



Enabling End-User Datawarehouse Mining
Contract No. IST-1999-11993
Deliverable No. D12.3

Description of the HCI for pre-processing chains Deliverable 12.3

Michael May, Detlef Geppert

Fraunhofer AIS
Institut für Autonome intelligente Systeme
Schloss Birlinghoven
D-53754 Sankt Augustin, Germany
{ Michael.May, Detlef.Geppert }@ais.fraunhofer.de
<http://www.ais.fraunhofer.de>

20. December 2003

Abstract

In this document the HCI for pre-processing chains (Chain Editor) and the framework for the whole HCI is described. Other parts of the system are the HCI for concepts and relationships (Concept Editor - see deliverable 12.4), which is integrated in a common environment with the Chain Editor, M4 Interface (deliverable 12.2), which is used by both editors and the M4 compiler (deliverable 12.1) which can be called by the HCI.

The document discusses the installation, configuration, usage of the main framework and the Chain Editor, the current state of the software and further improvements which will be implemented in the future.

Important: Due to the prolongation of the project and the fact, that the work on the software for the Chain Editor is still going on, this document is not final, too. This document is not meant to be a full user guide, but it gives a short overview how to work with the HCI and the Chain Editor. With the further development of the software, a complete user guide will be written.

Contents

Chapter 1	4
Introduction	4
Chapter 2	5
Installation of the HCI	5
Chapter 3	6
Getting started	6
Chapter 4	7
4.1 Using the Chain Editor	7
4.1.1 Cases	7
4.1.2 Chains	7
4.1.3 Steps	7
4.2 Compiler	8
Chapter 5	9
Current state and further improvements	9

Chapter 1

Introduction

The objective for work package 12 is to provide users with a graphical user interface for the specification of an entire pre- processing task. It should enable the user to transform business data into a form that is ready for mining (the learning stage in the knowledge discovery process). The users that are targeted to use this interface are the case designer and the database administrator. The case designer does the main part of the work. He specifies a conceptual model and decides which operators to apply and in what order. It is the task of the database administrator to map the conceptual model to the relational model: A necessary step before executing a case.

The most important components in the architecture that is developed in WP12 are the Chain Editor, Concept Editor, Compiler and the M4 Interface. The Chain Editor and the Concept Editor use the M4 Interface for accessing, creating and manipulating meta data objects.

The Concept Editor provides a HCI for building and representing a conceptual data model and to map it to the relational data model. It is also responsible for validation of the conceptual data model elements and the mapping of these elements to the database level.

The Chain Editor provides a HCI for creating and manipulating pre- processing chains for a case. It supports the case designer to define pre- processing steps and to build an ordered chain of pre- processing steps. It supports the case designer in building valid chains with testing the validity of steps.

Further information about the WP12 HCI requirements can be found in the document by Perot System Netherlands and the Fraunhofer Institut Autonome Intelligente Systeme titled *MiningMart Human Computer Interface, Requirements and specifications* that was presented in March 2002.

This document discusses the installation and usage of the common program environment and the Chain Editor. Furthermore it describes the current state and further developments of both parts.

Chapter 2

Installation of the HCI

The human computer interface can be downloaded as hci.zip. This file should be unpacked in a new directory on a windows machine, as the HCI has not been tested on a different system yet, although in principle it should run on other systems as well. However, the start script is available for Windows only at the moment (file start_hci.bat).

Before the HCI can be used, the start script file start_hci.bat has to be edited. Three lines have to be adjusted to the users own environment:

1. Set the variable mypath to the directory where the HCI is unpacked. Example: set mypath=C:\HCI.
2. Set the variable compilerServer to the machine where the compiler server is running. Example: set compilerServer=kirmes.cs.uni-dortmung.de.
3. Set the JBoss server and port. That is, replace server.gmd.de:1099 with the correct server name and port. Example: set jndi3=java.naming.provider.url=jnp://129.217.30.159:1099

Further a file db.config must be set up. For this the files db.config.example and db.config.template are provided in the HCI directory.

This description supposes that the user already has the JBoss- server and the compiler server running and that he installed the M4 meta data schema in an oracle database. For a closer look how to do this the document *Installing the MiningMart system, Technical Report 12.5* can be downloaded.

Chapter 3

Getting started

To start the HCI, the user has to run the file `start_hci.bat`. After a short moment the main frame of the application can be seen. Five Buttons are enabled, the user can create a new case, open a case from the database to edit or to view only, import a case from the file system or call the help system. This functionality is also provided via the menus `File => New`, `File => Open` and `File => Import`. Open a case is only possible if the user already has installed a case in the database or if he already has worked with the system and created a case. To import a case the user needs to have a file which is exported with a MiningMart system.

After the user has opened or created a case, three windows are shown. The one with the title "Concept Editor" belongs to the Concept Editor, the two others to the Chain Editor. For a new case he first has to describe the conceptual model and build a connection to the business database with the Concept Editor. How to work with the Concept Editor is described in the Deliverable 12.4. by Perot Systems, Netherlands.

One of the other two windows has as title the name of the case, the other has the name of the case (if the node with the case name is selected) or the name of a chain, if the node with the chain name is selected or if a node which is part of the chain is selected.

If a new case is created, the first chain has the same name as the case. How to change a chain name or the name of a node (step) is described in chapter 4.

Chapter 4

4.1 Using the Chain Editor

The methods and actions described in this chapter all belongs to the Chain Editor. The Chain Editor consists of two internal frames, one with the case name as title (tree view) and one with the chain name as title if a chain or a step in a chain is selected (graph view). The use of the menus and tool buttons depends on which window of the chain editor is active. Furthermore for some methods it is possible to use tool buttons and the menus to do the work. The rest of this chapter describes how to build, delete and manipulate the main objects of the chain editor.

4.1.1 Cases

The only way the object “case” can be manipulated is to change the properties of a case. To rename a case the user has to select the top node in the tree view, which stands for the whole case. Then he can either use the edit menu, and than the menu items open or properties or he can click on the tool button “Open Component” or “Edit Properties”. For a case or a chain this menu item opens the same window.

A property window appear where the user can change the name and enter or change a description for the case. After changing some information pressing “Save” will store the new information.

4.1.2 Chains

After opening or creating a case, at least one chain is visible in the tree view. The symbol for a chain is a folder in the tree view. A folder (or a chain) is the only object which can be inserted into a case using the menu “Insert => Chain” or the leftmost button in the button bar at the button of the HCI.

Selecting a chain enables some more menu entries. The user can delete and rename a chain. Renaming a chain is done the same way as for a case. He also can insert steps to a chain. He can either use the menu “Insert => MiningMart Operator => ...” This gives an overview of the existing operators which can be used in a step or the button bar at the bottom of the HCI. Both is described in more detail in the chapter 4.1.3.

After selecting a chain the user get an overview of the chain in the graph view and the frame with the graph view gets the chain name as title. It shows the steps which belong to the chain and the dependencies between the steps. This dependencies has to be inserted explicit with the arrow button from the button bar at the bottom of the HCI, with “Edit => Connections...” or the button “Connections” from the button bar at the top of the HCI. The latter two will open a menu where the connections (dependencies) can be edited, while the first one works in the graph view. First click on the arrow button, then press the left mouse button over one step and drag it to another step in the graph view. If this two steps can be connected, the grey arrow gets black and after releasing the mouse button, this two steps are connected. To do so it is necessary that the graph view is the frame with the focus.

4.1.3 Steps

First the user has to select a chain which he want to edit. Inserting a step into a chain can be done in two ways. If the tree view is the active frame, He can use the menu “Insert => MiningMart Operator => Operator Name”.

One special operator is called generic operator. In the current version this operator can only be

inserted via this menu entry. Selecting this operator he get a list of all m4 operators which are specified in the m4 database. The property window for this step is generated only with the information in the database. This provides an general way to extend the system with new operators. From the HCI point of view only the new operator with its parameters and its constraints has to be specified in the database to integrate a new operator.

After selecting an operator a new step with this operator is inserted in the tree view and the graph view.

The second way to insert a new operator is pressing a operator button from the button bar and than clicking in somewhere in the graph view. This will insert a new step in the graph view at the point where the user clicked. To do so the graph view has to have the focus. The step is inserted in the tree view, too.

Selecting a step will enable the user to delete (“Edit => Delete” or pressing “del” on the keyboard), open (“Edit => Open”, open button or double click on a step symbol in the graph view), connect a step (as described in the previous chapter) and edit properties (“Edit => properties” or property button). Open a step means open a parameter window where all the parameters for the operator of the step can be edited. For many operators individual editors and a single insert buttons exist, for the rest of the operators a generic view is created. Edit properties means renaming and changing description means the same as with chains and cases.

4.2 Compiler

If a case is open, the user is able to compile the whole case. Selecting a step enables him to compile this single step or to compile the selected step every steps behind it. The latter two assume that all steps from which the selected one is dependent are compiled correctly. To start the compiler he can use the menu “Compile” or the three compile buttons.

Chapter 5

Current state and further improvements

In the current state the Chain Editor and the common frame of the HCI supports the main functionality. The user is able to create cases, chains, to work with operators and steps and he can call the compiler. He can export and import cases and can use a help system.

However, there are still some open issues and some features which can be improved. Although the system is tested intensively, some more tests and minor bug fixes may be still necessary. The rest of the document gives an unsorted overview over the work which will be done in the next month:

- **Compiler window:** While the system is compiling steps, the user isn't able to see what's currently happening. In the future the HCI will provide a window where the user can see the compiler messages. For this window he can specify how detailed these messages are.
- **Validity checks:** Some checks are implemented, but this point has to be improved.
- **Description for an object:** In the current version the user is able to enter a description, but it isn't saved in the database.
- **Chains:** In the current version the user can insert different chains and can insert a chain into another, but there are no dependencies between them. In the next version a chain (folder) can be used to substitute a sequence of steps in a chain.
- **Generic Operator:** The parameters in the generic operator view aren't ordered in the current version.
- **Online help:** In the current version only parts of the Concept Editor and the existing operators are described. In the future a help system which guides the user in using the HCI will be created.