Cluster Based Analysis Of Petri Net Properties

Marcin Radom
Institute of Computing Science
Poznan University of Technology
Poznań, Poland
mradom@cs.put.poznan.pl

Agnieszka Rybarczyk
Institute of Computing Science
Poznan University of Technology
and
Institute of Bioorganic Chemistry
Polish Academy of Sciences
Poznań, Poland
arybarczyk@cs.put.poznan.pl

Piotr Formanowicz
Institute of Computing Science
Poznan University of Technology
and
Institute of Bioorganic Chemistry
Polish Academy of Sciences
Poznań, Poland
piotr@cs.put.poznan.pl

***Abstract*— Systems biology is an interdisciplinary field of study based mainly on biology, mathematics and computer science. In systems biology, the analysis of complex biological systems of interacting objects is being made. Before the analysis starts, a model is being created, representing the analysed system. The theory of Petri nets offers a necessary tool for such a task. Places in such a net correspond to the static elements of the system like chemical compounds. Transitions correspond to the reactions. The rules of transitions activation and firing allow the modelling of the biological system dynamics. Having a Petri net one can start the analysis which is based on the invariants, maximum common transition sets and t-clusters. Such analysis requires various tools using different file formats which adds to the complexity of such a task. We have developed a Java-based environment to help in the process of net creation, simulation and the analysis in the classical and time Petri nets, while the different formats in which Petri net data is being stored can be used there. Exporting data to other Petri net tools has also been implemented.**

Keywords–Systems biology; Petri nets; model analysis.

I. INTRODUCTION

Systems biology is an interdisciplinary field of study based on biology, mathematics and computer science which focuses on the analysis of complex biological systems. The first step before the analysis is the creation of a model of the system, and its further analysis can lead to the potential discoveries of the behaviors of the system [1] [2]. One of the possible tools for the creation and analysis of such a model is the Petri net theory. A structure of a Petri net is a bipartite directed graph of places and transitions. Places correspond to the static element of the system like chemical compounds, products, substrates, etc. Transitions on the other hand allow modelling of the system dynamics. They represent the reactions taking place within the system. Transitions in the Petri net can be activated and fired depending on the status of the places. The latter depends on the number of tokens in each place, i.e., the number of products of substrates a given place represents. A semi-positive (i.e., non-negative) vector describing the number of tokens in each place is called a marking. Finally, weighted arcs in a Petri net define how many tokens in places a fired transition consumes and produces.

Analysis of such a model is not an easy task and it is divided into different consecutive steps. First, the qualitative and behavioral properties of the net must be determined. Classical Petri net can always be described using a so called incidence matrix. Each entry of this (place $\times$ transition) matrix C gives the token change on a particular place by firing of the respective transition.

On the base of such matrix, the transition invariants can be computed. A t-invariant is a vector $x \in N^m$ where $C \cdot x = 0$ where m is the number of transitions. Every t-invariant corresponds to a set of transitions whose firing a number of times reproduces a given marking of the net. Firing number for each transition is defined within the x vector. Such a chain of reactions that each t-invariant describes, represents some basic behavior of the modelled biological system.

Next, the maximum common transition sets (MCT) can be computed based on the set of t-invariants. The MCT-sets partition the set of transitions into disjoint subsets whose biological meaning can be determined [2]. In other words, the MCT sets consist of the reactions linked with each other.

The most difficult task, yet giving the most valuable knowledge about the modelled system is the cluster analysis. In this phase the set of t-invariants is being divided into disjoint subsets called clusters. To perform this task various tools and methods are necessary. Invariants have to be converted into csv file, which then can be used in, e.g., R language environment. Using cluster algorithms and distance metrics from the R libraries and scripts specially prepared for this task, one can compute different clusterings (i.e., sets of clusters). In order to evaluate clusterings and choose the optimal one, the additional evaluation metrics are necessary. Often Mean Split Silhuette (MSS) and Celiski-Harabasz index [3] are used for such a task. Finally, when the correct clustering has been chosen, the biological meaning of clusters can be determined. This step can lead to the potential discoveries of the new facts concerning the modelled biological system.

All these steps are quite difficult and the number of methods and tools used adds to the overall complexity of Petri net analysis. Drawing a Petri net requires knowledge not only about the biological system, but also about current structural properties of the created Petri net that must be computed. All these problems on a way of successful Petri net model analysis were the reasons to develop a Java-based program, providing aid in all steps of the task: creating, simulating and analysing the given net.

 ISBN: 978-1-61208-409-1

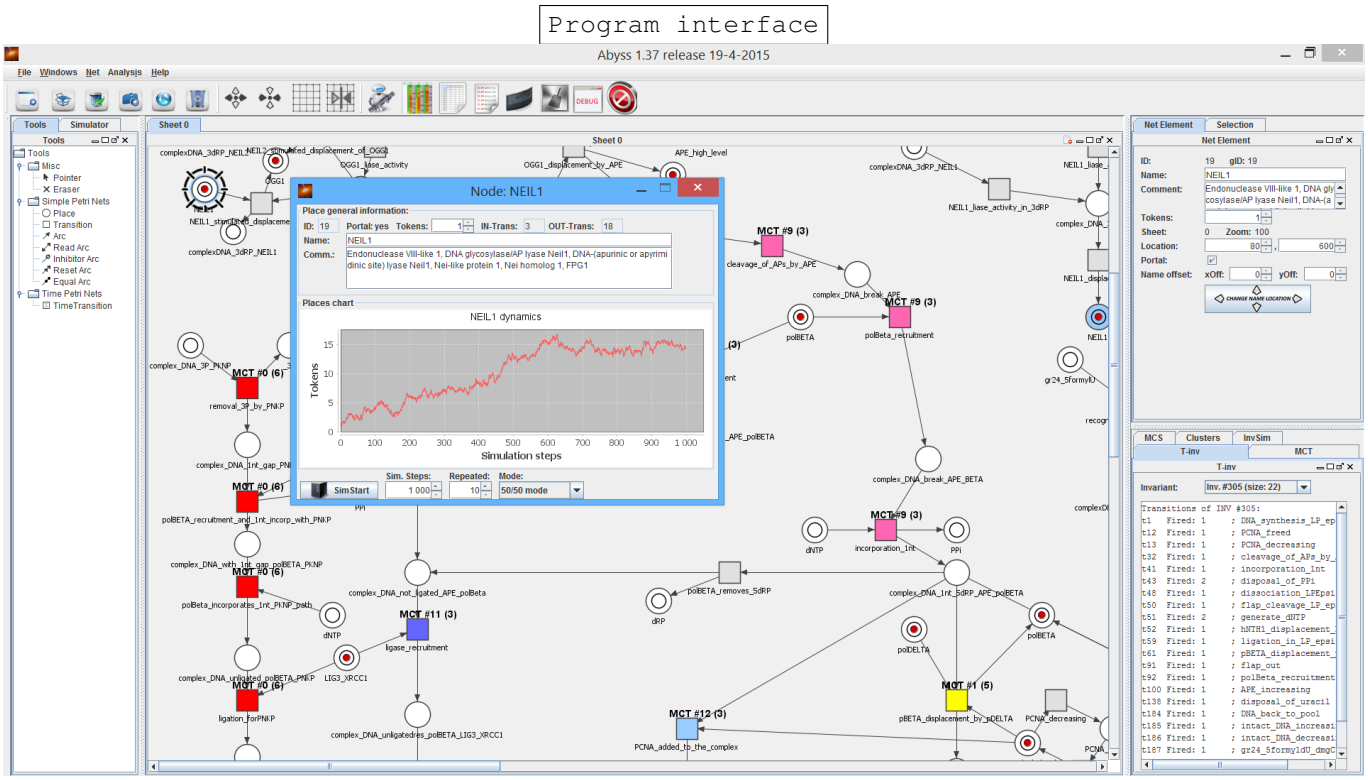

Figure 1. IPNE Java-based environment.

A. *Software*

The proposed tool is a Java-based program consisting of various modules supporting different phases of Petri net-based model analysis [4] [5]. In the current version, the user can draw a classical Petri net, time Petri net and also a net consisting of both classical and time components. The main window of the proposed tool is given in Figure 1. Net qualitative properties can be easily computed and displayed right from the very beginning of drawing the net, which often helps in detection and avoidance of potential problems within the net structure. One of the popular editors for Petri nets is a program called Snoopy [6]. Our application can both load and save nets in the Snoopy format.

When the net is ready, invariants can be obtained using either implemented algorithm within the program or from the external source like Integrated Net Analyzer (INA) [7]. MCT sets are computed as well by our algorithm and both the invariants and the MCT sets can be displayed on the structure of the net within program main window. Additionally, when displaying each given invariant, the number of firings for each transition is also given.

The most complex task lies in the cluster analysis as we have already stated. Our program integrates various R scripts used to compute clusterings using different cluster algorithms and different metrics. Seven cluster algorithms and eight distance metrics can be used, for each of them results for a range of numbers of clusters are given. This allows the user to choose the optimal clustering based on MSS and Celiski-Harabasz evaluation metrics which are computed as well. All these computations are performed in the background by our scripts within the R language, and when the computations are finished our program integrates the data into a table of clusterings. Every clustering (i.e., set of clusters for a given clustering algorithm, distance metric and the number of clusters) can be chosen for more detailed computations. Given the information about the appearance of each invariant within a specific cluster one can obtain important knowledge about the clusters which helps in finding their biological meaning. The number of MCT sets within the cluster and the number of firings of transition are given, along with the MSS measures for each cluster within the clustering. The user can send this data into the Petri net displayed in the main window of the program and analyse the structure of the clusters seeing each transition and its role within them.

II. Conclusion

The described tool is still being developed, but even now we obtained a quite powerful tool aiding the tasks of Petri net drawing, simulation and analysis. The most time consuming task, i.e., the clusters analysis have been made much simpler and faster. Therefore, more time can be devoted to detailed analysis, while all the computations and potential problems with different file formats describing Petri nets are solved in the background by the program. Each clustering and every cluster within it can be analysed thoroughly, while the data concerning them are given in separate windows.

Acknowledgment

This research has been partially supported by the Polish National Science Centre grant No. 2012/07/B/ST6/01537.

 ISBN: 978-1-61208-409-1

REFERENCES

[1] D. Formanowicz, A. Kozak, T. Glowacki, M. Radom, and P. Formanowicz, "Hemojuvelin-hepcidin axis modeled and analyzed using Petri nets," Journal of Biomedical Informatics, vol. 46, 2013, pp. 1030–1043.

[2] D. Formanowicz, A. Sackmann, A. Kozak, J. Blazewicz, and P. Formanowicz, "Some aspects of the anemia of chronic disorders modeled and analyzed by petri net based approach," Bioprocess and Biosystems Engineering, vol. 34, 2011, pp. 581–595.

[3] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," Communications in statistics, vol. 3, 1974, pp. 1–27.

[4] H. Andrzejewski, P. Chabelski, and B. Szawulak, "Integrated system for the creation, simulation and analysis of the Petri nets," Poznan University of Technlogy, B.Sc. (eng.) thesis, 2013.

[5] B. Szawulak, "Integrated system for creation, simulation and analysis of timed Petri nets," Poznan University of Technlogy, M.Sc. thesis, 2014.

[6] M. Heiner, M. Herajy, F. Liu, C. Rohr, and M. Schwarick, "Snoopy A Unifying Petri Net Tool," Lecture Notes in Computer Science, vol. 7347, 2012, pp. 398–407.

[7] "Integrated Net Analyzer," 2001, URL: http://www2.informatik.huberlin.de/ starke/ina.html [retrieved: 2015-03-02].

 ISBN: 978-1-61208-409-1