

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

HephaisTK : A Toolkit for Rapid Prototyping of Multimodal Interfaces

Dumas, Bruno; Lalanne, Denis; Ingold, Rolf

Published in:

Proceedings of the 2009 international conference on Multimodal interfaces (ICMI-MLMI '09)

DOI:

[10.1145/1647314.1647360](https://doi.org/10.1145/1647314.1647360)

Publication date:

2009

Document Version

Peer reviewed version

[Link to publication](#)

Citation for published version (HARVARD):

Dumas, B, Lalanne, D & Ingold, R 2009, HephaisTK : A Toolkit for Rapid Prototyping of Multimodal Interfaces. in *Proceedings of the 2009 international conference on Multimodal interfaces (ICMI-MLMI '09)*. ACM Press, pp. 231-232. <https://doi.org/10.1145/1647314.1647360>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

HephaistK: A Toolkit for Rapid Prototyping of Multimodal Interfaces

Bruno Dumas
University of Fribourg
Bd de Pérolles 90
1700 Fribourg, Switzerland
0041/26.300.92.96
bruno.dumas@unifr.ch

Denis Lalanne
University of Fribourg
Bd de Pérolles 90
1700 Fribourg, Switzerland
0041/26.300.84.72
denis.lalanne@unifr.ch

Rolf Ingold
University of Fribourg
Bd de Pérolles 90
1700 Fribourg, Switzerland
0041/26.300.84.66
rolf.ingold@unifr.ch

ABSTRACT

This article introduces HephaistK, a toolkit for rapid prototyping of multimodal interfaces. After briefly discussing the state of the art, the architecture traits of the toolkit are displayed, along with the major features of HephaistK: agent-based architecture, ability to plug in easily new input recognizers, fusion engine and configuration by means of a SMUIML XML file. Finally, applications created with the HephaistK toolkit are discussed.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces – Input devices and strategies, Interaction styles, Prototyping.

General Terms

Human Factors, Languages.

Keywords

Multimodal toolkit, multimodal interfaces, human-machine interaction.

1. INTRODUCTION

A number of tools for creation of multimodal interfaces have surfaced in the recent years. Bourguet [2] endeavored in the creation of a multimodal toolkit in which multimodal scenarios could be modelled using finite state machines. This multimodal toolkit is composed of two components, a graphical user interface named IMBuilder which interfaces the multimodal framework itself, named MEngine. Multimodal interaction models created with IMBuilder are saved as a XML file. Flippo et al. [5] also worked on the design of a multimodal framework, geared toward direct integration into a multimodal application. The general framework architecture is based on agents, while the fusion technique itself uses frames. Configuration of the fusion is done via an XML file, specifying for each frame a number of slots to be

filled and direct link to actual resolver implementations. Lastly, Bouchet et al. [1] proposed a component-based approach called ICARE thoroughly based on the CARE [3] design space. These components cover elementary tasks, modality-dependent tasks or generic tasks like fusion. Each of those approaches tackled only some of the four challenges when confronted to generic multimodal interfaces creation: modular architecture, possibility to plug in easily new input recognizers, configurable input fusion engine, and tool configuration.

Currently, the most advanced approach is the one of OpenInterface [7]. OpenInterface components are configured via CIDL XML files, and a graphical editor. Finally, OpenInterface uses the CARE design space for fusion.

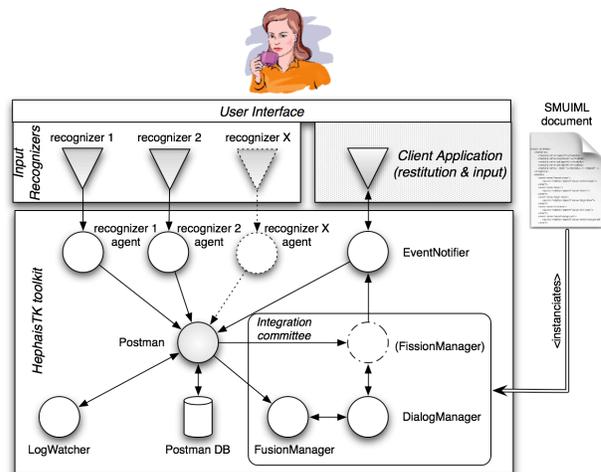


Figure 1: HephaistK toolkit architecture.

2. HEPHAISTK TOOLKIT

In order to provide a tool allowing developers to prototype multimodal interfaces in an easier way than by building them from scratch, a toolkit named HephaistK [4] has been developed. This toolkit has been designed to plug itself in a client application that wishes to receive notifications of multimodal events received from a set of modality recognizers. HephaistK is based on software agents, as depicted in Figure 1. Agents are dispatched to manage individual modality recognizers, receive and encapsulate data from the recognizers, and send them to an individual central agent named the "postman". This postman agent centralizes all data coming from the dispatched recognizers agents in a database,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '04, Month 1–2, 2004, City, State, Country.
Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.

and distributes the data to other interested agents in the toolkit, which can subscribe to be informed of specific types of data. An “integration committee” of three different agents achieves integration of data. A first agent manages fusion of input modalities, helped by a dialog agent; a fission agent encapsulates the fused data and sends it to the client application, which will be notified of the incoming data by means of event listeners, as would be done for a standard input source like the mouse.

HephaistK is an integrated tool, which needs to be configured in order to send satisfactory information to its client application. Thus, a configuration file, describing the human-machine multimodal dialog wished for the client application, needs to be specified when using the toolkit. This configuration script also has other goals, for example the specification of which recognizers need to be used with the toolkit. Hence, the SMUIML language has been developed as a configuration language for the HephaistK toolkit. The SMUIML language describes the different input events, how these input events have to be fused, the messages the toolkit will have to send to the client application, and, at a higher level, the human-machine dialogue.

Finally, a benchmarking tool is used for simulation needs. This benchmarking tool is able to simulate a number of input recognizers, and can send at predefined times input events into the toolkit. This allows testing the behavior of the fusion engine in front of specific use cases, as well as the description in SMUIML.

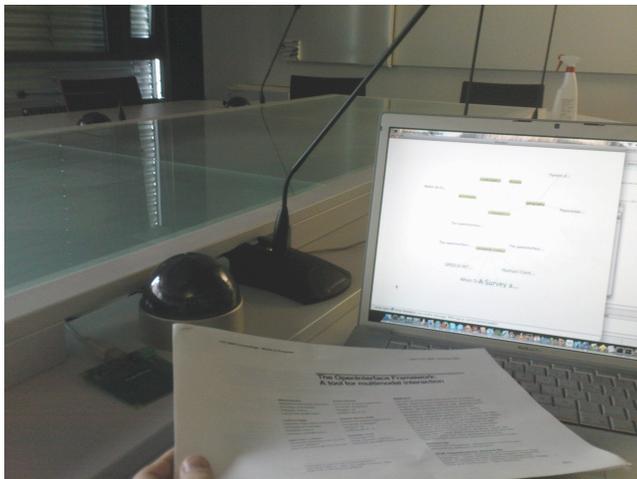


Figure 2. The Docobro multimodal document browser in use.

3. DEMONSTRATION APPLICATIONS

Three demonstration applications have been realized with help of the HephaistK toolkit. First, a multimodal music player was created in order to assess the correct operation of the toolkit. This music player could be used by means of different modalities : mouse, voice, or RFID objects. This multimodal music player application is still provided with the toolkit distribution as a toy example to show developers how HephaistK is used.

Since this first example, a number of different application for Smart Meeting Rooms were created with help of the HephaistK toolkit. Due to its extensible nature, HephaistK can accept inputs from a number of different sources, currently ranging from the standard keyboard and mouse, to Sphinx speech recognition, phidgets physical devices, RFID tag readers, or interactive tables such as ReacTIVision [6] or DiamondTouch. Adding other input

sources can be achieved by extending a set of predefined JAVA classes.

A typical meeting-oriented application created with help of the HephaistK toolkit is a tool for document management in Smart Meeting Rooms. This application, named Docobro (Figure 2), allows a user to visualize a number of documents, sort them by type or thematic, add a previously RFID-tag paper document to the virtual cloud of electronic documents, and open and display the electronic representation of a paper document. Docobro can be manipulated by means of standard keyboard/mouse, speech commands or RFID tagged objects.

The demonstration will feature Docobro, as well as a number of applications created with help of HephaistK, and will also focus on the toolkit itself. In particular, the different tools integrated within the toolkit (see Figure 3) will be shown.

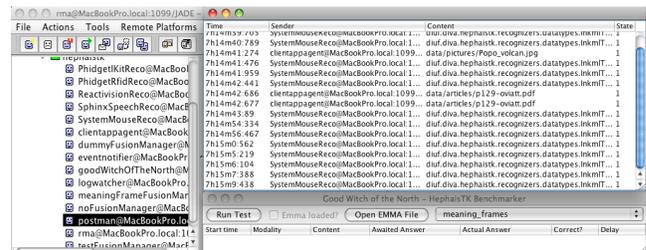


Figure 3. HephaistK logs and controls.

4. REFERENCES

- [1] Bouchet, J., Nigay, L., Ganille, T.: ICARE Software Components for Rapidly Developing Multimodal Interfaces. In: Conference Proceedings of ICMIT2004, State College, Pennsylvania, USA, Oct. 2004, ACM Press, pp. 251--258.
- [2] Bourguet, M. L.: A Toolkit for Creating and Testing Multimodal Interface Designs. In: companion proceedings of UIST'02, Paris, Oct. 2002, pp. 29--30 (2002).
- [3] Coutaz, J., Nigay, L., Salber, D., Blandford, A., May, J. and Young, R. Four Easy Pieces for Assessing the Usability of Multimodal Interaction: The CARE properties. In *Proceedings of the INTERACT'95 conference*, S. A. Arnesen & D. Gilmore Eds., Chapman & Hall Publ., Lillehammer, Norway, June 1995, pp. 115-1-20.
- [4] Dumas, B., Lalanne, D., Guinard, D., Ingold, R., Koenig, R.: Strengths and Weaknesses of Software Architectures for the Rapid Creation of Tangible and Multimodal Interfaces. In: Proceedings of TEI 2008, Bonn (Germany), February 19 - 21 2008 , pp. 47--54 (2008).
- [5] Flippo, F., Krebs, A. Marsic I.: A Framework for Rapid Development of Multimodal Interfaces. In: Proceedings of ICMIT'03, Vancouver, BC, Nov. 5-7, pp. 109--116 (2003).
- [6] Kaltenbrunner, M., Bencina, R. ReacTIVision: A Computer-Vision Framework for Table-Based Tangible Interaction. *Proceedings of TEI '07*. Baton Rouge, Louisiana, 2007.
- [7] Lawson, J-Y., Al-Akkad, A., Vanderdonck, J. and Macq, B. An Open Source Workbench for Prototyping Multimodal Interactions Based on Off-The-Shelf Heterogeneous Components. Proceedings of the First ACM SIGCHI Symposium on Engineering Interactive Computing Systems, ACM Press, USA, July 14--17, 2009.