Framing the Next-Generation 'Desktop' using Proximity and Human Perception

Thomas Pederson[1,2]
[1]Dept. of Computing Science
Umeå University
SE90187 Umeå
Sweden
top@cs.umu.se
[2]Dip. di Informatica
Università degli Studi di Bari
70125 Bari
Italy
pederson@di.uniba.it

Antonio Piccinno
Dip. di Informatica
Università degli Studi di Bari
70125 Bari
Italy
piccinno@di.uniba.it

Dipak Surie
Dept. of Computing Science
Umeå University
SE90187 Umeå
Sweden
dipak@cs.umu.se

Carmelo Ardito
Dip. di Informatica
Università degli Studi di Bari
70125 Bari
Italy
ardito@di.uniba.it

Nicholas Caporusso
Dip. di Informatica
Università degli Studi di Bari
70125 Bari
Italy
caporusso@di.uniba.it

Lars-Erik Janlert
Dept. of Computing Science
Umeå University
SE90187 Umeå
Sweden
lej@cs.umu.se

CHI 2007, April 28 – May 3, 2007, San Jose, USA
ACM 1-xxxxxxxxxxxxxxxxxx.

Introduction

Personal computing, and therefore Human-Computer Interaction (HCI), is becoming a seamlessly integrated part of everyday activity down to the point where "computing" is inseparable from "activity". A modelling problem occurs in these emerging mobile and ubiquitous computing situations because it is hard to determine the spatial and operational limits of an ongoing activity, for the human performing the activity, for the computer system monitoring and/or supporting it, as well as for the modeller observing it. Also, it is an open issue how best to model the causal relations between physical (real world) and virtual (digital world) phenomena that these "intelligent environments" can be programmed to maintain, whether defined by software engineers or the end-users themselves. We propose a modeling framework that adresses the above mentioned issues and present our initial attempts to create a User Interface Description Language (UIDL) based on the framework.

Egocentric Interaction

The Egocentric Interaction perspective on HCI has its beginnings in observing the interaction possibilities and limitations present around a specific human body and mind immersed in a world full of physical (real-world) and virtual (computational) objects of potential

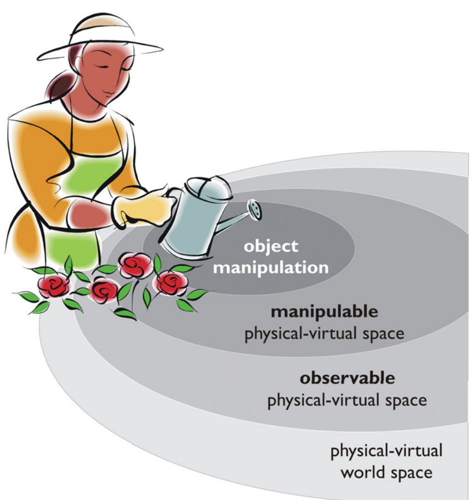

figure 1: A situative space model [2].

interest. Input and output devices are, just like the computer hardware itself, completely ignored in this conceptual framework similar to how the chair used by a typical desktop PC user is ignored in classical HCI models. This deliberate disregard for interaction infrastructure allows the modeling of physical and virtual objects as if they co-existed in the same Euclidean space. We would argue, without proof, that this might be very close to how domain experts see their current environment in the situation when no unexpected breakdowns (technological or other) occur [4]. While research in improving and evolving classical input and output infrastructure to fit the new computing applications surely has its place, we believe that greater advance can be made by rethinking the role of personal computing itself.

The "Desktop" for Next-Generation User Interfaces?
The situative space model (**figure 1**) is based on the physical-virtual design perspective briefly outlined above, meaning that physical and virtual domain objects are treated as being located in the same space. The model is for the emerging Egocentric Interaction paradigm what the virtual desktop is for the PC/WIMP interaction paradigm: more or less everything of interest to a specific human actor is assumed to, and supposed to, happen here.

An Example of Applying the Model to an Actor's Situation.
If a glass of juice is in the right hand of a specific human actor, and an email displayed on a cellular phone in the left hand (ready to view), both objects would be considered to reside in the object manipulation space in **figure 1**. A paper newspaper on the table just in front, and the keys in the same persons pocket would be modelled as inside the manipulable space but outside the object manipulation space. A painting on the opposite side of the table (but not the one behind the actor's back) would be in the observable space. Finally, all in principle perceivable objects in the physical-virtual world which at least for the moment do not happen to be perceivable by the specific human actor are regarded as situated in the world space, encompassing the spaces mentioned earlier.

In the physical world, the outer border of the manipulable space can be approximated and described in Euclidean terms: manipulable things are typically closer than things that are observable but not amenable to manipulation. This spatial relationship is reflected in **figure 1**. Determining a corresponding border in the virtual world is somewhat more complex and depends on situative access to input and output devices. Due to the nature of the application area towards which our current system development efforts are targeted, we have chosen to temporarily suspend the work on investigating how object manipulation and navigation should be best modelled in virtual environments (e.g. of the WIMP kind) to fit the situative space model. However, experiences from a first attempt [1] has convinced us that it should be possible.

Using the Model for Guiding Ubiquitous Computing Application Design.
We consider the borders of the observable space to define the set of objects that can possibly be part of a ubiquitous computing "application" at any given time for a specific human actor. If a computing system displays information outside the observable space, it

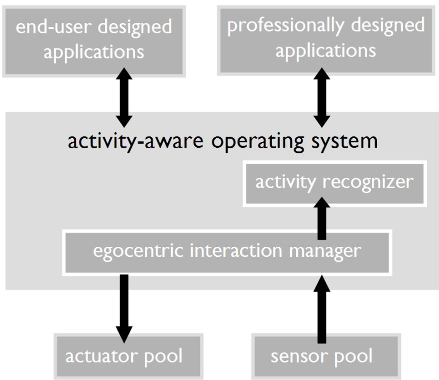

figure 2: An activity-aware wearable computing architecture including an Egocentric Interaction Manager [2].

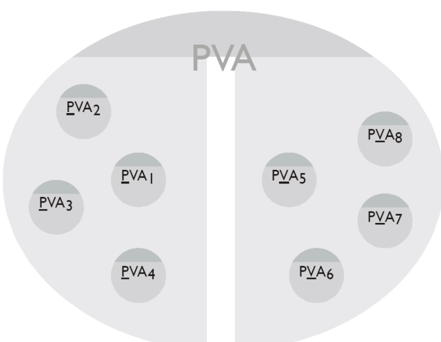

figure 3: A Physical-Virtual Artefact consisting of four physical manifestations shown to the left (e.g. four printouts of this document) and four virtual manifestations shown to the right (e.g. digital versions of this document in different formats and/or residing at different places in Cyberspace) [1].

will not be noticed. Likewise, if access to a desired virtual object is provided through an input device currently outside of the manipulable space, the human actor is forced to change physical location. As it happens, this view aligns well with the WIMP/direct manipulation paradigm for virtual-world interaction where successful application design as well as use very much depends on keeping the right objects "on screen" at the right time [3].

The Egocentric Interaction Manager
Within the context of developing an activity-aware general-purpose wearable computing architecture (**figure 2**), we have started to develop a computing component which will have as a task to keep necessary interaction resources within the manipulable space (in the case of input) and within the observable space (in the case of visual output). Because of limited actuation possibilities in the physical world, the component will at times rely on assistance from the human actor, such as to shift position to face a visual display, or to pick up a display-equipped device from the pocket, in the case when important visual information has to be communicated to the actor.

Physical-Virtual Artefacts
In the Ubiquitous Computing literature, it is often predicted that our everyday environments will become more "intelligent" in the sense that an increasing amount of actions belonging to specific higher-level personal human activities will be automated by computer systems that more or less frequently and autonomously intervene to the benefit of the human performing the activity. For the purpose of modeling pre-programmed or end-user programmed causal relations between physical and virtual objects (such as the effect of moving an arrow on a virtual desktop as a consequence of moving an ergonomically shaped object on a physical desktop, i.e. the essential functionality of a computer mouse), we have introduced the concept of Physical-Virtual Artefact (PVA):

A physical-virtual artefact is an abstract artefact that (1) is manifested in both the physical and the virtual environment, where (2) these manifestations to a large extent utilise the unique affordances and constraints (Norman, 1988) that the two different environments facilitate, and finally (3) where one manifestation of a specific physical-virtual artefact is easily identified if a corresponding manifestation in the other environment is known. [1]

This definition may be relaxed a bit to include also artefacts that are manifested in just one of the two worlds (e.g. a hammer in the physical world, or a web page in the virtual world), as well as artefacts that have more than one manifestation in one world (as illustrated in **figure 3**). The result, we believe, is a generalisation of what is the notion of an object of interest for a specific human actor (sometimes also referred to as "domain object") that lends itself well to modelling the majority of objects that appear in the situative space model during the course of any specific activity.

Towards a User Interface Description Language

In the last years, many efforts have been made to develop XML languages expressing implementations of user interfaces (UIs). While some of these have a W3C origin (e.g. HTML; XHTML; SVG, etc.), others are proposed by various research groups to better handle

```
<WorldSpace>
  ……
  <ObservableSpace>
    <PVArtefact ID = "B7934"
      name = "Java for Dummies"
      type = "book">
      <PhysicalManifestation>
        <material> paper </material>
        <coverType> hard </coverType>
        <manipulationType> can be opened
          </manipulationType>
        ………
      </PhysicalManifestation>
      <VirtualManifestation>
        <fileType> PDF </fileType>
        <fileName>
          Apogeo Java Book.pdf
        </fileName>
        <address> C:\Antonio\ </address>
        ………
      </VirtualManifestation>
    </PVArtefact>
    ……
    <ManipulableSpace>
      ……
      <ObjectManipulationSpace>
        <PVArtefact ID = "A1254"
          name = "My Office Monitor"
          type = "visual display">
          <PhysicalManifestation>
            <brand> CXT </brand>
            <displayType> CRT </displaytype>
            <maxResolution>
              1024x768
            </maxResolution>
            ………
          </PhysicalManifestation>
        </PVArtefact>
        <PVArtefact ID = "W3456"
          name = "Google Maps"
          type = "web page">
          <VirtualManifestation>
            <URL>
              http://maps.google.it/
            </URL>
            ………
          </VirtualManifestation>
        </PVArtefact>
        ……
      </ObjectManipulationSpace>
    </ManipulableSpace>
  <ObservalSpace>
</WorldSpace>
```

figure 4: A UIDL example transcript describing a situation where a book's both physical and virtual manifestations are present in the observable space, while a computer display and a web page is inside the manipulable space.

situations when the target devices are very heterogeneous, consisting of a variety of devices such as desktop computers, PDAs, and cellular phones. Nevertheless, most of these UIDLs are able to describe WIMP-like UIs only.

By combining the concept of PVA with the situative space model, and creating a suitable formal language, we should be able to describe the physical and virtual resources actually used within the course of an activity (by observing what artefact manifestations that become manipulated and observed)[1], as well as to specify the requirements for new "user interfaces" intended to support envisioned or existing physical-virtual activities, i.e. activities requiring frequent switching between actions in the physical and in the virtual worlds. A good representation of available interaction resources would also be the basis for the operation of the previously mentioned Egocentric Interaction Manager. A first sketch of such a markup language representation is shown in **figure 4**.

We are currently working on the definition of the properties for each PVArtefact and on the XML Schema containing the rules to which each artefact type has to comply to. Note that there is nothing preventing a PVArtefact from have several physical and/or virtual manifestations (as pictured in **figure 3**) in this representation.

[1] We have in fact received promising results by using the situative space model for the purpose of activity recognition in a kitchen environment simulated in Virtual Reality [5].

Future Work: Extending the Model with Sound and Gestures

Up until now, only the visual presence of objects have been considered in the manipulable and observable spaces, while object manipulation has been limited to object grabbed/not grabbed. Apart from developing the UIDL language, we are interested in increasing the modeling power in the object manipulation "space" by adding some notion of gestures with (and without?) grabbed objects. We are also interested in increasing the number of interaction modalities covered by the model by starting to consider sound as a complement to vision for determining the borders of spaces. Initial theoretical investigations show that sound modelling could be done by adding a complementary observable (or rather, perceivable) space and an additional sound-related manipulable space to the model.

References

[1] Pederson, T. *From Conceptual Links to Causal Relations — Physical-Virtual Artefacts in Mixed-Reality Space*. PhD thesis, Umeå university, Sweden, report UMINF-03.14, ISSN 0348-0542, ISBN 91-7305-556-5. (2003).

[2] Pederson, T. & Surie, D. Towards an activity-aware wearable computing platform based on an egocentric interaction model. In *Proc. UCS 2007*, LNCS, Springer (2007).

[3] Shneiderman, B. The future of interactive systems and the emergence of direct manipulation. *Behaviour and Information Technology*, 1, 237-256 (1982)

[4] Suchman, L. A. *Plans and Situated Actions*. Cambridge University Press (1987).

[5] Surie, D., Pederson, T., Lagriffoul, F., Janlert, L., Sjölie, D.: Activity Recognition using an Egocentric Perspective of Everyday Objects. In *Proc. UIC 2007*, Springer (2007).