

ProCKSI - Bug #39

Standardisation of big dataset takes too long for browser

06/03/2008 12:35 PM - Anonymous

Status:	Closed	Start date:	
Priority:	High	Due date:	
Assignee:	Anonymous	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:	9.0		
Resolution:	fixed		
Description			
<p>I have been analysing the dataset of 361 proteins in particular I was running DALI Lite calculations. The results were finished in about 5 and half hrs time. However, while standardising the results I faced the "Internal Server error!" with no results from standardising. Upon refreshing and updating the contents of the analysis manager no SSM and SSV results were found. The request number was 133.</p>			

History

#1 - 07/19/2007 02:13 PM - Anonymous

Problem:

It seems that we have reached the limit of the size of the dataset that can be handled with the current system design. The standardisation process seems to take too long as for the browser to keep the connection open.

Intermediate Solution:

- An intermediate solution might be that you standardise each similarity measure separately and not all of them in one go. No guarantee that it will work with these large datasets.
- Additionally, one can try and go "back" with the browser button and then *update* the analysis page with the given link. If the standardisation has been completed successfully, you will be able to access the SSM files.

Parmanent Solution:

This problem will be scheduled for a release in the future especially dealing with problems for large datasets.

#2 - 11/27/2007 09:57 AM - Anonymous

- Status changed from New to In Progress

#3 - 04/08/2008 10:24 AM - Anonymous

Refer to #76 for the proposed solution strategy.

#4 - 06/03/2008 12:35 PM - Anonymous

- Status changed from In Progress to Closed

- Resolution set to fixed

(In r690) f - Prepare/analyse big datasets (fixes #39 and #76): * analysis.cgi

f - similar to dataset.cgi: send time consuming analysis process into background; show/refresh "waiting/processing" page in the foreground (browser)

* dataset.cgi

b - end backgrounding when errors have occurred

f - Scheduling from within the workbench and the command line: * general in admin/ directory:

- c - uses Messages objects for messages collection/output (fixes #4)
- f - scheduling can be set into "full" or "secure" administration mode, whereas "secure" does not allow certain actions, e.g. cancel/delete a running task/request * tasks.cgi
 - f - restart tasks that have finished with errors
 - f - cancel running tasks if administration mode is "full"
 - f - delete tasks is implemented but not available in the browser as it can delete the "chains" directory when deleting a "structures" or "contacts" task
- c - createOutput() => addInputData() * experiments.cgi
- f - can potentially handle multiple experiments, thus experiment.cgi => experiments.cgi
- f - delete experiment, which deletes all data and meta-data from disk and database and logs user automatically out as the user is only logged in for ONE experiment * Scheduling/Tasks.pm
 - f - restart() sets status to "prepared" (P) and re-initialises the times
 - f - validates matrices when finishing a task and only registers valid ones into the database

e - CGI/Templates

- first steps towards more/better reusable classes for CGI/Templating
- c - CGI.pm => CGI/General.pm
- f - TEST: CGI/Cache.pm, which allows caching HTML from templates to disk
- f - TEST: CGI/Process.pm, which allows backgrounding processes, especially useful for big datasets.

BUGFIXES:

- b - online.html: incl. server name in link for "New Experiment" button
- b - [[TypedVectorMatrix]].pm: throw Exception::InvalidMatrix if target identity (first row, first column) is different from target-target comparison (first row, second column)
- b - experiment|dataset|tasks.cgi: distinguish between 'execute' and 'cmd' actions
- b - [[DaliLite]]: link to error files instead of linking to somewhere on the HTML page
- b - [[MaxCMO]]: add raw results before sequence comparison in output
- b - myexp.cgi: delete session before preparing output in logout()
- b - [[FormSpecial]].js: disable reference checkbox if all other checkboxes are disabled
- b - HTML: Phylipp => PHYLIIP (fixes #83)
- c - HTML: help (closes #84)

GENERALISATION:

- c - CONF_MAIN => CONF
- c - CONF_SETTINGS => SETTINGS