*** are used to interconnect iTASK modules and to connect them to accessors, local workspace and target computers. The iTASK Module connects to lower level devices as a USB host and as a USB Client to the local workspace and to a higher-level iTASK Module or IT device. This arrangement enables the iTASK Module to connect to any computer as a standard HID device and to enable other devices to connect to it as standard HID devices.
>
> ***Wireless Communications*** can be supported internally by a Zigbee network built on an 802.15 transceiver, or externally by USB wireless modules that replace any individual USB connection.
>
> ***RS-232 and PS/2 communications*** are provided to allow the iTASK module to work with the existing special access devices that are used by people with disabilities.
>
> ***IR and X10 communications*** are provided to enable an iTASK Module to control home entertainment systems and appliances. The X10 may be replaced by one of the emerging home automation protocols when a clear winner is seen.

Networking

ITASK modules can be used individually, as depicted in Figure 2, but they are more likely to be used in an iTASK Collaborative network as depicted in Figure 4. The iTASK collaborative network is designed to provide advanced personal control functions and operates independently of any standard network that may be used for data transfer between computers and automation systems. The iTASK collaborative network is optimized for low latency so that commands will feel natural and responsive. Accessors that provide special input and output modalities can be added to or removed from the system at any time without any impact on the system. Individual iTASK modules and devices can be moved around the network and will be automatically located and recognized the next time they are used.

Any accessor or local workspace can interact with, and control, any computer, appliance or IT device connected to any iTASK module on the collaborative network.

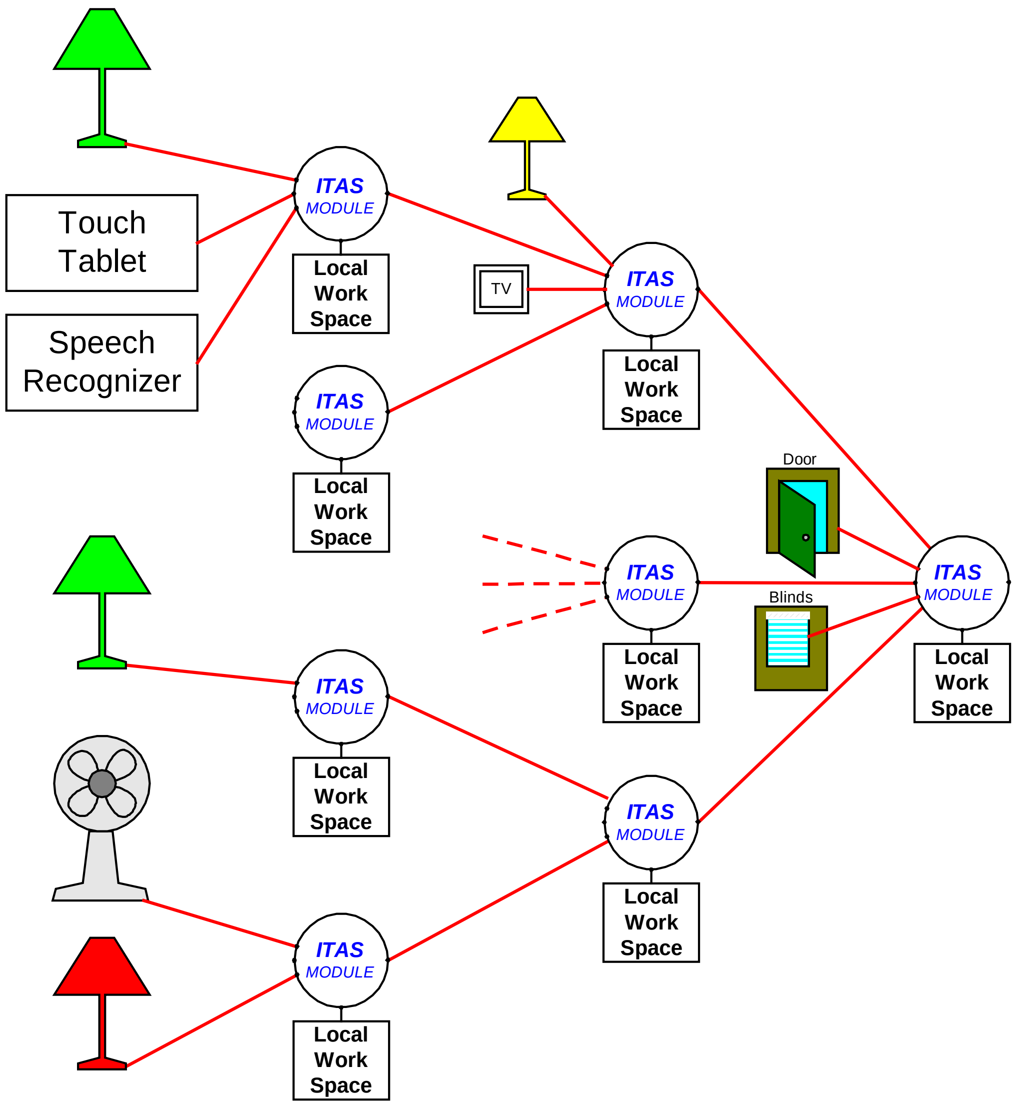

Figure 4. iTASK Collaborative Network

IMNIP

Archimedes researchers have developed a new type of natural language processor called an ***Integration Manager and Natural Interaction Processor (IMNIP)*** that enables people to interact with and control devices using their own natural words and gestures. An IMNIP deduces user intent by analyzing user actions. This ability to derive intent is the key difference between iTASK and other options for creating and interacting with smart environments. Two patents are pending for this system.

Each iTASK module contains two ***IMNIPs:*** one for deducing user intent based on an analysis of one or more streams of multimodal input data along with the current system context; and one for combining messages of intent received from other iTASK modules and generating commands for performing the intended actions on the connected IT devices and appliances.

The major components of an IMNIP are depicted in Figure 5.

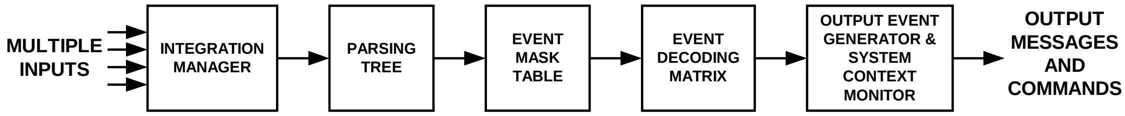

Figure 5. IMNIP components

Integration Manager – receives text messages from multiple accessors and feeds them to appropriate inputs of a parsing tree.

Parsing Tree – parses each word of each incoming message. Recognizes words or synonyms that are included in the corpus for the particular IMNIP.

Event Mask Table – Each recognized word triggers a series of pointers that identify all potential events that depend on the recognized word.

Event Decoding Matrix – The event-decoding matrix links all active keywords and all potential events. When each keyword is recognized, every occurrence of that word in the event decoding matrix is flagged. The matrix is tested whenever an input message is completed to see if any command has all of its keywords flagged.

Output Event Generator and System State Monitor – When the system detects that all of the keywords in a command have been flagged, it checks the current context of the system and generates a context-appropriate message, command, or string of commands (macro). Automatically remembering the context of the system simplifies command entry by enabling a user to enter commands such as “turn on the TV” followed by “turn up the volume,” knowing that the volume adjustment will be made to the TV.

Controlled Device – Commands or messages are transmitted to one or more controlled devices. The controlled device may be connected to any iTASK module that is included in the collaborative network.

Significant Properties of an IMNIP

An IMNIP can recognize user's intentions that have been expressed in many different ways.

a) Recognition of keywords received in any order from a single source.

b) Recognition of keywords received in any order from a multiple sources.

c) Recognition of keywords received in a precisely defined order from a single source.

d) Recognition of keywords received in q precisely defined order from multiple sources.

Each of these options is necessary to cover the different ways in which people express their intentions. The last two are particularly important for recognizing critical commands that could cause problems if accidentally triggered

Multimodal commands simplify intent recognition both for the person and for the intent recognition system. Due to practical limitations in the various recognition technologies, different input modalities handle some types of input better than others. For example, a speech recognizer reliably recognizes spoken commands such as “turn that on” but does not work well for recognizing the names of objects. The iTASK uses the word "that" as a placeholder for the name of an object. The command tells the system that a particular action is intended and that the name of the object to be acted upon is to be obtained either by context or through some other input modality. Typically, a pointing device, such as one of the following, would be used to indicate the name of the intended object:

a) A simple touch tablet overlaid with printed pictures or photographs of potential targets – touching a picture generates a text message containing the name of the device.

b) A tablet computer with interactive images of potential targets that can be selected with a mouse or finger – clicking on, or touching a picture generates a text message containing the name of the device.

c) An image recognition system that can determine where a person is pointing or looking – pointing at an object generates a vector through the object in 3D space. This vector is used to generate a text message containing the name of the object.

d) A head trackers or eye gaze tracker – looking at an object generates a vector into a 3D space that is used to determine its name.

Derived intent can be used in several different ways.

a) Intent can be translated into text messages that can be displayed on the local workspace or sent to other iTASK modules.

b) Intent can be translated into spoken messages for feedback to the user or for building communication devices.

c) Intent can be translated into animated figures that generate sign language for deaf people.

d) Intent can be translated into commands.

e) Intent can be translated into macros that can be expanded into strings of commands that are sent to one or more devices.

An IMNIP can support any size corpus and can be implemented on processors ranging from small 8-bit microprocessors through to large servers.

Summary of iTASK

ITASK modules use a "divide and conquer" strategy to break a single large, and basically unmanageable intent recognition problems into many small recognition problems that can be solved individually.

ITASK modules support many different input and output (I/O) interfaces and allow them to be combined in any way that satisfies the specific needs of individual users. This provides two significant advantages: (i) different users can adopt completely different input and output modalities to interact with the same device, and (ii) individual users can be equipped with a personally configured set of input and output interfaces matched to his or her particular needs, abilities, preferences, language and culture. Input and output devices that use iTASK modules are called Accessors. An accessor may incorporate all of the necessary iTASK functions internally or it may use an externally connected iTASK module

Accessors provide a bridge between user actions and user intent. User actions indicated by one or more of the supported I/O modalities are combined and analyzed to determine intent, and then translated into a text description of that intent. Each accessor handles a particular set of I/O modalities and contains an IMNIP with a corpus of words related to specific I/O actions within those modalities. Intent derivation is distributed across all of the iTASK modules in a system. Each iTASK module derives as much of the intent as it can from the input information it receives and either queries the user for more input or passes a partially completed text message up to the next iTASK module in the system. For instance, a system that enables a person with quadriplegia to use a computer might have a speech recognition accessor and head tracking accessor as depicted in Figure 2. While it is possible for any task to be performed using only speech input, it is usually much easier to split tasks in ways that make the best use of each modality. In this particular example, it is best to use the head tracker to move the mouse cursor, and the speech recognizer to click the button. This translates into actions such as looking at a button on the screen and saying "click this" or looking at a word and saying "make this word bold" or "underline this sentence." The same accessors could also be used to control household appliances and audio/visual devices with user actions such as looking at the VCR and saying "please record the news on channel five."

iTASK supports many different types of target device independently of the input and output strategies adopted by a particular user.

1. A person can use the same accessor or group of accessors to interact with any target device that is equipped with an iTASK module.
2. An iTASK that connects to a target device knows how to interact with all user-accessible functions in the particular target device but it knows nothing about users or other target devices.
3. Each person's personal accessor interface strategy is completely independent of the proprietary interfaces supplied with commercial software.
    a. Each user can develop unique personal strategies to perform particular tasks.
    b. Personal task definitions achieve the same results when used with different applications or with similar applications from different vendors.
4. Personal accessors and task definitions are independent of the type or version of operating system and applications that are used on a target device.

iTASK-to-iTASK Communications

Text messages based on human language is used to pass information back and forth between iTASK modules. This makes it simpler for manufacturer's to build compatible devices since they need only know a list of words that are understood by the system. A major advantage of this approach is that it lessens the likelihood of deliberate obsolescence.

iTASK Vocabulary

iTASK implementations can be scaled to match the vocabulary requirements of any particular situation or implementation. In some situations, the iTASK functions are completely contained within a single intelligent device while, in others, the iTASK functions are provided by a separate iTASK module. The same interaction occurs between iTASK modules regardless of whether they are internal or external. An intelligent lamp, for example, only needs to understand its name and a few commands like on, off, bright and dim. All of the electronics necessary to implement the lamp control electronics, an iTASK, and a simple speech recognizer could be fabricated on a single chip. Audio visual (AV) entertainment systems made up of components from different vendors would most likely use a single external iTASK module that receives command messages from any accessors or other iTASK modules, derives user intent, formulates the desired command sequences, and transmits infrared commands to the targeted AV devices.

iTASK Applications

While the original motivation for developing the iTASK was to make computers more accessible for individuals with disabilities, it is now being used to create intelligent living, learning and working environments that are more accessible for everyone. We believe the iTASK holds the key to solving several serious problems that face society on a global scale.

Disability Access

Countries all around the world are beginning to acknowledge the rights of citizens with disabilities to have the same access as everyone else to resources such as accommodation, information, education, employment, leisure and travel. As yet, most countries are still struggling to find systematic ways for providing these rights. Trying to make large systems universally accessible is extremely difficult because of the wide diversity of disabilities that must be accommodated and the connectivity and protocol differences in the equipment. By separating personal accessor needs from system interfaces, the iTASK offers the most promising option for widespread implementation of universal access. A typical example is depicted in Figure 2. Equipping a disabled individual with a selection of accessors that best overcome his or her specific problems ensures the user will always have the same high level of performance and personal customization regardless of the type or location of the IT device or appliance that is being accessed.

By functioning independently of the devices that are being accessed, iTASK modules preserve the huge investment that has gone into developing accessibility strategies over the past three decades. Present practices in the computer industry are extremely wasteful because perfectly good access devices are constantly being made obsolete when computer operating systems are changed without updating the device drivers that allow them to recognize and interact with peripherals. If there is no available driver, the peripheral device can no longer be used. iTASK isolates the accessible devices from the operating system and allows them to be used with any type of hardware and any type or version of operating system.

Independence for Aging People

A global crisis in the support and care of aging people is developing due to a dramatic shift in the balance of younger and older people in the population. With people having fewer children and living longer, traditional practices for supporting aging people are breaking down. Working families and wide geographic separation make it impractical for aging parents to live with their children. Institutional support is failing because there are not enough assisted living facilities available to accommodate the increasing number of aging people who can no longer look after themselves. Furthermore, even if sufficient facilities could be made available, there would not be enough available or willing young people to provide the necessary workforce.

The most promising strategy for limiting the demand for additional support for aging people is to extend the length of time they are able to remain independent in their own homes. iTASK can play a key role in enabling aging people to easily and confidently interact with smart, supportive living environments. iTASK provides a cost effective way for enabling aging people to control even the most complex technologies, using their own words and gestures.

Educational Applications of iTASK

Another global crisis is looming in education. There are not enough young people being educated with the skills necessary for an information-based society. The technically advanced countries are not educating enough young people with math and science skills, for instance, to provide the skilled workers required to sustain the ongoing development of new technologies. Emerging countries are desperately trying to catch up with the West but, to do so, they need effective systems for teaching literacy in languages such as English or Spanish, and also effective ways for making their populations literate in information technology. Computers have been promoted as the ideal tool for education and vocational training for almost three decades but are still not delivering on the promises. While there are many reasons for this, the main ones are system complexity, built-in obsolescence, and lack of suitable software. Teachers and students spend too much time dealing with the quirks of the technology instead of being able to focus all of their efforts on the material they are supposed to be learning.

Archimedes researchers have developed a new type of educational computer system that uses iTASK modules to solve the inherent problems in current educational computer systems. The "Intent Driven Environment for Active Learning" (IDEAL) impacts computer based education in many different areas:

a) IDEAL is accessible to all students. All of the accessors developed for people with disabilities are automatically available to IDEAL.

b) IDEAL separates interaction and control functions from information delivery. The control functions are handled by an external iTASK collaborative network and the information delivery is handled by a standard WIFI network. This separation ensures that evolution and changes to commercial networking standards have no direct impact on the IDEAL system.

c) IDEAL software works on any computer or device that can access a network, run a web browser and play Flash movies. This leads to an extremely important result: IDEAL eliminates version obsolescence since the brand, version or age of the computer and its operating system have no direct impact on its ability to run IDEAL software.

d) Using the IMNIP's ability to integrate multiple streams of real-time input data, the IDEAL system incorporates real-world experimentation and measurments directly into learning materials. Linking practical hands-on activities directly to theory strongly reinforces the learning experience and makes it much easier to remember.

e) IDEAL uses the IMNIP's ability to form collaborative networks to enable students to work alone, in small groups, in large groups or as a complete class.

f) IDEAL enables the teacher to interact directly with the whole class, with groups of students or with an individual student.

Simplifying Complex User Interfaces

Aerospace and automotive design engineers use advanced Computer Aided Design (CAD) systems to do most of their work. As the capabilities of these systems has grown, so too have the demands on the user interface. User productivity has been severely impacted by the complexity of these interfaces. Companies like Boeing and Ford have identified two particular problem areas:

User interface options are inadequate or difficult to use when manipulating graphical representations of virtual objects. a] Connecting special purpose user interface devices to off-the-shelf computer systems is complicated by the need for special driver software, compatibility issues and unintended interaction beween different input modalities. iTASK solves these problems by isolating the low-level functions of each input modality within its own accessor. An IMNIP within each accessor translates personal operating strategies into standardized descriptions that can be cleanly integrated with other input modalities through a higher level iTASK module. The IMNIP in the iTASK module that combines the multimodal inputs determines user intent and triggers the intended commands or macros. This approach enables individual CAD users to easily develop highly personalized multimodal interaction strategies using completely standard hardware and software components.

Engineers are overwhelmed by frequent updates and modifications to the user interaction models used in the software. to keep track of changes in updated interfaces. They no sooner master one interface than a new version is introduced, forcing them to waste time searching through new menu structures to locate familiar functions. iTASK solves this problem by introducing a layer of abstraction between a personally organized interface and the real interface. Individual designers often use quite complex macros to streamline tedious or complex tasks. These macros are often broken, however, by upgrades to the CAD software. Since the iTASK isolates the user interface from the details of the application software, updates to an application such as CAD can be handled by simply updating the command definitions in the IMNIP. This makes transitions between different versions of the CAD interface completely invisible to the user. The same benefits apply for any type of application.

Access to Public Services

Organizations such as telephone companies that must provide services to large numbers of people are challenged by the problem of how to accommodate widely different personal needs when catering to the population at large. While some of these personal needs can be globally accommodated by judicious design choices, many of the special needs can be met only on a case-by-case basis. Several large organizations have been exploring the use of user profiles that dynamically adapt standard interfaces to accommodate individuals with special needs as and where needed. The storage, detection, maintenance and use of user profiles adds greatly to the complexity of interfaces that must be used by large numbers of people. Inevitably, there are compromises that limit the effectiveness of the interface for all users.

The iTASK provides a much simpler universal accessibility solution without compromising the interface for other users or necessitating the service provider to obtain, maintain and protect personal interaction data. The specification, provision and maintenance of personal interfaces remain the responsibility of the individual user in the same way that spectacles are the responsibility of the individual. Each user is free to choose interaction modalities that best match his or her unique needs and have them implemented in whatever way they prefer and can afford. Benefits to the user include: having a single, consistent interface that works in all situations, having the freedom to mix and match different interaction modalities, the need to learn only one interface, the security of knowing that they do not appear to be different to any other person using the system and don’t need to provide additional personal data, and, in most situations, having an interface that is much smarter and easier to use than the interface supplied by the service provider.

