Taverna Optimization Plugin user manual

Sonja Holl
Contents

1 Introduction .................................................................................................................. 2
2 Installation ................................................................................................................. 2
3 Basic usage guide ..................................................................................................... 2
4 Known problems and open issues ........................................................................... 5
1 Introduction

This document describes how to install and use the Optimization Plugin for the Taverna Workbench 2.3.

The plugin enables users to optimize workflows via the optimization perspective.

2 Installation

Please make sure that you have installed the Taverna Workbench 2.3 from http://www.taverna.org.uk/.

Start the Taverna Workbench and open the PluginManager via Updates and plugins from the Advanced menu. Use Find new Plugins then Add update site and type as URL: http://unicore-dev.zam.kfa-juelich.de/taverna/plugins/ then install the optimization plugin (may take a few minutes). Then restart Taverna.

Sometimes, Taverna cannot manage to directly download the plugin (it is queued) then please uncheck the "install" check box and recheck it. Afterwards restart Taverna. The plugin will be installed during start up.

3 Basic usage guide

If you want to use any UNICORE components with the workflow, please also download the UNICORE-Taverna Plugin from http://unicore-dev.zam.kfa-juelich.de/taverna/plugins/, following the instructions on http://unicore.eu/community/development/Taverna/taverna.php.

Configure and run the Optimization

First of all, please open a workflow in the Workflow Diagram or design a new workflow, which you want to optimize. After opening the workflow, open the optimization perspective, at the top left in the Taverna Workbench.

Figure 1: Click the "Optimization" button to open the optimization perspective.
Then select the applications that should be optimized and enter the details for the optimization process (cf. Figure 2). The Evolutionary Steps is a number to define how many generations should be created during the optimization process. The Generation Size defines the number of chromosomes within one generation. The No change within x generations defines the number of generations that will be executed until the best value does not change within the last x generations. Maximum runtime in minutes defines the overall runtime of the optimization process including ALL executions (includes queuing). After configuration please click "Configure/Refresh Optimization Details".

Figure 2: The optimization perspective: (1) The workflow selected to be optimized. (2) Select the applications to be optimized during the optimization process. Thereunder, the parameter of the optimization algorithm can be specified. (3) Specify ranges and dependencies of application parameters. (4) Set ranges for dependencies.

A new window will be created (cf. Figure 2 (3)) to enter the ranges for the parameter. Please check all the parameter that should be optimized within the optimization process. Please do not forget to select the output parameter, which should serve for the fitness evaluation during the optimization process under the fitness tab. Additionally, it is recommended to set ranges for all the parameter and dependencies if available. The round to radio button will round the value until the second value after the dot. The function type will select only values along a specific function as input for the respective parameter. When selecting a function type, the min and max value refer to the value of the function. (E.g. Function is \( x^2 \), min = 3, max = 6, min value = 9, max value = 36).

Figure 3 illustrated how dependencies of two parameters can be set. You can only set a dependency to two parameters of the same application. If you want to use the absolute value you should check the “abs” button.
Finally, click the "Run optimization" button (cf. Figure 4), located next to the normal workflow run button. The green arrow is on a black background. A normal configuration window opens and you can enter the input values for the remaining parameters.
3.1 Requirements

Please be sure that the Workflow Engine (including the Service Orchestrator!!!) is installed on the target machine. Additionally, the admin of your site has to configure all applications via the execution environment in the IDB. Otherwise you will not be able to access any application from within Taverna.


4 Known problems and open issues

Dependencies
Dependencies can only be set for parameters that belong to the same application. Additionally, dependencies can maximum have two variables. This limitation might be updated in the future. You have to touch both variables before running the optimization.

Merge
If the workflow contains a merge activity, the sub-workflow required to perform the optimization cannot be created. Please remove the Merge activity!

Parameter Type
Currently, the file and string parameter type is not configured properly and cannot be used.

Output
When your workflow contains more than one output, the optimization cannot be configured properly. Please remove the unused outputs during the time of optimization.

Naming
The naming of activities, parameter and sub-workflow has to be unique.

Grid execution
Please do not use the option “Optimization on the Grid”. This option requires having a Taverna Server installed on your remote cluster. This Taverna Server needs to be configured correctly and requires an own certificate!