18

Sensory Augmented Computing

B. Schiele

"Computing belongs in furniture and foot-ware much more than it does on the desktop."

MIT Media Lab 1995

18.1 Introduction

The vision of Ambient Intelligence is to embed computing and communication capabilities into nearly everything, namely the environment, objects, or even clothing. Great advances in mobile computing, communication, and device technology for example allow to access a large variety of computing services without the constraint of sitting in front of a desktop computer or being in a particular smart or intelligent environment. However, many research challenges remain. A particularly challenging research topic within Ambient Intelligence is the question of how to interact unobtrusively and in a seamless way with users. Quite obviously current desktop interaction techniques do not generalize well to the more versatile settings of Ambient Intelligence.

In order to realize the vision of Ambient Intelligence context awareness is often seen as a means to make the computing tasks sensitive to the situation and the user's needs. Ultimately, context awareness may support and enable seamless interaction and communication between human users and ambient intelligent computing environments. The notion of implicit interaction, for example, suggests to sense "an action, performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input" (Schmidt 1999). That means, the user interacts with physical objects in a natural way, but a computer system also can extract inputs from these actions. Others such as Hinckley et al. (2000), Schmidt et al. (1999), Rekimoto (1996) and Harrison et al. (1998) propose physical interaction, e.g., tilting a device for configuring a device's functionality, as new and convenient forms of interaction for mobile user scenarios. System input generated from interaction with physical objects has been used for coupling physical objects with computer applications such as tangible user interfaces (Ishii and Ullmer 1997), computer-assisted furniture assembly (Antifakos et al. 2002), tracking a patient's medicine cabinet (Siegemund and Flotifakos 2003) or workflow monitoring in a chemical lab (Arnstein et al. 2002).

Empowering a computer system to process physical user inputs requires augmentation of today's computer nerve-endings, such as mouse and keyboard, by sensors: perception, reasoning, and interpretation of real world phenomena enable computer systems to observe the user's physical environment and serve the user in more appropriate ways than it is possible today. Current technology offers an impressive range of sensors and sensor modalities. Furthermore, it is also widely believed that many sensors will become so cheap and small that they can be deployed unobtrusively and in large numbers. Computing which has access and makes use of this vision of ubiquitous sensors is what I call sensory augmented computing.

This chapter is devoted to Sensory Augmented Computing since the accessibility to a large variety and diversity of sensor information has great potentials to change the way we interact with computers. The general vision is that the use of sensors and elaborate perception techniques will play an important role in order to derive interesting and high-level context. Using a multitude of sensors, distributed throughout the environment may enable applications to be aware of the situation of the environment and the users. Obviously, a computer interface which uses user models, contextual, and situational information to its fullest is a long-term research goal. The chapter starts with an overview and classification of interesting research in the area of sensing for Ambient Intelligence. Then we describe in more detail one of our own sensory augmented computing research projects, namely multi-sensor context awareness for proactive furniture assembly.

18.2 Sensing Opportunities for Ambient Intelligence

As mentioned above, we strongly believe that computers should have access to a large variety of sensors in order to see, hear, interpret, and eventually understand more about humans. Already today, there exist many sensors and sensing devices. Besides the prominent examples of vision and audio sensors there exist a large variety of other sensor modalities, which could be embedded in many objects and devices. This section gives an overview of some work related to sensory augmented computing for Ambient Intelligence. For the discussion we will characterize sensing opportunities with two criteria: the logical view and the physical view.

18.2.1 The Logical View: Dimensions of Sensing

The first criteria we use to characterize various sensing technologies is the type of information that can be extracted from them. Since we are particularly interested in information those sensors contain about humans we concentrate on various aspects of "human-sensing." We have identified the following six different aspects (Michahelles and Schiele 2003): human ID, object usage, location, bio signs and emotions, human activity, and interaction among different humans.

Human or user *ID* has been widely used, e.g., for customizing and personalizing services without requiring explicit user inputs (Richardson et al. 1994; Bohnenberger et al. 2002). In fact, we use a more general definition of *ID* ranging from differentiating people to actual person identification. The second dimension is *location*, which is the most prominent and widely used form of context information (Abowd et al. 1999) used in ubiquitous computing applications such as Want et al. (1992) and Davies et al. (1998). It does include 3D coordinates but also semantic location descriptions. The third dimension, *activity*, describes the task the user is performing which ranges from simple moving patterns (Van Laerhoven et al. 2001) to precise job descriptions. The fourth dimension, *object use*, comprises collocation of a user to an object (Richardson et al. (1994), carrying an object (Langheinrich et al. 2000) and the actual use (Antifakos et al. 2002). The fifth dimension, *bio signs/emotions*, describes the internal state of the user. Research in this area is still in its infancy. First results could be obtained with heart rate and skin-resistance, for reasoning about a user's affects (Picard and Klein 2002). The sixth dimension, *human interaction*, characterizes the relationship between humans including simple collocation, listening to a speaker, gazing, and actual interaction as discussion.

18.2.2 Physical View: Placement of Sensors

We differentiate four different sensor placements. *In environment* refers to stationary installed sensors, e.g., in the floor, walls, where placement can only be changed with effort. Whereas *in environment* installations work with all users at the stationary location *on human* has the opposite characteristics: only users wearing the sensors can participate, therefore they are not bound to a location. *On object* is in between the two previous categories, as objects can be personal and can be carried by a human, such as a key, but also stay at a certain location, e.g., chair. This distinction depends on the object. Additionally, *mutual collaboration* defines sensing system that always require more than one unit in order to operate properly, e.g., triangulation of signal strength for localization.

Using the six logical dimensions of sensing and using the four sensing placements (*in environment*, *on object*, *on human*, and *mutual collaboration*) the following discusses some work of sensory augmented computing for Ambient Intelligence.

18.2.3 Sensors in AmI Research

For recognizing a person's *ID* the best results can be achieved with biometric sensors (Wayman et al. 2003), such as fingerprint or iris scan. Methods based on vision (Donato et al. 1999), audio, or load-cells embedded into the floor (Cattin 2001) deliver less quality. Inertial sensors placed *on object* and *on human* can be used to sense typical movements, e.g., perceiving the signature at a pen, for identification. Scheirer et al. 1999) report on using vision. Kern et al. (2002) report on using audio worn *on human* for people identification. Location systems as described by Hightower and Borriello (2001) can also be used for identifying people at different locations. These systems require both sensors worn by humans and installed base stations.

For detecting *object use* load-cells (Schmidt et al. 2002) have been proven useful installed both *in environment* and *on object*. Object classification with vision is well established in static settings, occlusion during dynamic use can be challenging. Audio is another option, if the *object use* generates characteristic sounds. Inertial force sensors placed *on object* have been successfully used for *object use* as reported by several authors including Hinckley et al. (2000), Schmidt et al. (1999), Rekimoto (1996), and Harrison et al. (1998). Obviously, motion during *object use* can be also sensed *on human* but with less quality. Audio *on human* is also possible (Lukowicz et al. 2002) but is an indirect measurement compared to *on object* placement. Location systems can give hints as well for *object use*, e.g., teleporting X Windows to user's current location (Richardson et al. 1994).

Location is the most explored sensing dimension in ubiquitous computing. Load-cells (Schmidt et al. 2002), vision (Brumitt et al. 2000), and audio (Darrell et al. 2001) have been explored in different projects. Coarse location can be also gained through passive-infra-red sensors, mechanical switches, or IR-barriers. *On object* and *on human* the primary outdoors is GPS, more low-level information delivers humidity, inertial, or pressure sensors (Vildjiounaite et al. 2002). The variety of location systems based on *mutual collaboration* is huge: differential GPS, ultra-sound, radio, etc. There exist various systems that integrate several of the standard techniques such as GPS, GSM, or WLAN; see for instance LaMarca et al. (2004) and Fox (2003).

Sensing *bio signs/emotions* with *in environment* sensor-settings is difficult: Donato et al. (1999) and Fernandez and Picard (2003) report on vision and audio for reasoning on user's *bio signs/emotions*. Augmented objects measuring force and touch (Ark et al. 1999) can give some hints about *bio signs/emotions*. However, most promising are *on human* measurements such as reported by Healey and Picard (1998) and Michahelles et al. (2003).

Activity can be well sensed with special purpose system, such as commercially available smart white boards. Load-cells, passive infrared, pressure and capacity sensors can be used for low-level detection only. *On human* sensing has been well explored for motion activity (Farringdon et al. 1999). Location system can give hints reasoned from semantical location descriptions.

Interaction among humans has not been explored very well. *In environment* sensing systems based on vision, load-cells, and audio could help to perceive characteristics of interaction, such as collocation, gestures, or speech. The *on object* field is blank, as objects are not involved here. *On human* the same sensors can be used as for *activity* if measurements are correlated among interactors. Location systems do not really help here, as collocation is not significant for interaction.

18.2.4 Discussion

Quite interestingly each sensor placement *in environment*, *on object*, *on human* and *mutual collaboration* is meaningful for at least one of the six sensing dimensions. *In environment* placement is the primary choice for *ID* sensing. Regarding the other five sensing dimensions the power of *in environment* placement is mainly based on video and audio methods. However, the perception quality relies on computational expensive methods. Nevertheless, once an environment has been augmented with sensors, e.g., Smart Rooms, applications work without additional instrumentation of users or objects. It also can give hints for human–human interaction. As our focus is on human sensing it is obvious that *on object* points out useful for *object use*. As physical interaction with everyday objects mostly involves movements, such as grasping, moving, or turning the dominant sensor choice for *object use* are inertial sensors and force strips to a certain extent. *On human* is suited for direct measurements of human-centric sensing aspects, such as *bio signs/emotions* and *activity*. Applicable sensors include inertial sensors, audio, biosensors, and also video to a certain extent. *Mutual collaboration* sensors, such as the location systems, have similar characteristics as video and audio with even lower quality: location can provide coarse information about *object use*, *activity*, and *in environment* due to the strong implications of physical location. However, in direct comparison with *on object* and *on human* sensing location system are in an inferior position.

18.3 Proactive Furniture

As an example of sensory augmented computing this section describes a context-aware system enabling proactive instructions for assembly. We demonstrate the system with flat pack furniture but the approach generalizes to a variety of cases where there exist instructions that need, should, or could be followed by a person.

Proactive instructions are taken as example for various reasons. First of all, many of today's instructions, handbooks, and even reference manuals are rarely used even though many of us would and could profit from getting the appropriate instructions at the appropriate moments in time. Secondly, the level of instructions required varies depending on particular person. So the system should proactively adapt its instructions to the current user. Thirdly, most of today's instructions are mostly linear in the sense that they do not model and allow variations in the way or the order people perform actions. This is a common problem for many paper-based but also computer-based instructions. And fourthly, since instructions are typically detached from the physical object the users have to make the connections between the "virtual instructions" and the physical objects and actions themselves. For all those reasons modeling the various states and actions of an assembly, recognizing them using multiple sensors embedded in the involved physical objects, and giving appropriate instructions and feedback depending on the actions performed by the user is a highly promising approach to overcome many problems with today's instructions.

As the running example for proactive instructions or more generally for proactive guidance we chose the example of presenting instructions during the assembly of *Do-It-Yourself* furniture. We use the parts of the assembly as the interface to the instructions, by sensing what the person is doing and how far he or she has got with the assembly. Quite obviously the idea of perceiving the user's actions and presenting instruction based on the actions applies to many other applications. The Labscape Project, described by Arnstein et al. (2002), is such an example, where the actions in a biological laboratory are monitored. Here, both a logbook of daily activities is created, and instructions are presented in situ. Other examples in the field of aircraft maintenance have also been discussed; see for instance Lampe et al. (2004). More examples will be presented later.

The two main questions and challenges we are addressing in the following section can be formulated as follows: (1) Can sensing be implemented reliably enough, so that the user can interact with such systems? (2) How can feedback be given to help the user understand what the system is doing? In our example, the second challenge translates into, how can the assembly instruction be presented. Section 18.4 gives an overview over the sensing task and the technology used. Section 18.5 presents our vision of situation-aware affordances and summarizes a user study in which we compare our prototype with traditional paper-based instructions.

18.4 Sensing a Furniture Assembly

This section starts off with a brief overview of the assembly instructions provided by IKEA. We propose our own assembly plan, offering more possible solutions of assembling the wardrobe. How this plan is used and which user actions need to be perceived is explained and demonstrated with experiments.

18.4.1 The IKEA PAX Wardrobe

The PAX wardrobe is a simple wardrobe that can be used for many different purposes. Many types of shelves can be inserted at different heights. Our discussion will be concerned with the assembly of the main wardrobe without its shelves.

The wardrobe consists of six wooden boards, two metal corners, cams, cam-bolts, dowels, screws, and nails. For a standard assembly of the wardrobe, a screwdriver and a hammer are the only tools required. If the wardrobe is mounted to the wall a drill is also needed. In Fig. 18.1 Steps 1 up to 6 of the assembly instruction are depicted. Steps 1 and 2 show the preparation of the

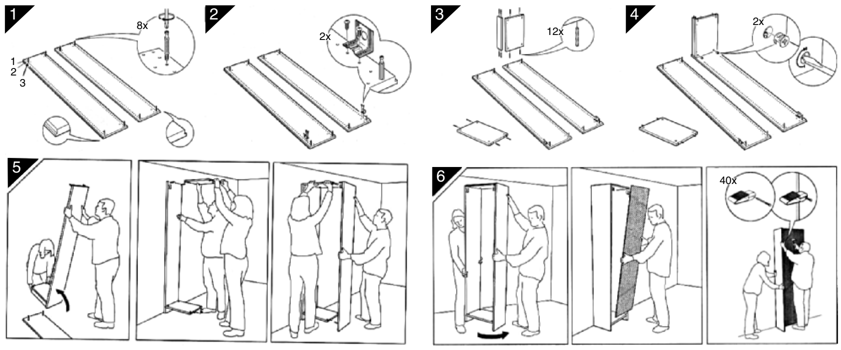

Fig. 18.1. Steps 1 to 6 of the IKEA assembly instructions. Steps 1 and 2 (upper row, left) show the preparation of the two sideboards. Steps 3 and 4 (upper row, right) show how one horizontal board and the base strip are attached. Steps 5 and 6 (lower row) of the assembly instructions show how the compound is lifted into an upright positions, the remaining sideboard and horizontal board are attached, and the back panel is nailed on (Reproduced with the permission of the IKEA corporation.

two sideboards. The first step is to insert the four cam-bolts in each board at the right positions. Then the two metal corners have to be attached. Steps 3 and 4 show how one of the horizontal boards and the base strip are attached to a sideboard. Before attaching these boards they need to be prepared with dowels. The last step is to tighten the cams to fix the board. Step 5 shows how the compound part from Step 4 is lifted into an upright position. It is important to lift the wardrobe into an upright position before continuing with the assembly, because in rooms with low ceilings it is not possible to lift the fully assembled wardrobe. After lifting the wardrobe upright the top (horizontal) board subsequently the remaining sideboards are attached. Step 6 shows how the back wall has to be nailed on by using the nail-holder to position the 40 nails correctly.

Looking at the instructions offered by IKEA it is clear that they are well optimized. However, they only represent linear sequence of actions similar to many other types of instructions. In its current paper format, the user has to make a connection between the physical world of the wardrobe and the virtual domain of the instructions explicitly. Making this connection is typical burden, which is not necessary. To overcome these problems we suggest an assembly plan, which models all possible paths the user can take, in the following section.

18.4.2 Assembly Plan

To successfully assist the user the assembly plan has to modeled. Here we present an assembly that models the different states of the assembly as well as the different actions performed by the user. Those actions are modeled as state-transitions.

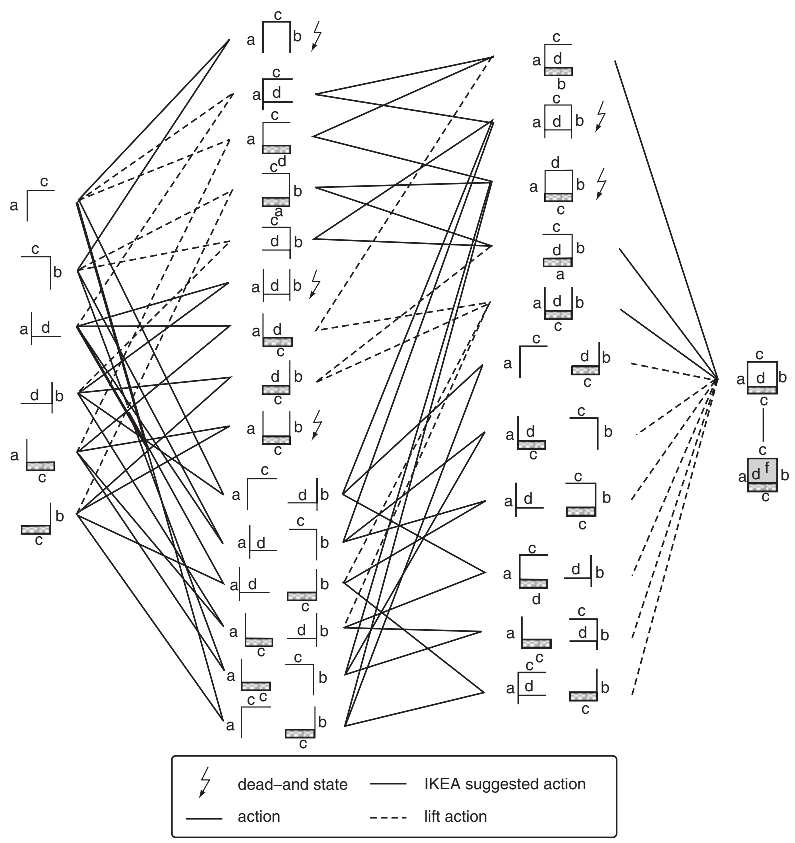

Fig. 18.2. Assembly plan for the IKEA PAX wardrobe

To create such a plan, we gave the wardrobe parts the following identifiers: sideboards a and b, horizontal boards c and d, base strip e and back panel f. Figure 18.2 shows the main part of the full assembly plan. The graph consists of icons representing partial states of the wardrobe and interconnecting lines. These lines describe actions that need to be completed to move from one state into the next.

Actions always consist of joining a not previously used board to the compound depicted or joining the two compounds together. The actions of preparing the boards (adding dowels, cams, or screws) are not shown in the graph. The only restriction for those actions is that they have to be completed before the board is used. We had to distinguish between the situations of connecting the sideboard a and the one of connecting sideboard b to the horizontal board. This is important since the board orientation matters.

The graph is read from left to right. The wardrobe can be in any one state in one column at each step in time. An action transfers the wardrobe to the next column on the right or to the end state. The dashed lines in the graph symbolize actions in which the user has to lift up one of the compound parts before adding the extra board or joining the two compounds. This is due to the restriction that the compound has to be lifted before the two horizontal parts are added to one sideboard. States from which one cannot continue are marked as dead-end states with a lightening-bolt. If the user reaches such a state he has to go one step back before he can continue.

The plan offers a total of 44 possible assembly sequences, and shows 14 sequences leading to dead-end states. The four sequences marked with thick lines are the ones proposed in the assembly instructions by IKEA. It is worth noticing that the lift action occurs in the second step in all the IKEA sequences. Besides the IKEA-sequences there are four other sequences that also have the lift action as the second step. From our experience we can say that these sequences are just as simple for the user to set up as the ones proposed by IKEA. Knowing with which state the user is occupied, it is possible to implement a variety of different types of assembly instructions. The user can be given information about the best action to take next. Alternatively, the user could only be informed when he or she has arrived in a dead-end situation. For quality monitoring reasons, simply noting all states the user passed through may be of interest. The following is concerned with how the assembly state is inferred using sensors.

18.4.3 Observing the User's Actions

The following presents a sensor-based approach for perceiving the actions in the assembly plan. First, we show how the actions can be subdivided into partial actions. We then show which sensors can be used to recognize the partial actions.

In Table 18.1 the actions that have to be recognized to trace the full execution plan are listed. Preparing the different parts, lifting a compound part into an upright position, joining a sideboard or the base strip to a horizontal board have to be detected. Furthermore, we distinguish the case when the second sideboard is added to the compound and the action of nailing the back panel to the rear of the compound resulting in the finished wardrobe.

Most of the actions in Table 18.1 can be subdivided into partial actions. These actions are relatively simple and self-contained such as tightening a screw, hammering in a dowel or a nail, turning a board, or joining two parts together. In the following we will show how these simple partial actions can be detected using the appropriate sensors. In the third column the table gives the sensor configuration we used in our experiments. The fourth column provides some sensor alternatives, which may influence the precision of the perception and the total sensor cost.

Taking a look at the simple example of preparing a horizontal board (inserting four dowels), we see that this can be recognized using only one accelerometer attached to the board itself. Alternatives would be to enhance the hammer with an accelerometer or to use an electric contact that reacts

Table 18.1. Assembly actions and possible sensor configurations

Action	partial actions	our sensor configuration	alternative sensor configuration
prepare sideboard	screw 4 cam bolts screw 2 screws	screwdriver with gyroscope	contact sensors
Prepare horizontal board/base strip	hammer in 4 dowels possibly turn board	accelerometer on board	accelerometer on hammer dowel contact sensors
lift compound part		accelerometer on board	
join sideboard and horizontal board	join parts tighten 2 cams	force sensors screwdriver 2 accelerometers	distance sensor contact sensors
join base strip to sideboard	Join may be hammer	force sensor 2 accelerometers	
nail wall to back	hammer 40 times	accelerometer	

as soon as the dowel has been fully inserted. What makes this example interesting is that one can insert the four dowels in many different ways. As the dowel insert-points lie on opposite sides of the board, one usually has to turn the board during its preparation. This action of turning the board can also be easily recognized using the accelerometer attached to the board. How many times and when exactly the board is turned, however, can be varied by the user. Next we discuss an approach for how such different sequences of partial actions can be incorporated into the perception process.

18.4.4 Sensor Experiments

In this section we show how sensors can be used to detect the partial actions presented above. After that, we describe how Markov chains can be used to detect sequences of these partial actions. For our experiments we used off-the-shelf sensors. The available sensors were attached to parts of the wardrobe and to the tools used during assembly. Data was collected on a standard PC. In a second step we then developed an interactive prototype using wireless communication technology developed in the Smart-its project (Smart-Its, online; Beigl et al. 2003).

Figure 18.3a shows two 2d-accelerometers connected to the sideboard and horizontal board of the wardrobe. The accelerometer used is the MEMS accelerometer ADXL202 from Analog Devices on the evaluation board.

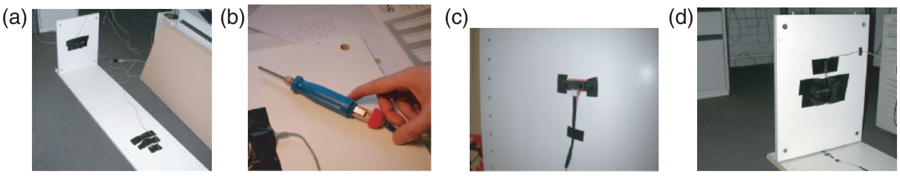

Fig. 18.3. (**a**) Horizontal board and sideboard both equipped with 2d-accelerometers. (**b**) Screwdriver enhanced with gyroscope. (**c**) Sideboard with attached distance sensor. (**d**) Horizontal board with accelerometer and sideboard with attached force sensor

Figure 18.3b shows a screwdriver enhanced with the gyroscope ENC-03JA from Murata. Furthermore, Fig. 18.3c shows the Sharp GP2D12 infrared distance sensor attached to the sideboard. In Fig. 18.3d a horizontal board being joined to a sideboard is shown. The horizontal board is equipped with an accelerometer and the sideboard is enhanced with a standard force-sensing resistor (FSR) to measure when the boards are joined.

18.4.5 Detecting Partial Actions

It is worthwhile going through a few of the experiments to see how the partial actions can be detected. For example, the preparation of the sideboard consists of screwing in four cam-bolts and two screws. In Fig. 18.4a the output of the gyroscope enhanced screwdriver is plotted over a time period of 5 min. By calculating the standard deviation over a time-window one can easily recognize that the user was using the screwdriver six times. We also conducted experiments to detect whether a user is opening or tightening screws. It shows that these actions are also easily distinguishable as the standard deviation of the gyroscope signal is clearly negative when tightening a screw and clearly positive when opening a screw.

The action of joining a horizontal board to a sideboard is shown in Fig. 18.4b. This plot incorporates the output signal of the gyroscope-enhanced screwdriver, the force sensor, and the two dimensions from the accelerometer attached to the horizontal board. One can reconstruct that the horizontal board was moved into place, and then the screwdriver was used to tighten the cams, which in turn increased the pressure on the force sensor. In another experiment we included the infrared distance sensor mentioned above. We used this sensor to detect the orientation of the horizontal boards with respect to the sideboards. To do this, IR-receivers were placed on both sides of the horizontal boards.

Beyond the sensors described above, Table 18.1 presents a variety of alternatives. Depending on the required system reliability and the product cost, different design choices can be made. For example, metal contacts could be used as a very cost-efficient sensor, to detect when two objects are connected.

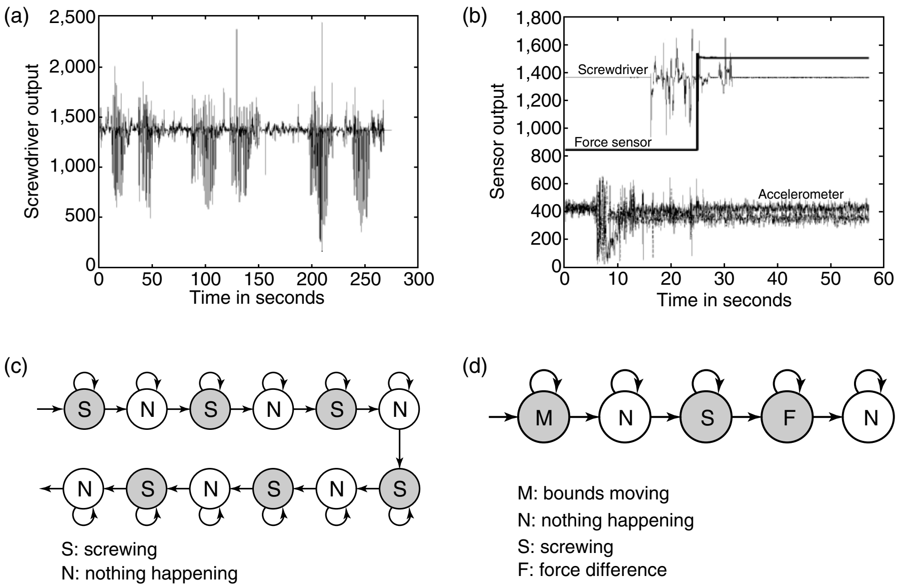

Fig. 18.4. (a) The output signal of the gyroscope enhanced screwdriver during the preparation phase of a sThis action consists of tightening four cam-bolts and two normal screws. **(b)** The outputs of the screwdriver, force sensor and accelerometer are plotted during the action of joining the horizontal board to the sideboard. **(c)** and **(d)** The Markov chains for the sideboard preparation and the joining actions, respectively

From various experiments we can conclude that all partial actions necessary can be perceived quite easily and reliably; see also Table 18.1. There are surely still problems and ambiguities, for example one cannot differentiate if someone is hammering in a nail or a dowel by using only an accelerometer. Similarly one cannot precisely count how main nails/dowels have been hammered in, as the user might start hammering, take a break, and then continue hammering in the same part. Nevertheless, one must say that the partial actions being recognized can be distinguished to an adequate degree. Following, we present a method that allows us to detect the complete actions by modeling valid sequences of possible partial actions.

18.4.6 Detecting Complete Actions

One approach for detecting actions with a higher confidence is using Markov chains to model sequences of partial actions. With this technique the chronological order of the partial actions can be taken into account. Simple Markov chains were designed for each action. Figure 18.4c shows the Markov chain for the action of preparing the sideboard. It incorporates the partial action of tightening a screw, six times.

Similarly the Markov chain for the joining action is shown in Fig. 18.4d. Here we model how the pressure on the force sensor rises when the screwdriver is used.

To test our Markov chains we performed the actions described in Table 18.1 several times and recorded them using the described sensor configuration. Applying simple classifiers to the data, such as Bayesian or threshold classifiers, we generated sequences of partial actions. These sequences were then fed to the Markov chains for recognition. The actions were all easily recognized. This is due to the fact that the detection of the partial actions is quite reliable and that the order of partial actions is distinctive for each action.

18.5 Situation-Aware Affordances

The previous section showed how the actions modeled in the assembly plan can be divided into partial actions and how the necessary action sequences can be recognized using Markov chains. As a result the system can recognize the various states as well as the actions of the assembly. In this section we present the concept of situation-aware affordances as a technique to make physical objects more interactive. We start with a general discussion of the concept and then show, how the concept can be implemented for the specific example of proactive furniture.

18.5.1 General Concept

Interactive environments such as the Aware-Home (Abowd et al. 2000) and smart offices (Johanson et al. 2002) introduce new and diverse tools into everyday life. It is crucial to design these environments in a way that people can explore, understand, and predict functionality and effects. Beyond training and instruction manuals, appropriate design has proven essential to make such systems more usable and intuitive (Van Welie 1999). One way to approach this is to provide objects with clear affordances as have been defined by Gibson (1986) and made popular later by Norman (1988). Affordances give people visual cues about how to use objects and thus offer a simple form of instructions. For example, buttons are here to be pressed, and a coffee-cup handle is here to lift up the coffee-cup.

Although carefully designing objects with discernible affordances can lead to better results in many cases, affordances are mostly static and bound to a single object. In contrast to that, in interactive environments objects can adopt different roles at different times and may be involved in multi-step tasks. Classic object affordances can display information regarding the use of single objects. However, what would be needed for interactive environments is a situation-aware notion of affordances that can also reflect relations among several objects and changes in the environment.

Object affordances are closely related to usage instructions. Several guidelines for designing instructions have already been proposed. Actually, an ideal design of objects should not require any instructions at all: the user should be able to guess and understand the functionality at a glance. However, it is hard to eliminate instructions in general. It would already be an achievement to integrate them into the related objects. Instructions could be split into smaller portions – hints – that subtly but infallibly guide users towards correct conclusions. These hints should be tailored to the users momentary task. Each hint helps in one dedicated situation in contrast to manuals covering all error cases.

As an overall requirement, successful interactive instructions have to follow three principles described by Constantine and Lockwood (2003): explorability, predictability, and intrinsic guidance. Explorability enables users to explore, experiment, and discover functionality without penalization of unintentional or mistaken actions. Predictability builds upon intuition: a user can draw conclusions based on first impressions without having to understand all details. Intrinsic guidance is integral and inseparable of the user interface. Instructions are provided as needed without requiring any special action or initiative on the part of the user.

The following presents a specific solution on how an implementation of situation-aware affordances may look like for the example of proactive furniture.

18.5.2 Specific Solution

Our approach is to show and evaluate how the notion of affordances can be extended to situation-aware affordances including dynamic cues, so that workflows, and relations among objects can be presented. Due to the physical nature of the assembly, the symmetry of boards, and the interchangeable activities, several sequences of assembly steps are possible. However, steps depend on each other, such that previous steps constrain the assembly of the parts in consecutive steps. Thus, the role of parts changes during the assembly process and have to be visualized to the user.

In Sect. 18.4 we showed how the states of assembly could be sensed using integrated sensors. Knowing the state of the assembly, instructions can be given to the user at any time. Displaying instruction on a computer screen does not overcome the disadvantage of conventional paper instructions. By looking at the instructions the user is distracted from his original task of assembling the furniture. The flow of action is disturbed. Augmented reality presents one way of bridging the gap between the instructions and the real world. It has been used to integrate information into a user's physical environment often (Tang et al. 2003; Zauner et al. 2003). However, AR is cumbersome and, typically, computationally expensive. Audible instructions offer a cheaper way of immersion but have to tackle the problem of addressing the appropriate parts by a vocabulary the user is familiar with or has to learn in advance.

Our vision and aim is to integrate instructions directly into objects. We study how affordances of physical objects can be exploited and enhanced by dynamic cues resulting in situation-aware affordances. In particular, we evaluate the effectiveness of LEDs attached to objects as a way of extending static affordances. We use LEDs attached to the furniture parts in the furniture assembly task to guide users through the assembly process.

For the example application of proactive furniture, a set of video based mock-ups was designed. Figure 18.5a shows a screenshot from one of the videos produced. With these we investigated the feasibility of visual markers on objects for enhancing affordances. After showing the different videos to several people at multiple occasions, we defined a set of visual guidance principals we found appropriate. To guide the user through the furniture assembly we identified the following five types of feedback:

1. Direction of attention
2. Positive feedback for right action
3. Negative feedback for wrong action
4. Fine grain direction
5. Notification of finished task

Users unwrap the furniture package and their attention gets directed immediately (feedback type 1) to the parts they are supposed to start with. User's actions, such as turning and moving boards are sensed and blinking green light patterns indicate which edges have to be connected in which manner.

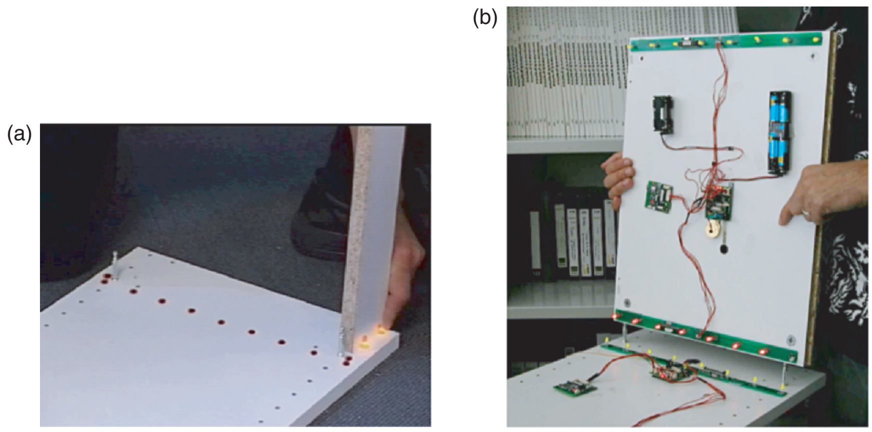

Fig. 18.5. (a) A short video showing the instructions in use was produced. Different colors and flashing sequences were tried out. The video ws showed to a side public to gain experience. The figure shows red flashing lights, while someone is trying to assemble the parts in the wrong way. **(b)** A person assembling two boards with the LED instructions. The short board shows that its orientation is wrong by flashing the lower strip of LEDs red

If boards are aligned in the proper way, a synchronized green light pattern on both edges indicates a well-performed action (feedback type 2). If the user takes a wrong action, a red light pattern appears representing a mistake (feedback type 3, see Fig. 18.5b). Additionally, a green flash pattern shows the alternative. After boards have been aligned together in the right way, individual green lights direct user's attention to the holes where the screws have to be inserted and tightened (feedback type 4).

Coding information with colors has to be done with care as different cultures map meanings to colors in different ways. Even so, Helander (1987) points out that red, yellow, and green should be reserved for "Danger", "Caution", and "Safe", respectively. In accordance with these guidelines we use red to signal an error or wrong position in assembly and green for a correct position. Compared to Tarasewich et al. (2003) we do not evaluate the information rate of LEDs as such, but use them to display instructions. In contrast to tangible bits by Ishii and Ullmer (1997) or the work presented by McGee (2002), we introduce and evaluate the concept of situation-aware affordances to visualize usage of objects rather than using objects to facilitate new forms of human–computer interaction.

18.5.3 Summary of a User Study

To evaluate the use of LED-based instructions we used our interactive prototype. To present instructions to the user we have developed a custom layout board carrying eight dual green/red LEDs; see Fig. 18.5b). These LED strips are connected to the described sensing hardware and are attached to the ends of all three boards to give instructions to the user. For more details on the used hardware, see also Beigl et al. (2003), Holmquist et al. (2003), and Michahelles et al. (2003). Besides only presenting information using the LEDs we have the possibility to provide visual and auditory instructions on a laptop computer.

The study was carried out with 20 participants with different backgrounds. Fourteen of the participants are male, six are female. The average age of the participants is 26.16 years with a standard deviation of 1.64 years. Five of the participants had computer science backgrounds, four were from engineering disciplines and nine from other fields. The participants had different levels of experience with flat-pack furniture assembly. All participants reported normal or corrected-to normal vision. For more details of this study please refer to Michahelles et al. (2003).

The overall goal was to compare the usability and effectiveness of classic paper instructions with our situation-aware affordance approach. To this extent, the user study was conducted in two phases. In the first phase the assembly time between an assembly conducted with classic instructions and one with LED-based instructions was compared. In the second phase of the experiment, participants were encouraged to perform the setup again three times using instructions presented in different modalities. The first modality employed only the LEDs to display the situation-aware affordances. The second modality displayed interactive instructions on a computer screen. The information for the instructions was based on the same sensor setup as with the LEDs. The third modality extended the LEDs with auditory spoken instructions. Overall, the user study presented by Michahelles et al. (2003) revealed that there is a measurable time gain when using LED-based instructions. Beyond that, we saw how errors during assembly can be reduced using instructions in the right place. Designing the sensors and instructions for this purpose, it may even be possible to totally prohibit errors during assembly. For applications beyond furniture assembly, such as airplane or power plant maintenance, this is a critical issue.

Besides these performance-related gains we found other problems solved through the LED-based instructions. The questionnaire showed that determining *which part fits where* is one of the main problems using today's instructions. Interestingly, 75% of the participants found that the LED-based instructions help with exactly this problem. Because the LEDs light up on both boards that need to be joined, finding out what goes where becomes a straightforward task. The instructions do not need to be mapped to the objects anymore, as they are simply integrated into them.

The comparison of instructions presented in different modalities led to further insights. Participants stated that instructions presented on the computer screen helped for orientating the boards correctly, but the screwing direction remained unclear. The various participants received presenting instructions with audio in the form of spoken words differently. About half of the participants found spoken instructions useful. About a quarter found them disturbing. The general opinion is that spoken instructions have to appear at exactly the right time, in order not to disturb too much. In the questionnaire the participants mentioned that they could be useful for instructions that cannot be presented visually. Getting the screws was mentioned as an example.

18.6 Conclusions and Discussion

In this chapter, we have argued that perception and context awareness have great potential to change the way we interact with computers in general and in the context of Ambient Intelligence in particular. The first part of the chapter gives an overview of various sensing and perception opportunities classified by the physical placement of the sensors and by the logical view of the information sensed. The second part of the chapter then describes in more detail an example from our own research for proactive instructions for the running example of a flat pack furniture assembly. For this example we described the modeling, recognition, as well as a proposed feedback mechanism.

We believe that the example of proactive furniture together with the concept of situation-aware affordances can be generalized to several other applications. A wide range of assembly and maintenance tasks could benefit from embedded sensors and proactive instructions. Sensors could be used to monitor the assembly of aircraft or the installation of roof racks on cars. Integrated into machine parts sensors could then be used to continuously monitor system performance. Proactive instructions can offer the user freedom in his choice of actions, while still guaranteeing a correct assembly. Generally, security-critical applications could benefit from information presented at the right time and at the right place. Beyond today's applications we believe that situation-aware affordances have the potential to let the user explore the functionality in future interactive environments in a more intuitive way. Smart homes and office environments will need simple and effective ways of letting the user know how they can be controlled.

As stated above, modeling and recognizing context information is often seen as one of the most important ingredients of Ambient Intelligence. However, the number of examples where context information really is used is rather limited and restricted, for example, to location information, which can be sensed with a predictable or measurable accuracy. The main reasons for this may be characterized by the fact that context information is often too uncertain, ambiguous, and cannot be extracted reliably. One might argue that today's technology is not good enough yet to overcome these problems but as I will argue in the following there are fundamental and inherent problems which are either very difficult if at all solvable. In my opinion, there are at least the following five fundamental issues for context-aware systems:

– *Unobservable information.* One of the most fundamental problems is that not all relevant information is measurable or observable. For example, a human's mental state is not observable or the personal interests and objectives are difficult to reliable estimate for a computer or even for another human observer.
– *Missing information.* Even for observable information in most circumstances a context-aware system will not have access to all relevant information such as complete history information. Even when a particular piece of information is in principle observable this information might not be available since it may not be stored or measured.
– *Unpredictable behavior.* Humans are notoriously unpredictable and change objectives, goals, and motivations often. This poses a great challenge for context-aware systems.
– *Ambiguous situations.* Clearly, many situations do not have a single interpretation but do have multiple interpretations. Those do not necessarily have to contradict each other but they leave enough room for drawing different conclusions.
– *Context is changing constantly.* An interesting but often overlooked issue is that context is changing constantly with every interaction or communication we might have with a particular environment or device. This is probably most obvious in the case of human-to-human communication where every single discussion or communication we have with a person changes our knowledge and understanding of the particular topic as well as of that person so that any subsequent discussion is influenced by that

change in context. However, I do not know of any context- aware system today that does take this into account effectively.

The above list of fundamental issues suggests that context-aware systems should be designed and evaluated much more carefully than it is done today. While these problems really are fundamental and important they are seldom raised and discussed with respect to context-aware systems.

In the following I would like to briefly discuss which and how fundamental challenges are addressed in the described example of proactive instructions for furniture assembly. The issue of unobservable and missing information was alleviated by the fact that we explicitly reduced our modeling and recognition task to the furniture and the actions performed with the various objects involved. This consideration of modeling only what can be modeled reliably clearly helps in many circumstances. Nevertheless, this separation of what can and cannot be modeled is often not done well. In order to reduce the ambiguity of situations to a minimum we basically used a set of sensors that enabled the recognition of the different actions and states with a close to perfect reliability. The unpredictability of human behavior was not an important issue here since we did not aim and need to predict human behavior. Rather than to consider the fact that context is changing constantly over time we allowed the system to adapt its behavior depending on the actions performed by the user. While this does not really address the issue of constantly changing context directly this appears to be sufficient for simple applications such as the one described here.

In summary, we can state that context-aware systems using perception techniques certainly have the potential to change the way we interact with computing in general and with ambient intelligent computing in particular. A first challenge is to make sensing more robust and reliable, for example, by fusion of multiple sensor modalities. A second challenge we have pointed to is the fact that there exist various fundamental challenges inherent in the use of sensing information and context information. While the first issue is well known the second issue seems to be largely underrepresented in our community and we hope that future research will enable us to deal better with these fundamental challenges.