Tutorial: Computing alignments using decomposition techniques

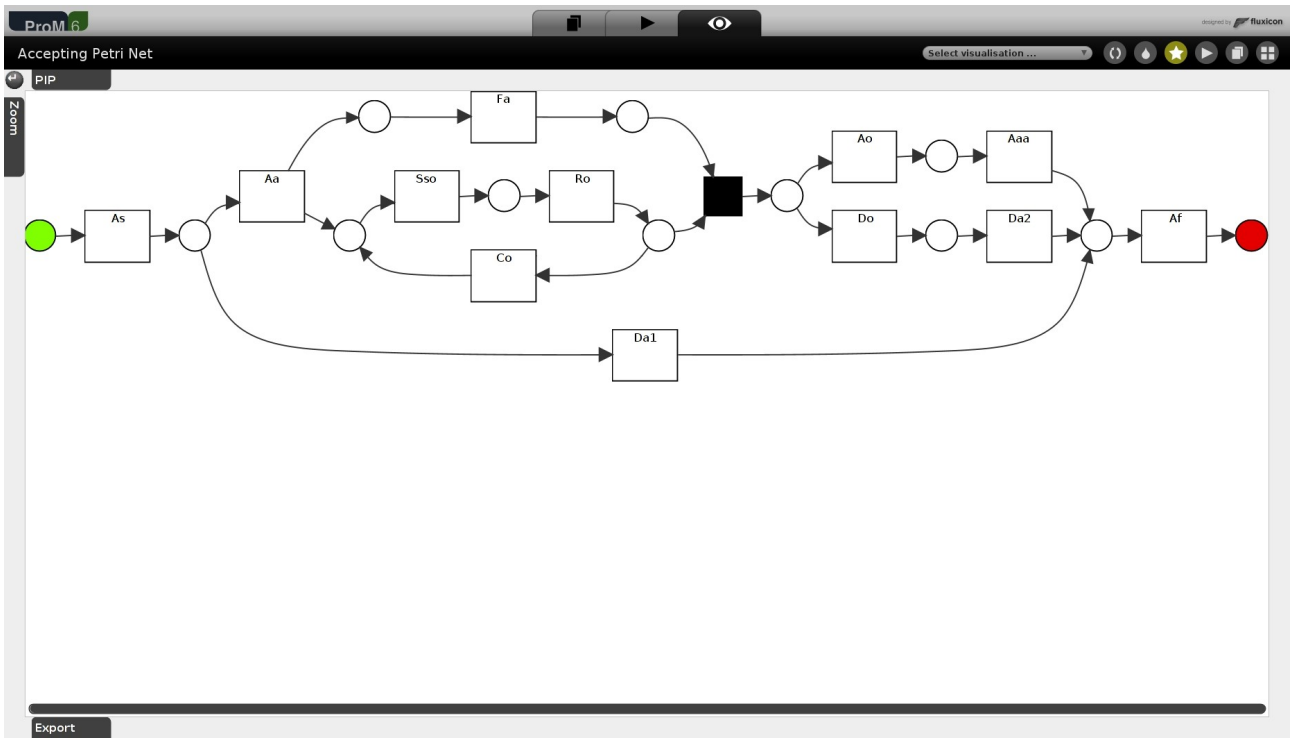
This simple tutorial will provide you the necessary steps to use ProM6.8 for computing the optimal alignments of a process model with respect to a given log. The underlying theory can be found in Chapter 9 of the book "Conformance Checking – Relating Processes and Models".

Dataset

Process model associated to the file M9.apnml, and event log associated to the file M9.xes.gz, available at the webpage of the book.

Loading the Model and the Event Log

Once you have downloaded the dataset of the tutorial, you should be able to upload them into the ProM6.8 platform. The process model should look like:



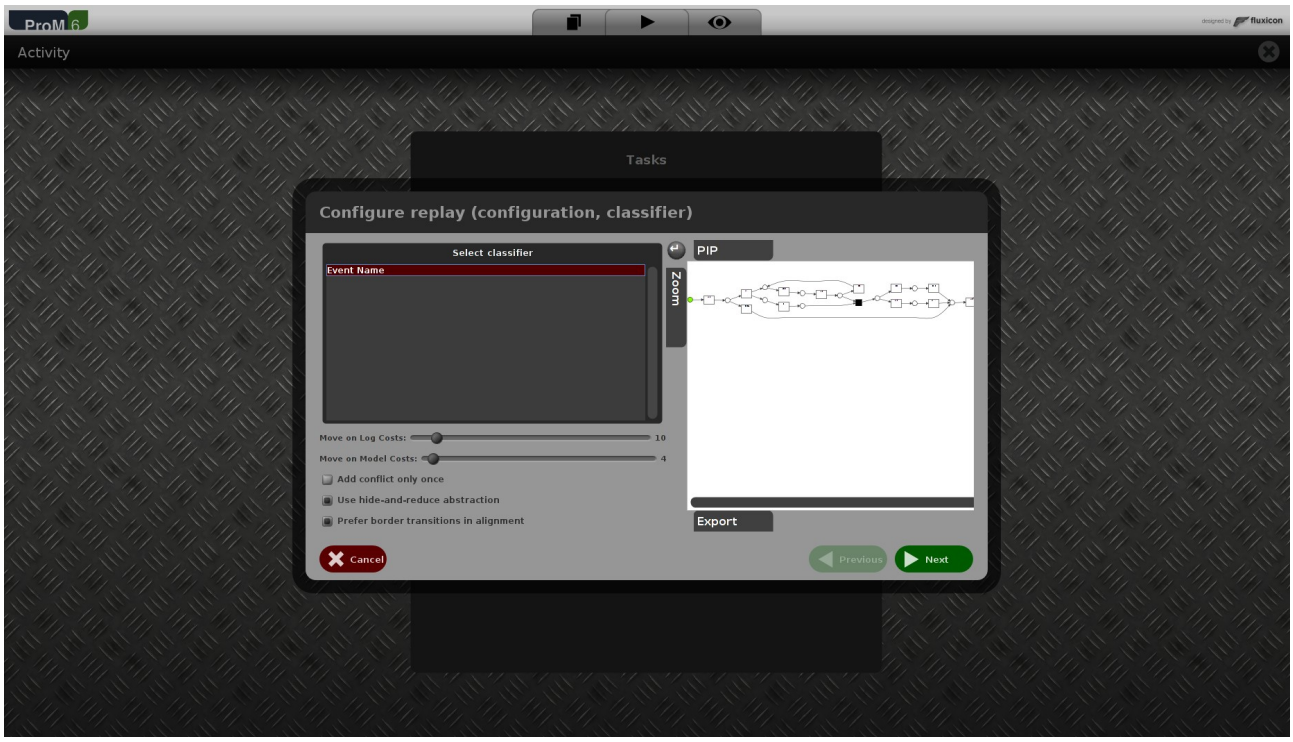
The event log contains 1000 cases, and 12 different activities.

Aligning the Event Log and the Process Model

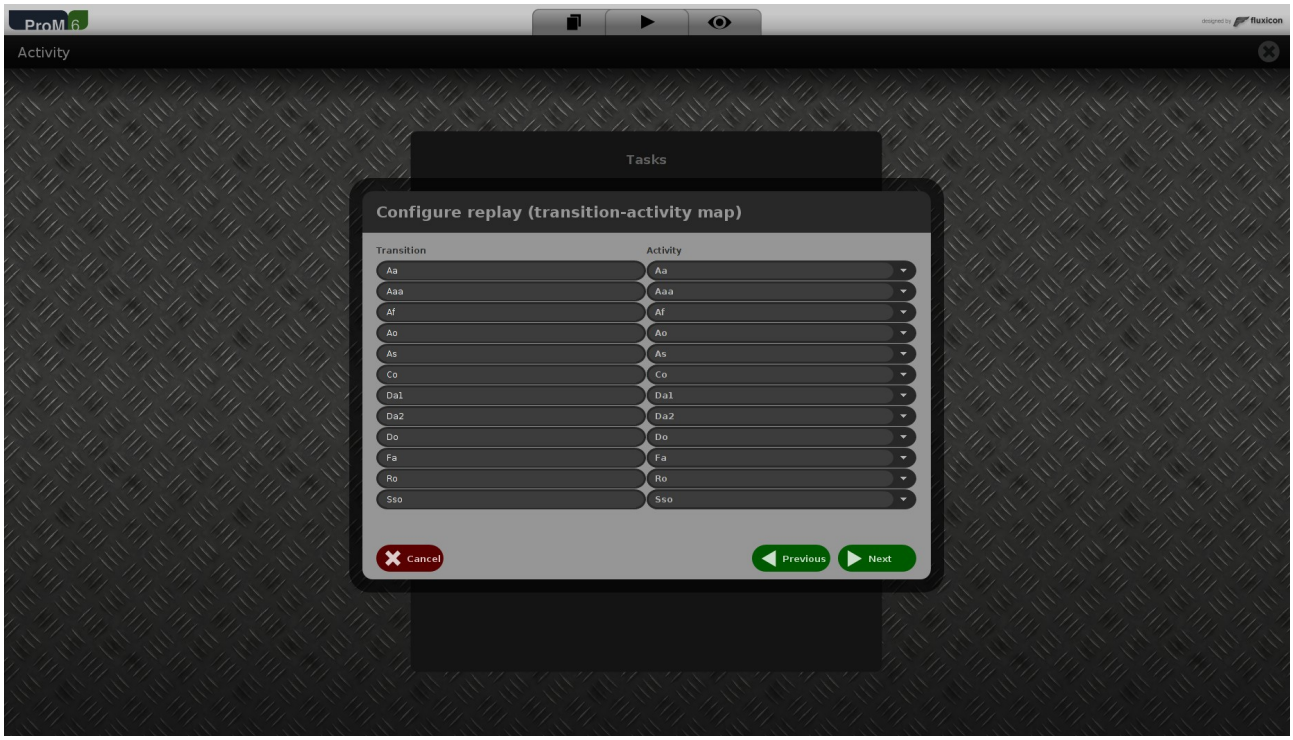
The second step is to compute an alignment of the event log and the process model using decomposition techniques. For this, click on the Play button and add both the event log and the process model. Then select as Action the plugin "Replay using Recomposition". The window should look like the following:

The plugin can then be started by clicking the Start button at the bottom part of the screen. There are a number of configurations for the “Replay using Recomposition” plugin because it includes both the decomposition of the model and log and the alignment of the data.

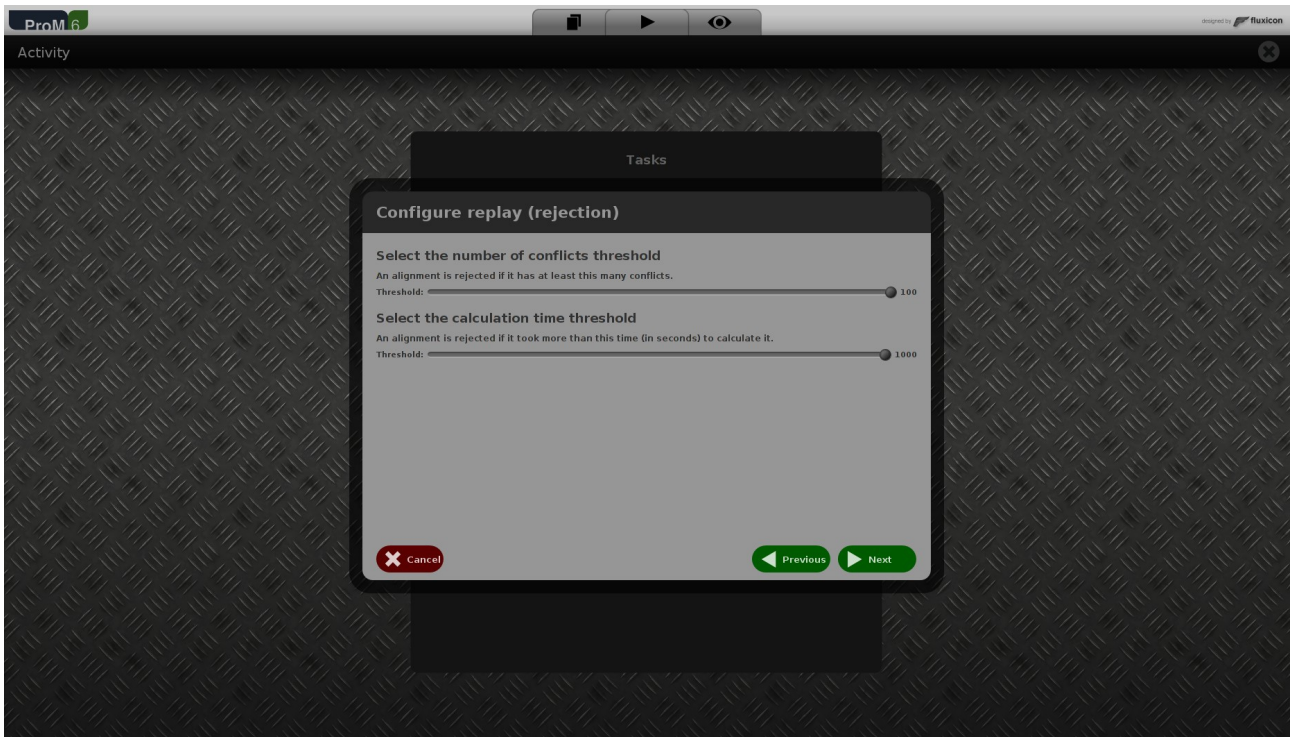
The first things to configure are the costs assignment to log and model moves as well as the event classifier to map log events to model activities. There are also three optional parameters related to the decomposed replay approach and the way in which statistics are to be recorded. For general purposes, it is fine to leave them at default settings. The window should look like the following:



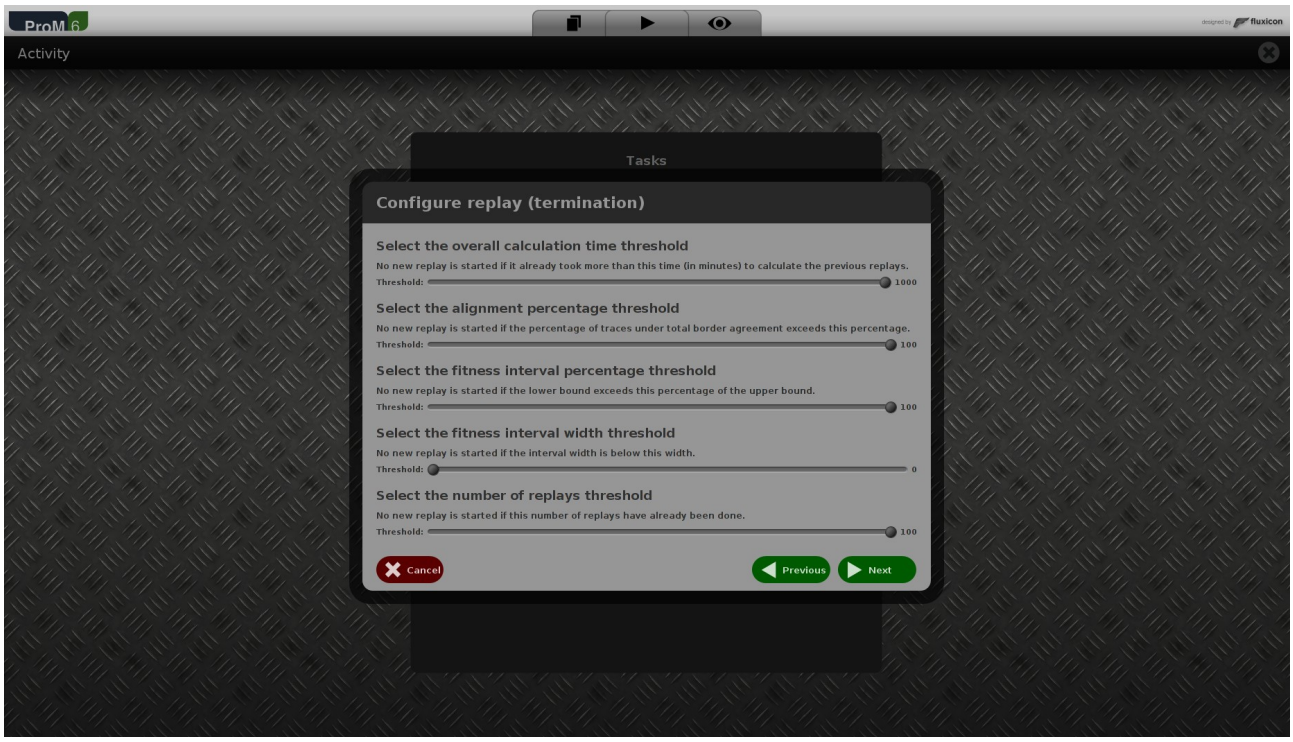
Pressing the Next button will show the transition-activity mapping computed by the previously selected event classifier. Here, final changes can be made to ensure correct mapping. The window should look like the following:



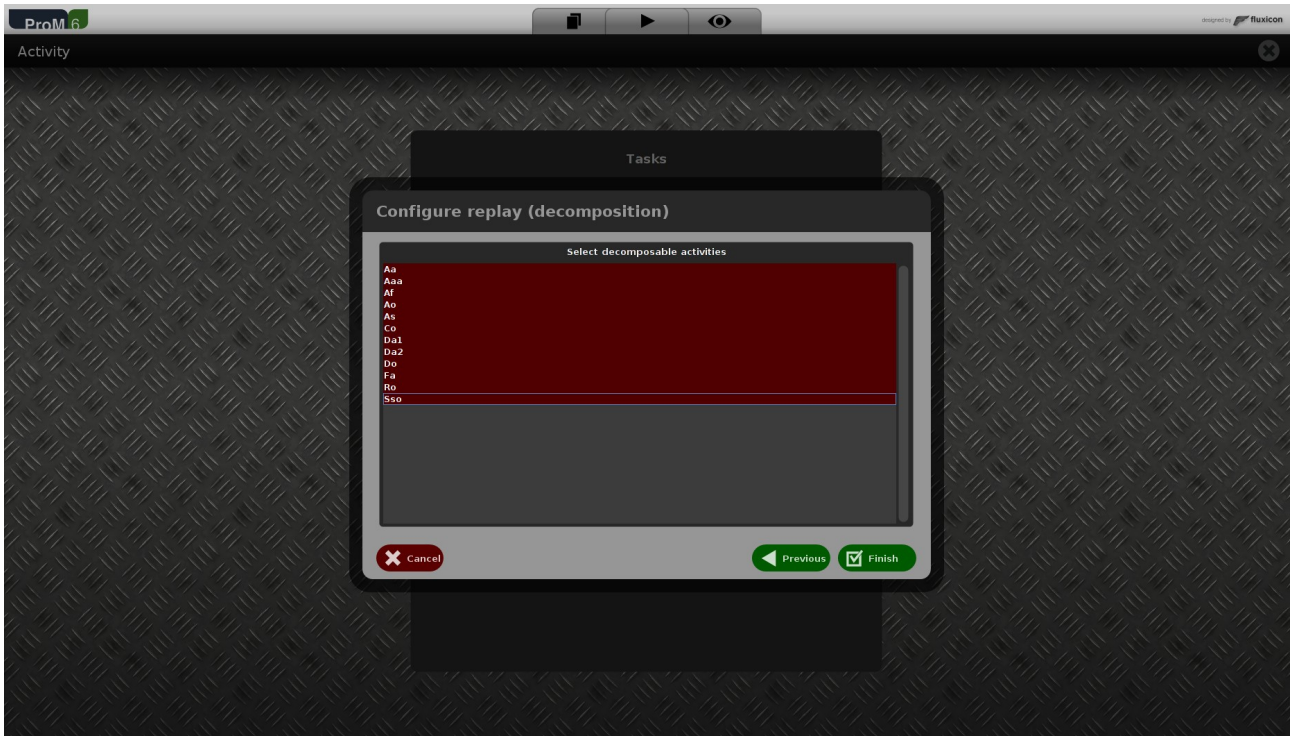
Pressing the Next button will show the configuration of the rejection conditions for individual log traces. In particular, one can adjust the maximum number of conflicts and computation time allowed before ceasing the exact alignment computation. The window should look like the following:



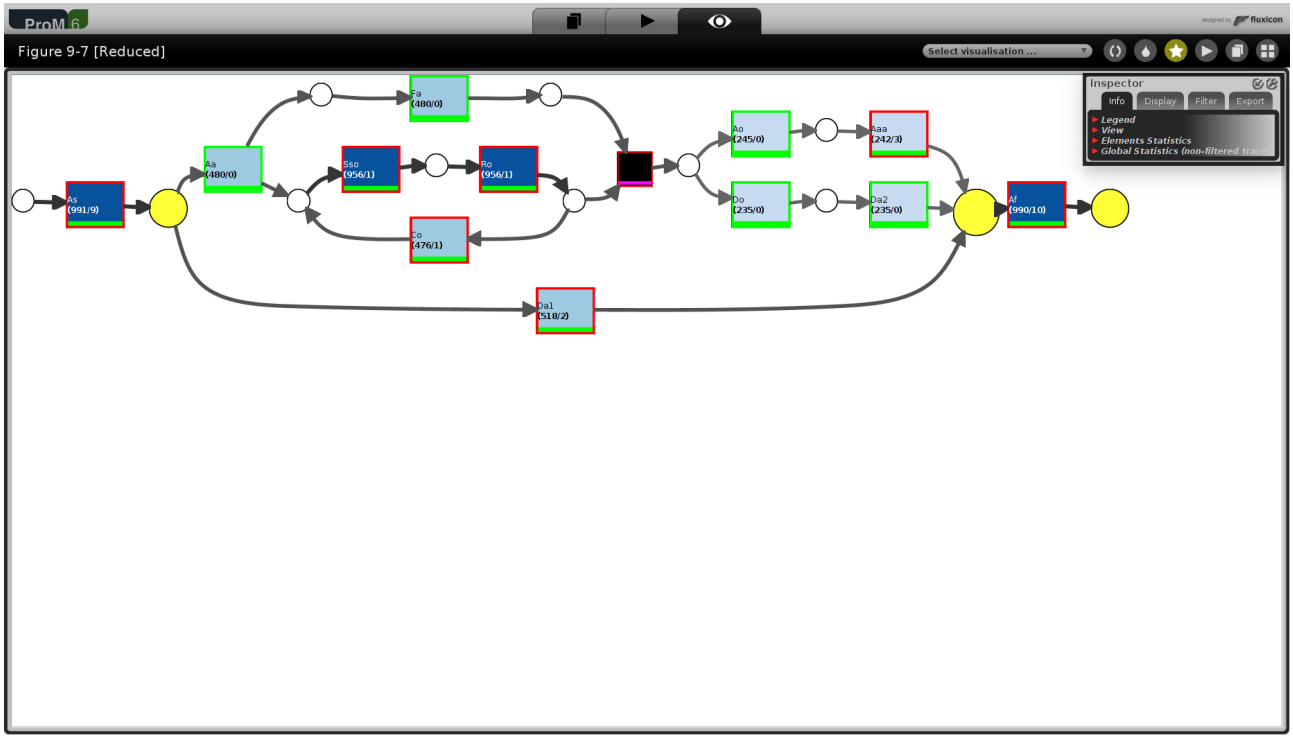
Pressing the Next button will show the configuration of the termination conditions on the overall level. There are five configurable parameters. First, one can adjust the overall computation time threshold and the exact alignment percentage threshold. Setting the alignment percentage to 100% computes the alignments of all log traces. Some result precision can also be traded off for faster computation times by adjusting the fitness interval percentage and width threshold. Adjusting the interval percentage to below 100% or the interval width to above 0 will terminate alignment computation before all the log traces are aligned exactly. Lastly, the maximum number of replays can be adjusted to prevent the algorithm from going through many iterations. The window should look like the following:



Pressing the Next button will show the last configuration window. Here, the decomposable activities can be selected to decompose the event log and process model. The selected activities will be the ones on which the process model will be split upon; the event log is decomposed by projecting it onto the sub-nets created by the model decomposition. The window should look like the following:



The Finish button can then be pressed to start the alignment computation. Once the alignment between the event log and process model is finished, the replay result will be shown. As recalled, alignment results can be approximate, i.e., not exact, due to the plugin configuration. However, the interpretation of both approximate and exact alignment results are similar. Same as The Lab section of Chapter 7 of the book, the alignment results can be projected onto the process model or the event log. The model projection should look like the following:



The model projection should show that there are log moves in places p1, p10, and p11. Moreover, there are a few model moves associated with the activities As, Sso, Ro, Co, Da1, Aaa, and Af. Note that there are only model moves associated with the skip activity because it is invisible and therefore unobservable from the event data.

To analyze the alignments, project the result onto the event log. Other than the alignments, alignment statistics are also shown on the left side-panel. In the dropdown box, the fitness percentage interval can be found by looking at the Raw Fitness Perc Lo and Raw Fitness Perc Hi. The fitness cost can also be found similarly by looking at the Raw Fitness Cost Lo and Raw Fitness Cost Hi. The log projection should look like the following: