A Statechart based representation for SBML descriptions

Fabio Fioravanti [1], Manuela Helmer Citterich [2], Enrico Nardelli [3]

[1] *Dept. of Sciences, Univ. Chieti-Pescara "G. D'Annunzio"*

[2] *Dept. of Biology, Univ. Roma Tor Vergata*

[3] *Dept. of Mathematics, Univ. Roma Tor Vergata, nardelli@mat.uniroma2.it*

1 INTRODUCTION

We consider SBML models represented in the Reactome web site [1]. SBML descriptions do not represent executable models, while the importance for biologists of having formal models, which both are executable and have a graphical representation, is more and more recognized [2, 3].

And a tool based on formal semantics, which allows a biologist to both visualize and simulate a biological model, can be very useful, because it helps to remove ambiguity and to increase confidence on the model itself. This holds, in particular, for wiki-like collaborative environments, where it facilitates the consensus reaching phase of the model construction process.

One of the main goals of the SIMBIOSYS project is to build a tool able to automatically derive a StateChart [4] representing the SBML description of a biological pathway or sub-event. This would allow to have both an executable representation of the SBML description and a representation that can be read at various levels of aggregation, as described below.

2 METHODS

The hierarchical structure of a given biological pathway or reaction is extracted from the Reactome site by parsing the appropriate html. Such hierarchical structure is used to generate a statechart which, extending the standard visualization, has two different views: the *compartment view* and the *reaction view*:

- the compartment view shows for each compartment all the species present in it
- the reaction view shows the hierarchical structure of reactions and expresses their dynamics.

It is important to stress that these are two different viewpoints of the same statechart representing the considered biological reaction. For each species present in each compartment the statechart has one variable representing its quantity.

3 RESULTS

3.1. The compartment view

The purpose of this view is just to show how species are distributed in compartments, hence no reactions are shown. Note that if the same molecule is present in different compartments, it is considered as different species.

The root state of the statechart is a *concurrent macro-state* (also called AND state) whose sons are concurrent macro states representing each one of the compartments defined in the reaction. Each AND state representing a compartment contains one atomic state for each species present in the compartment itself.

3.2. The reaction view

This view is an executable one. That is, given some user provided specifications of the quantities of some species, the statechart concurrently executes all enabled reactions and determines how the quantities of the various species are modified as a result of the reaction dynamics. Concurrent executions means that non-determinism is used whenever a plurality of reactions are enabled but only some can actually be executed due to the overall quantity constraints on species.

In this view the root state of the statechart is a concurrent macro-state (AND state) representing the whole reaction. A reaction can be either an atomic reaction or a composition of (sub-)reactions which can be executed in parallel.

If the reaction is an atomic one, that is a reaction which is not the composition of simpler reactions but which is a relation between *reactants*, *products*, and *modifiers*, then the macro state has two or three sons: one containing the reactants, one the products, and an optional one with the modifiers. Each of these is again an AND state containing one atomic state for each, respectively, reactant, product, and modifier involved in the reaction.

If the reaction is not an atomic one then its sons are the AND states representing the reactions it is made of. Each of these may either be an atomic one (and the previous paragraph applies) or a non atomic one (and this paragraph recursively applies).

To each atomic reaction a *transition* is associated. Such a transition has exactly one branch ending on the AND state representing the entire atomic reaction. To this ending branch the actions modifying the quantities of species involved in the reactions are associated. The transition has also exactly one branch starting from each reactants and modifier present in the atomic reaction: all this branches converge in a *join connector* the ending branch of the transition departs from. To each starting branch of the transition a condition is associated giving the quantity constraint the species the branch start from has to satisfy for the reaction to happen.

3.3. Views at higher aggregation level

Given the hierarchical structure of the reactions, one may consider the given overall reaction at an aggregation level higher than the atomic one of the atomic reactions. This means reading the reaction view without going down until the atomic level.

In such a case our tool automatically derives what we call *abstract transitions*, that are transitions showing the relations between non-atomic reactions. We call them abstract transitions because they are not directly executable, but their semantics is given by the composition of the underlying reactions.

3.4. An example

We start from the SBML description of the biological reaction "EGFR interacts with phospholipase C-gamma [Homo sapiens]" present at: http://www.reactome.org/cgi-bin/eventbrowser?DB=gk_current&FOCUS_SPECIES=Homo%22sapiens&ID=212718.

The statechart representing the SBML is shown in Figure 1, where the number below each species name in the atomic states is the species stoichiometry in the reaction. For clarity the arrows representing the transitions have been omitted.

The view at a higher aggregation level, where the atomic reactions are not shown, is presented in Figure 2: here the automatically derived abstract transitions are shown.

A complete reaction view with transitions, including conditions on starting branches and actions on ending branches is presented in Figure 3 for an atomic reaction.

ACKNOWLEDGEMENTS

Supported by MIUR PRIN 2008 Project: "Simulation and modeling of biological systems using StateCharts (SIMBIOSYS)".

REFERENCES

[1] Matthews L, Gopinath G, Gillespie M, Caudy M, Croft D, de Bono B, Garapati P, Hemish J, Hermjakob H, Jassal B, Kanapin A, Lewis S, Mahajan S, May B, Schmidt E, Vastrik I, Wu G, Birney E, Stein L, D'Eustachio P. *Reactome knowledgebase of human biological pathways and processes*. Nucleic Acids Res., 37:D619-D622, 2009. http://www.reactome.org/

[2] David Harel, Yaki Setty, Sol Efroni, Naamah Swerdlin, Irun Cohen, *Concurrency in Biological Modeling: Behavior, Execution and Visualization*, Electronic Notes in Theoretical Computer Science, 194:119-131, 2008.

[3] Hillel Kugler, Antti Larjo, David Harel, *Biocharts: a visual formalism for complex biological systems*, J.R.Soc.Interface 7(48):1015-1024, Jul.10, published online 18dec09.

[4] David Harel, *Statecharts: a visual formalism for complex systems*, Sci. Comput. Program. 8:231-274, 1987.

TABLES AND FIGURES

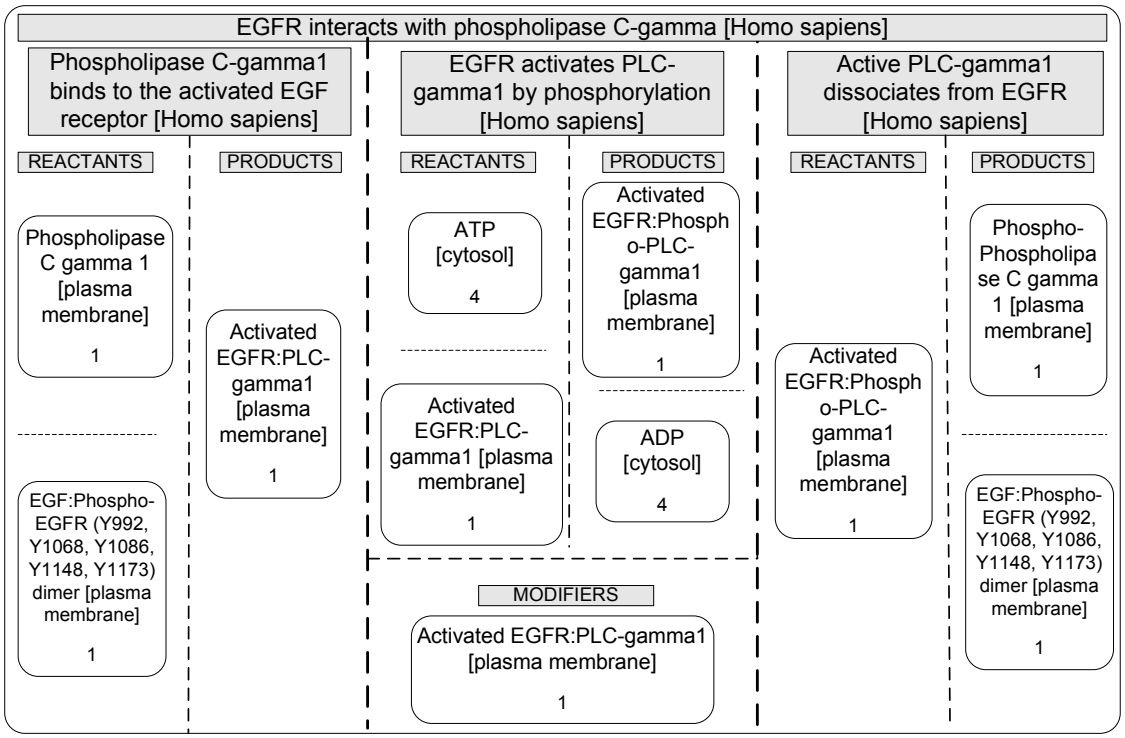

Figure 1: The statechart representing the SBML description for "EGFR interacts with phospholipase C-gamma [Homo sapiens]"

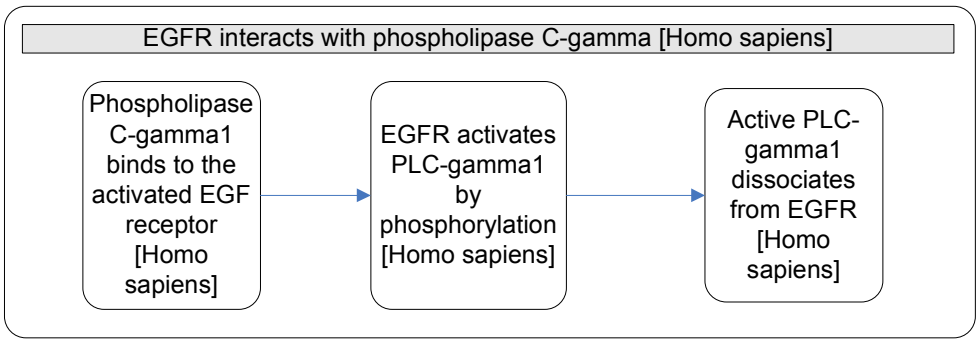

Figure 2: The previous statechart shown at a higher aggregation level, without the atomic reactions

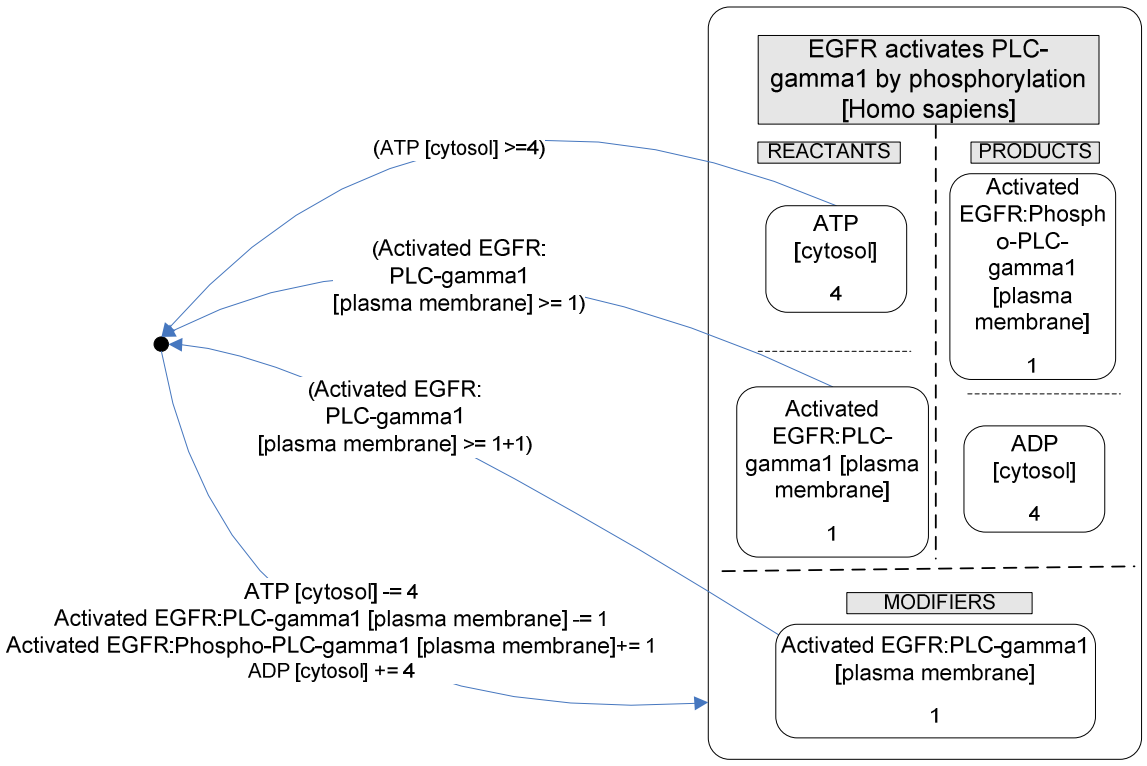

Figure 3: A statechart showing a complete reaction view with transitions