



SOMA₂

SOMA2 Case Example: A Basic Project

Tapani Kinnunen
CSC – IT Center for Science Ltd.

SOMA2 Case Example: A Basic Project

Version: 1.4 Aluminium

Written by: Tapani Kinnunen

Copyright (C) 2006-2012 CSC – IT Center for Science Ltd.

All rights reserved.

Table of Contents

| | |
|--------------------------------------|---|
| Table of Contents | 3 |
| 1. Introduction | 4 |
| 2. Launching SOMA2 Application | 4 |
| 3. Starting New Project | 4 |
| 4. Navigation and Help System | 4 |
| 5. Inputting Molecules | 4 |
| 6. Creating a Workflow | 5 |
| 7. Submitting a Project..... | 6 |
| 8. Project Overview..... | 7 |
| 9. Results..... | 7 |
| 10. Project Completed | 8 |

1. Introduction

This document describes a step-by-step introduction to SOMA2 modelling environment. The document includes all phases of a basic modelling project that can be done with the SOMA2 system. All the available features of the SOMA2 environment are not described in detail in this document. If additional information is required, the reader should see the SOMA2 user manual that can be directly accessed from the SOMA2 help system. The reader should have access to the SOMA2 system to fully benefit from the walkthrough style explanation.

2. Launching SOMA2 Application

First start a web browser. Any of the common browsers should work fine with SOMA2. Then type address <https://sui.csc.fi/group/sui/soma2> to the address field of the browser. This will load the login page of the system. By supplying your username and password and proceeding, the SOMA2 application is launched.

3. Starting New Project

The first page shows a project list that can be used to examine the projects configured earlier. If this is the first time that you use SOMA2, only a welcome message is shown. A new project is also started from this page. Click “New Project” button to initialise a project creation. This takes you to “New Project” page where a name and model template are set for the project. Use the provided textfield and type in a name for your project. The default is to start from empty template with no predefined settings. In this demonstration project the empty model template is fine. Then proceed by clicking “Initialise Project” button.

4. Navigation and Help System

After the project is initialised, the next page is “File Upload Page”. But before proceeding, get familiar with the navigation and help system of SOMA2.

The navigation links are located in the upper part of every page of the SOMA2 environment. Before you actually commit your project for computation, the navigation links include “Project List”, “Add Molecules”, “New Step”, “User Info” and “Help”. These links are used to move between different phases of the project and can be accessed at any point while working with the SOMA2.

Then move the mouse pointer over the navigation link “Help” and click mouse button. As you see, a popup window opens and shows the help text. To close the window, click “close” icon in the popup window. This is the way to use the general help texts in SOMA2 system. The content of the help text is changed according to the page where the user currently is. The help text boxes always include a direct link to full SOMA2 user manual which is available online.

5. Inputting Molecules

“Add Molecules” navigation link is used to access the three ways of inputting molecular data to SOMA2. First one is to upload a molecule file from your own local computer. If you did not try any of the navigation links, you should be on the “File Upload Page”. If you are on some other page, click “Add Molecules” navigation link to return to the “File Upload Page”. To perform the file upload, first check the supported molecule file formats using “[Supported Filetypes]” link on the page. Then click “Browse...” button and use the opened window to search for a molecule file in your local computer. After this select the file and double-click the left mouse button or press “open” to start the upload of the file to the SOMA2. After this “Molecule Files:” list appears showing the uploaded file. If you want to add explicit hydrogen atoms

and/or create initial 3D-structure for your input, check “Add Hydrogens” and/or “Clean Structure” checkboxes before selecting a file to be uploaded.

All the basic conversions and preparations for the file are made automatically. If the status of the uploaded file in the details of the “Molecule Files:” list is “ok”, all the preparations are successful. If the uploaded file contains only a single molecule, both the molecule name and the file name are shown in the details. If the uploaded file contains more than a single molecule, only the file name is shown. Please note that the initial preparation is performed in the background and it can take some time. During the preparation the status of the file is “processing”. If “ok” status would fail to appear, the file can be deleted using “[DEL]” link in the details of the “Molecule Files:” list. Even the preparations are successful, click an embedded link in “Molecule Name / File Name” column in “Molecule Files:” list. This loads “File Details” page where you can examine the possible molecular properties that the input file includes. If properties exist, you can also change their names and data types using the provided web form elements. A property can also be checked to be inherited to all the descendants that are generated in the calculations performed with the SOMA2. If there are any properties that are not needed, you can delete them by checking “Delete” box on the page. If you make any changes, they are applied by clicking “Apply Changes” button. Clicking “Cancel” revokes the changes. Clicking either one of the buttons will take you back to the “File Upload Page”. The automatic preparation procedure of the input files applies to all three ways of inputting molecular data.

The second way of inputting molecules is to sketch them within the SOMA2 environment. Use “Sketch Molecule” button shown in the toolbox of the “File Upload Page” to change the view. While the view changes from the “File Upload Page” to “Sketch Molecule”, a sketcher applet tool is loaded and it can be used to draw a molecule. Help for using the sketcher tool can be accessed by choosing “Help” from the sketcher applet menu. So draw some structure with the sketcher. Then click the right mouse button while the mouse pointer is over the applet to open a menu. Choose “Transform/Rotate in 3D”. Then examine the structure with the three dimensional rotation tool by pressing the left mouse button and moving the mouse. As you see, the drawn structure is flat. By default, the applet draws 2D structures. Choose “Structure/Add/Add Explicit Hydrogens” from the applet menu to add all the hydrogen atoms. Then choose “Structure/Clean 3D/Clean in 3D”, again from the applet menu. Then rotate the model again and as you can see, the model now has a 3D structure and explicit hydrogen atoms. To save the sketched structure to the SOMA2, supply a name for the molecule in the provided textbox below the applet tool and click “Save Molecule” button. After this the structure is added to the “Molecule Files:” list. Instead of using the applet tool for adding hydrogens and creating 3D-structure, “Add Hydrogens” and “Clean Structure” checkboxes can be used also in this view. Sketched molecules are shown with file names such as “Sketched_1.sdf” in the “Molecule Files:” list.

In addition, each item in the “Molecule Files:” list also has “[DEL]” and “[Save]” options. These can be used to delete a file from the SOMA2 system or save file to the user. If inputted file contains only a single molecule, also “[Edit/View]” option is available. This can be used to edit or view the molecule with the sketcher tool in the “Sketch Molecule” page. During the initial preparation of a file, only “[STOP]” action is available. This can be used to terminate the preparation process of the file.

6. Creating a Workflow

Now you have added the molecules to SOMA2 and next phase in the project creation is to create a workflow of programs. In this demonstration the workflow generation is started after adding the molecules but it could be also done vice versa; the navigation between different phases and steps of a project is completely up to you.

Start the workflow creation by clicking “New Step” navigation link. This will load a page with a popup menu titled “Select Program:”. Then choose “3D-Property” from

the menu. With this selection, the configuration page of program "3D-Property" is loaded and as you can see, a link titled "3D-Property" is added to the navigation menu. So when a new program is added to the workflow, it is also added to the navigation. This will enable an easy switching between the configuration pages of different steps included to the workflow.

The program configuration pages of the SOMA2 have some general features. Every program has a brief description, which explains the usage of the program. If applicable, a link to the online manual of the current program is available after the description text. In addition to the general help system available through the navigation menu, all the user adjustable parameters have their own specific help texts. These are used by moving the mouse pointer over the question mark (?) characters shown next to each parameter. The parameter specific help texts are included to offer a quick runtime help for the users hence they usually are not complete. Then look at the parameters shown on the page. The checkboxes can be used but if you try the textfields, you will see that they are locked. Then check the parameter "Total surface area". Instantly after this, parameters "Area resolution" and "Probe radius" also become adjustable. So the two parameters affect the calculation of the "Total surface area" and can be modified since the "Total surface area" is now selected. This demonstrates the built-in automation of the SOMA2 environment; the user can adjust only the parameters that are valid for the current configuration. But as you see, there are already some default values included to the additional parameters, as are for almost all of the parameters of all of the programs that are available through the SOMA2. So for example here, you do not need to change anything. However, change the value of the parameter "Area resolution" from "180" to "-7". As the parameter value changes, its validity is checked. Now you will see an error message indicating that the supplied value is too small. Also the minimum possible value is shown. If you change "-7" back to "180", the error message disappears. Almost all of the user adjustable parameters in the SOMA2 are equipped with the information on the threshold values of the parameters to help the user setting up a valid configuration. With the current configuration, the total surface area of each molecule will be calculated with the default settings.

Now you have configured the first program of the workflow. Then configure another program. Choose "2D-Property" from the menu titled "Add New Step After This Step:" that is available in the toolbox of the configuration page. This will load the configuration page of the program "2D-Property" and "2D-Property" link is also added to the navigation menu. With the others controls in the toolbox, the current step of the workflow can be changed or deleted. The control tools also allow adding steps before or after the current program step. Then choose options for the program "2D-Property". Just select the properties that will be calculated. For example, uncheck "chiral atoms" and check "cLogP" value. There are no additional adjustable parameters for this program so the configuration is ready when the properties are selected.

The workflow that you now have created includes two programs. First, for the original input molecules, the total surface area is computed by the program "3D-Property". Then its results are input for the program "2D-property" that calculates the cLogP value for the molecules. Results of the second program will then be the final results of the workflow.

7. Submitting a Project

For your project, you now have added the molecules and also created the workflow to be executed. Then you can proceed to commit the project by clicking "Submit Project" from the navigation menu.

Now a page titled "Submit Project" is loaded. On this page you can see a brief review of the project that is about to be submitted. If SOMA2 system finds any inconsistency or erroneous parameters from the input molecules of the workflow, the "Submit Project" page also shows error and warning messages. If you have proceeded as

guided in this document, there should be no errors or warnings shown on the page. If you wish, you can still use the navigation menu to look or modify the program configurations or even add more molecules to the project. When you are ready, just use “Commit Project” button in the page’s toolbox to start execution of the workflow.

8. Project Overview

After committing the project a page titled “Project Overview” will be loaded. You will also see that the navigation menu changes a little. “Add Molecules” is replaced with “Project Overview” and “Submit Project” is replaced with “Results”. Also program links on the navigation are appended with link “View Config”.

The “Project Overview” page shows the current status of the project and statuses of each individual program that the workflow includes. Project and program statuses are updated automatically. Now you just need to wait until all of the program steps and the project reach “finished” status. When the project reaches the “finished” status, you can access the results of the workflow. Use either the navigation link “Results” or the corresponding button in the toolbox of the page to see the results of the whole workflow.

9. Results

“Results” page in SOMA2 shows the results in the spreadsheet like table, which is titled “Result Data:”. Also here the table is sorted by clicking the column headers and the table has its own scrollbars. Additional feature in the main result table is that property data fields can be shown or hidden according to the user selections. For example, move mouse pointer over column header of “Step Number” in the result table. Then click arrow icon that appears and choose “Columns” option from the menu. You should see list of all available properties for the data table. De-select “Step Number” by using provided checkbox. As you see, “Step Number” column no longer is shown in the table. Re-select “Step Number” and you should see it again in the table. Order of the shown columns can be set by dragging and dropping columns to desired positions.

The results in the table are presented in generations of molecules with property data field “Molecule Name” as the identifier for a molecule. The number suffixes of the molecules indicate the generation of a molecule. For example, in this demonstration the results of program “3D-Property” have a suffix “.1” and the results from program “2D-Property” have a suffix “.1.1” The original input molecules are always presented with their original names. The property “Step Number” can be used to quickly identify from which program each result molecule originates. In this demonstration, results from the program “3D-Property” have the “Step Number” value “1” and results from the program “2D-Property” have the “Step Number” value “2”. Sorting the table according to the “Step Number” enables easy picking of the results from one individual program.

Each of the individual molecules can also be visualised within the SOMA2 environment. As you see, also in the main result table there is embedded link in the “Molecule Name” property of each result. Click the link to visualise a selected molecule. This will load “Result Details” page that launches the viewer applet showing the molecular structure of a result. Default is to show the structure with 3D viewer but you can switch to the sketcher by changing the active selection from menu titled “Select Viewer:”. Moving the mouse pointer over the molecule followed by pressing left mouse button and moving the mouse enables the three dimensional rotation of the model. The “Result Details” page also shows a summary of the computed properties for the current result molecule. The links above the applet tool show the possible parent, sibling or children result molecules of the current result. The links can be used to directly load the “Result Details” page of a corresponding result molecule.

You can return to the “Results” page by clicking “Results” in the navigation or “[Results]” link below the applet tool.

After you have evaluated the results in the table you can also save them to your local computer. The results are saved by selecting a desired molecule file format from the menu titled “Format Selection” shown in the toolbox. Default is the “sdf” file format but other formats are also supported. Then select the preferred results by selecting the rows of the table. After this, click “Save Selected” button in the toolbox to start automatic download of the selected results.

10. Project Completed

At this point, you have performed a full project with the SOMA2 environment. Now you should have a basic knowledge of all the different project phases from inputting molecules to viewing the results. If you need any additional information of the features that the SOMA2 offers for the user, please consult the SOMA2 user manual.