Migrating Characters Performing Explicit Physical Object References

Michael Kruppa
International Post-Graduate College Language Technologies and Cognitive Systems
Saarland University, Saarbrücken, Germany
and DFKI GmbH, Saarbrücken, Germany
mkruppa@cs.uni-sb.de

Abstract

The main goal of the work presented in this paper is to determine an optimal strategy for virtual characters performing judicious combinations of speech, gesture and motion in order to disambiguate references to objects in the physical environment. The work is located in the research area of mobile computing and deals with the combination of mobile and stationary devices.

1 Introduction - The Migrating Character Concept

In the research areas of mobile-, ubiquitous- and pervasive computing, computer technology becomes merged with the physical world. The main benefit of this development is, among the fact that the technology becomes accessible almost everywhere, the possibility to build applications which incorporate the physical environment around the user and hence offer location sensitive services. The basis for these applications is formed by a combination of sensory data and knowledge about the physical world (reactivity). In addition, efficient methods allowing for explicit references to physical objects will be a key element within reasonable solutions for mobile applications. Since virtual characters have proven to successfully disambiguate references in virtual 3D worlds (see [2] and [3]), these characters seem to promise similar results when performing references in the physical world. The Migrating Character Concept described in this paper allows virtual characters to dislocate themselves in physical space (mobility). Furthermore, the Migrating Characters are capable of performing many different types of references, depending on the available technology. A character may for example exist on a mobile device, referring to physical objects by performing gestures on photos or abstract object representations on the screen, or it may also appear on a wall (using a projector) or on a stationary screen right next to the referred object. Based on a user model representing both the user's actual context and preferences and an ontology representing the world knowledge, a set of rules determine an optimal referencing solution in an arbitrary situation (adaptivity). The Migrating Character Concept is based on three major elements: mobility, reactivity and adaptivity. We will describe each of these in the following subsections.

1.1 Character locomotion - Mobility

Character locomotion is the key element of the Migrating Character concept. It allows virtual characters to move throughout the physical world and also to assist users by means of deictic gestures. I classify character locomotion into two categories: active and passive. The active locomotion category subsumes all methods allowing the character to dislocate itself, regardless of the users movements (e.g. a virtual character "jumping" from one device to another or dislocating itself by means of a steerable projector), whereas passive character locomotion depends on the user's movements (e.g. a character on a mobile device, carried by the user). Both categories yield different advantages and implications:

- Using active locomotion, the character is capable of positioning itself freely in the environment. Active locomotion can be the result of an explicit user request or an action performed by the character itself. In the first situation the user might have decided to direct the character to another location, e.g. from a small mobile screen, where only the character's head could be displayed (as in [1]) to a large stationary screen to support a full embodiment of the character. Locomotion in this situation can be used to guide users through an environment. Of course it has to be ensured that the character does not loose contact to the user, which requires some kind of user tracking, either in a global or local coordinate system (e.g. distance between the user and the character).
- Using passive locomotion, the character is sure to be close to the user, however it depends on the user in order to reach a certain position.
- Depending on the chosen locomotion method, the character uses either the same (passive locomotion) or a different (active locomotion) frame of reference of the user and must change its gestures/utterances accordingly. The decision whether active or passive locomotion is possible, depends also on the actual context. A character can only migrate if appropriate screen space is available in the proximity of the user.

These observations demand different character behaviors depending on the actual/available locomotion method. We assume that the character is always driven by a certain objective. In case it is necessary to move to another location in order to fulfill a specific goal, the character could either move to the new location actively (in this case it should however ensure that the user is following) or it could try to convince the user to move to the new location and hence move the character passively. In either way, the character needs to be aware of the users movements/actions in order to react appropriately.

1.2 Physical Context - Reactivity

The first step towards performing a reference to a physical object is sensing the object's presence close to the user. Secondly, information about the object itself is necessary (e.g. *What is it?; How big is it?; Which objects are next to it?*). Depending on the position, size, proximity and similarity of physical objects, different strategies need to be chosen in order to disambiguate references. The objects could be equipped with active senders, emitting all the necessary information in a narrow range. However, it is also possible to determine the users position and orientation, and store the information on physical objects in a database. Organizing this "world knowledge" in an hierarchical structure like an ontology will allow a system to determine relative position information (e.g. *User is in room x; User is close to object y; User can see object z*). Depending on the position, size, proximity and similarity of physical objects, different strategies need to be chosen in order to disambiguate references. While the user position and orientation may limit the choice of appropriate devices to be used by the character due to physical restrictions like visibility, it may also indicate that a user dislocation is necessary prior to performing an object reference. References to physical objects performed by the character are based on the frame of reference of the user and the (possibly different) frame of reference of the character and the location of the physical objects. The characters gestures and spoken utterances, along with a possible character dislocation, are chosen accordingly.

In addition, different types of environments (i.e. a public or a private situation) may impose further restrictions on the choice of appropriate devices for a reference task. However, since users may have very different opinions on what is adequate in a specific situation, the character should also take into account the personal preferences of the user, when deciding on a specific referencing strategy.

1.3 Personal Context - Adaptivity

When deciding on a specific reference strategy, the interaction history between the character and the user is of utmost importance. Keeping track of former references will allow the character to refer to previously mentioned objects more easily (e.g. *The red box we just saw on the other side of the room*). In addition, personal preferences may also influence the reference strategy decision. For example, a user might prefer to reduce the number of references, where the character utilizes a public audio system, to a minimum in public situations.

1.4 Finding a reference strategy

To determine an appropriate reference strategy, a number of steps need to be taken, depicted in figure 1. After determining an object (reference target),needing to be referenced for clarification, the reference planning mechanism starts by locating the user and the physical object to be referenced. Based on a "world knowledge ontology", the next step is to identify possible ambiguities caused by similar objects close to each other. Based on these possible reference ambiguities, a list of reference methods is calculated, ordered by the presumed usefulness in resolving these ambiguities (e.g. sometimes a spoken reference will be sufficient, sometimes a combination of character movement, gesture and speech may be necessary and in some cases it may even be inevitable to ask the user to move to another physical location prior to performing a reference to a specific object). In a next step, only those reference methods remain on the list, which are technically feasible at the location where the reference occurs. Since the main goal is to disambiguate references, the user preferences may only be applied if several reference strategies have been identified, guaranteeing for an explicit physical object reference.

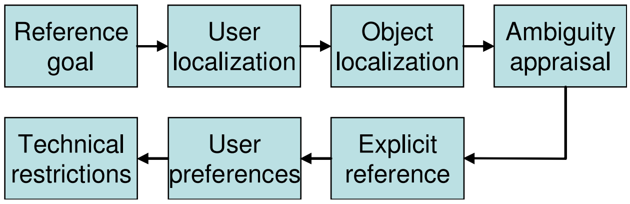


Figure 1: The sequence of steps determining the best reference strategy

2 A prototypical implementation example

Within the scope of two prototype implementations, several different methods for physical object references by means of virtual characters have been realized and evaluated. While in the first project (PEACH, see [4]) a virtual character is used on a Personal Digital Assistant (PDA) to guide users through the physical space of a museum, the second project (VRI, see [5]) implemented a virtual character capable of moving along arbitrary walls by utilizing a steerable projector and a spatial audio system. In the PEACH prototype, the character performs references to physical objects by a combination of gestures and speech related to photos/videos of physical objects shown on the screen of the PDA (see figure 2 A). In addition, the character may "jump" from the PDA onto stationary screens in order to refer to virtual objects shown on those screens or to physical objects located close to the stationary screen (see figure 2B). In contrast, the character in the VRI project is capable of positioning itself right next to a physical object and is hence capable of giving very precise references (see figure 2C and 2D).

In order to allow the Migrating Characters to perform explicit references to physical- and virtual objects the following mechanism was implemented:

Whenever a physical object reference is necessary, a rule-based physical reference system is instantiated. The data fed into the system is taken from an online ontology and transformed into facts which are then asserted and evaluated by the rules defined in the rule-based system. These facts describe the room in which the user is located at the moment, the objects located in that room, the reference target and the actual situ-

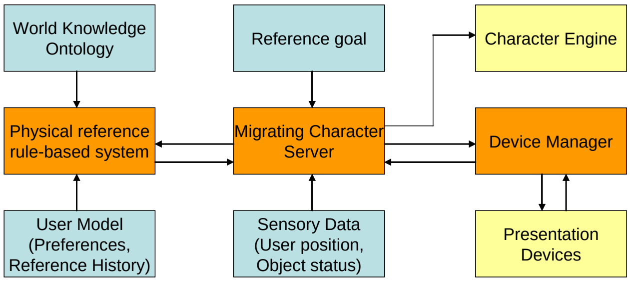
Figure 2: Examples of Migrating Characters performing physical object references

ative context of the user. Whenever instantiated, the rule-based system will determine the next step which is necessary in order to reach the final goal of a unique physical object reference. Based on the result produced by the rule-based system, the Migrating Character will perform the according actions and then (in case this wasn't the final step of the reference task), the rule-based system will be instantiated again with the updated world model and user context.

Starting from the initial reference target, the first three rules within the rule-based system determine whether a user dislocation is necessary (i.e. user must move to another room, turn towards another wall or move closer to a wall if the target object is too small from the user position). If any one of these three rules is activated, the result determined by the rule-based system will not be a reference instruction for the Migrating Character but instead it will be an instruction for a necessary physical user context change. Based on this instruction, the user is asked to dislocate herself accordingly by the Migrating Character. Whenever the user context has changed, the rule-based system is restarted. Once, none of the first three rules are activated due to the user's physical context, the result determined by the rule-based system will be a reference instruction for the Migrating Character. In case the object to be referenced is not close to a similar object (this is calculated based on the size and type of the target object and surrounding objects), a simple spoken reference will be sufficient. Otherwise, based on the user's focus history (i.e. last two objects which have been referenced), the availability of a picture of the target object, the availability of the necessary hardware and physical space for a character dislocation right next to the object, and the users preferences, different strategies are chosen to disambiguate the physical object reference. If all strategies fail, the worst case is a possibly ambiguous spoken reference by the Migrating Character.

The rule-based physical reference system has been realized in JESS[1] and was integrated into a java server which handles communication with the knowledge base and sensor services. Whenever necessary, the reference system is invoked by the java server. The reference system itself consists of $\sim$ 20 rules and functions.

The interplay of the single components of the Migrating Character Referencing System is depicted in the following graph:



3 Conclusions

The Migrating Character technology described in this paper offers a new approach to the problem of disambiguating references to physical objects in mobile applications. By utilizing virtual characters capable of performing a combination of utterances, gestures and character locomotion in physical space, possibly ambiguous references to physical objects can be supported by the virtual character and hence become explicit. An empirical study testing the effectiveness and reliability of the described object referencing technique is to be conducted in the near future.

References

[1] Malaka, R., Häußler, J., Aras, H.: Smartkom mobile - intelligent ubiquitos user interaction. In: Proceedings of the international conference on Intelligent User Interfaces. (2004)

[2] Lester, J.C., Towns, S.G., Callaway, C.B., Voerman, J.L., FitzGerald, P.J.: Deictic and emotive communication in animated pedagogical agents. Embodied conversational agents (2000) 123–154

[3] Towns, S.G., Vorman, J.L., Callaway, C.B., Lester, J.C.: Coherent gestures, locomotion, and speech in life-like pedagogical agents. In: Proceedings of the international conference on Intelligent User Interfaces (IUI). (1997)

[4] Kruppa, M., Krüger, A., Rocchi, C., Stock, O., Zancanaro, M.: Seamless personalized TV-like presentations on mobile and stationary devices in a museum. In: Proceedings of the International Conference on Hypermedia and Interactivity in Museums (ICHIM). (2003)

[5] Kruppa, M., Spassova, M., Schmitz, M.: The Virtual Room Inhabitant. In: Proceedings of the 2nd IUI workshop on Multi-User and Ubiquitous User Interfaces (MU3I). (2005)

[6] Heckmann, D.: Integrating privacy aspects into ubiquitous computing: A basic user interface for personalization. In: AIMS03 Workshop, in conjunction with UbiComp2003. (2003)

[1] A rule engine for the java platform - http://herzberg.ca.sandia.gov/jess