Multiple Sclerosis disease: a computational approach for investigating its drug interactions.

Simone Pernice[1*], Marco Beccuti[1], Greta Romano[1], Marzio Pennisi[2], Alessandro Maglione[3], Santina Cutrupi[3], Francesco Pappalardo[4], Lorenzo Capra[5], Giuliana Franceschinis[2], Massimiliano De Pierro[2], Gianfranco Balbo[1], Francesca Cordero[1], Raffaele Calogero[6]

(1) Dept. of Computer Science, University of Turin, Turin, Italy, *pernice@di.unito.it
(2) Computer Science Inst., DiSIT, University of Eastern Piedmont, Alessandria, Italy
(3) Dept. of Clinical and Biological Sciences, University of Turin, Orbassano, Italy
(4) Dept. of Drug Sciences, University of Catania, Catania, Italy
(5) Dept. of Computer Science, University of Milan, Milan, Italy
(6) Dept. of Molecular Biotechnology and Health Sciences, University of Turin, Turin, Italy

Abstract. Multiple Sclerosis (MS) is a chronic and potentially highly disabling disease that can cause permanent damage and deterioration of the central nervous system. In Europe it is the leading cause of non-traumatic disabilities in young adults, since more than 700,000 EU people suffer from MS. Although recent studies on MS pathophysiology have been performed, providing interesting results, MS remains a challenging disease. In this context, thanks to recent advances in software and hardware technologies, computational models and computer simulations are becoming appealing research tools to support scientists in the study of such disease. Motivated by this consideration, we propose in this paper a new model to study the evolution of MS in silico, and the effects of the administration of the daclizumab drug, taking into account also spatiality and temporality of the involved phenomena. Moreover, we show how the intrinsic symmetries of the model we have developed can be exploited to drastically reduce the complexity of its analysis.

Keywords: Multiple sclerosis; Computational model; Colored Petri Nets

1 Introduction

Multiple Sclerosis (MS) is a long-term and autoimmune disease of the Central Nervous System (CNS). During the progression of the dis-

ease, cells of immune system attack the principal components of the CNS, the neurons, removing the enveloping myelin and preventing the efficient transmission of the nervous signals. Relapsing-Remitting MS (RRMS) is the predominant type of MS since it is diagnosed in about $85\% - 90\%$ of MS cases [12]. In RRMS, the disease alternates two phases: (1) relapse phase is characterized by a disease worsening due to the active inflammation damaging the neurons; (2) in the remission phase there is a complete or partial lack of the symptoms [7]. Recently, many treatments were proposed and studied to contrast the RRMS progression. Among these drugs, daclizumab [4] (commercial name Zynbrita), an antibody tailored against the Interleukin-2 receptor (IL2R) of T cells, exhibited promising results. Unfortunately, its efficacy was accompanied by an increased frequency of serious adverse events as infections, encephalitis, and liver damages. For these reasons daclizumab has been withdrawn from the market worldwide.

In [11] we proposed a model to investigate the effect of the daclizumab administration in RRMS. It involves the following seven main actors of MS: Epstain-Barr virus (EBV), Effector T lymphocytes cells (Teff), Regulatory T lymphocytes cells (Treg), Natural Killer cells (NK), Oligodentrocytes cells (ODC), Interleukin-2 (IL2) and daclizumab (DAC). In details, the EBV was considered since several studies [13] commonly agree on the hypothesis that viruses may play a role in RRMS pathogenesis acting as environmental triggers, and in particular the presence of this virus represents a well established risk factor in MS [8]. Effector T cells (Teff) are instead immune cells with a protective role against pathogens in healthy people. However, in RRMS a hypothesis is that the reactivation of EBV latent infection could bring to the activation of autologous Teff lymphocytes against myelin, due to a structure similarity between one viral protein and myelin protein (molecular mimicry). Regulatory T cells (Treg) are immune cells acting as balancing of the immune response since they contribute to suppress and modulate the Teff cells activity when no longer needed, or when there is a high risk of inflammation that can cause injuries to the tissues of the host. Other important actors within this context are the natural killer (NK) cells, a family of immune cells that acts as host-rejection of infected cells. Oligodendrocytes (ODC) are instead cells supporting

the neurons since they produce and are able to partially restore the myelin around the neurons whenever a not excessive damage occurs. IL2 is an immunomodulatory cytokine released by Teff in order to self-stimulate to duplicate and to propagate their immune actions. Finally, we included in the model the drug daclizumab, a humanized monoclonal antibody used in MS as drug against the Interleukin-2 receptor (IL2R) that is able to break the autoimmune reaction by suppressing the immune cells proliferation [4].

Thus, to help scientists in improving their knowledge of these phenomena, in this work we extend the RRMS models presented in [11] considering the cells movement into a three-dimensional grid. In details, in this paper we show how the use of a graphical formalism, i.e. the Extended Stochastic Symmetric Net (ESSN) formalism [10, 11], allows to easily deal with this complex three-dimensional model whose direct definition in terms of ODE system becomes clearly unfeasible even for a small three-dimensional grid. Indeed, for instance considering three-dimensional grid with dimension $3 \times 3 \times 3$ the ESSN model is a bipartite graph with only 38 nodes (i.e. 13 places and 25 transitions) and approximately 90 arcs, while its underlying deterministic process comprises 433 ODEs. Moreover, the high level of parametrization and flexibility provided in the model through this graphical formalism enables to study different grid dimensions in a fairly easy manner and without the need of redrawing the whole model. Similarly, in the analysis phase the ESSN model provides a powerful methodology that automatically exploits the system symmetries to reduce the complexity (in terms of number of equations) of the underlying deterministic process. Indeed in [2] we proposed an algorithm that directly derives a *compact* ODE system from a ESSN model in a symbolic way, through algebraic manipulation of ESSN annotations.

2 Scientific background

In this section we introduce the Petri Nets (PNs) formalism used to describe our model. PNs and their extensions are effective formalisms to model biological systems thanks to their capability of representing in a simple and clear manner the system features and

to provide efficient techniques to derive system qualitative and quantitative properties. In details, PNs are bipartite directed graphs with two types of nodes called *places* and *transitions*. Places, graphically represented as circles, correspond to the state variables of the system, while transitions, graphically represented as boxes, correspond to the events that can induce a state change. The arcs connecting places to transitions (and vice versa) express the relations between states and event occurrences. Places can contain tokens, drawn as black dots. The state of a PN, namely a *marking*, is defined by the number of tokens in each place. The system evolution is provided by the firing of an enabled transition, where a transition is enabled if and only if each input place contains a number of tokens greater than or equal to a given threshold defined by the cardinality of the corresponding input arc. The firing of an enabled transition removes a fixed number of tokens from its input places and adds a fixed number of tokens into its output places (according to the cardinality of its input/output arcs).

In this work we focus on Stochastic Symmetric Nets (SSNs) a high level formalism that extends PNs with *colors* and *stochastic firing delays* [6]. Colors provide a more compact, readable and parametric representation of the system thanks to the possibility of having tokens with different characteristics.

More specifically, the *color domain* associated with place p, denoted $cd(p)$, specifies the color of the tokens contained in this place, whereas the color domain of a transition defines the different ways of firing it (i.e. the possible *transition instances*). In order to specify these firings, a color function is attached to each arc which, given a color of the transition connected to the arc, determines the number of colored tokens that will be added to or removed from the corresponding place. A color domain is defined as Cartesian product of *color classes* which may be viewed as primitive domains. A color class can be partitioned into *static subclasses*. The colors of a class have the same nature (e.g. T cells), whereas the colors inside a static subclass have the same potential behavior (e.g. Teff). Stochastic firing delays, sampled from negative exponential distributions, allow to automatically derive the underlying Continuous Time Markov Chain (CTMC) that can be studied to quantitatively evaluate the system behaviour. In the literature, different techniques

are proposed to solve the underlying CTMC; in particular, in case of very complex models, the so-called deterministic approach [9] can be efficiently exploited. According to this, in [3] we proposed a method to derive a deterministic process, described through a system of Ordinary Differential Equations (ODEs), which well approximates the stochastic behavior of an SSN model assuming all reactions follow the Mass Action (MA) law. In the same paper we also described an efficient translation method based on the SSN formalism, which is able to reduce the size (in terms of equations number) of the underlying ODE system through the automatic exploitation of system symmetries. Practically, the complete set of ODEs, which can be derived from an SSN model is partitioned into equivalence classes of ODEs which have same solution so that a representative equation, called *symbolic equation*, can be pointed out for each equivalence class. Then, a reduced ODE system may be derived including only these symbolic equations whose solution mimics the behavior of the original model. Recently this result was further improved in [2] where a new algorithm is discussed to generate the symbolic equation for each equivalence class of ODEs without deriving the complete ODE system. This is achieved thanks to a recent extension of a symbolic calculus for the computation of SSN structural properties [5].
Furthermore, in [11] we introduced the Extended SSNs (ESSNs) to deal with more complex biological laws splitting the set T of all the transitions of the model into two subsets: T_{ma} and T_g. The former subset contains transitions (that are called *standard*) whose rates are specified as MA laws. The latter includes instead all the transitions (that are called *general*) whose random firing times have rates that are defined by means of general real functions. In our definition, we assumed that the general function associated with a transition $t \in T_g$ is a real function which depends only on time and on the input places of t. So, if $x_{p,c}(\nu)$ represents the average number of tokens of color c in the place p at time ν, then the rate at which the instance $\langle t, c, c' \rangle$, $t \in T_g$ will move tokens with color c in place $x_{p,c'}(\nu)$ is given by $f_{\langle t,c,c' \rangle}(\hat{x}(\nu), \nu)$, where $\hat{x}(\nu)$ is the vector characterized by the average number of tokens of the input places of transition t.

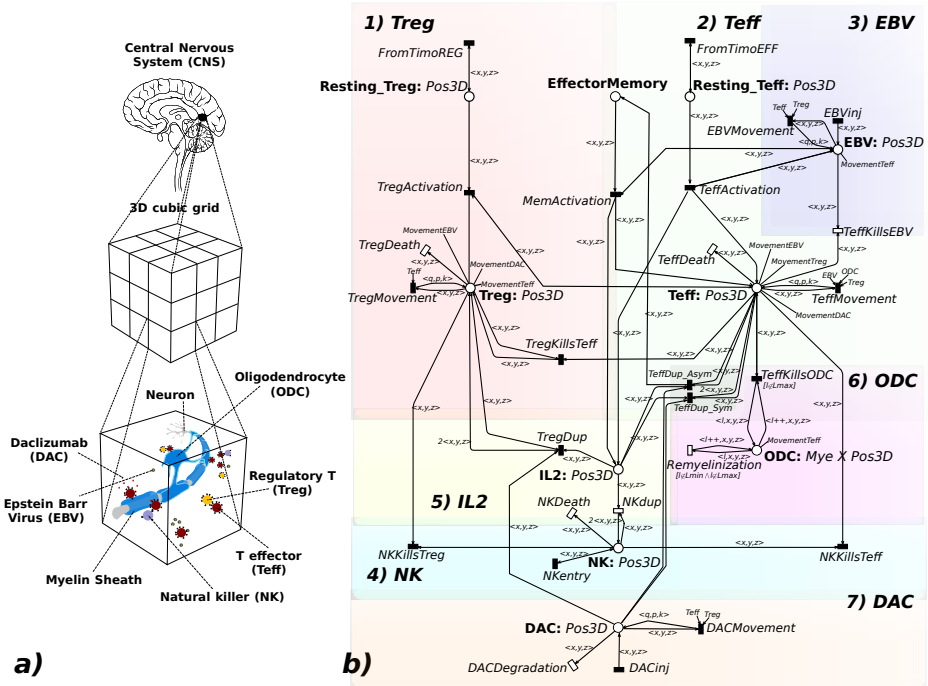

Fig. 1: a) Representation of the three-dimensional model. b) The ESSN model.

3 Materials and Methods

In this section, we report our extension of the Relapsing-Remitting Multiple Sclerosis (RRMS) model defined with the ESSN formalism and originally presented in [11] assuming that the cells may move within a cubic grid. In Fig. 1a) is depicted a portion of the CNS, showing: the neuron with its myelin sheath, and the 7 elements characterizing the MS disease distributed within a 3D cubic grid. The respective ESSN model is shown in Fig. 1b), consisting of 13 places and 25 transitions. For the sake of clarity, the white transitions are standard transitions, while the black ones are general transitions. This model is organized in seven modules corresponding to the biological entities characterizing RRMS. Briefly, the EBV module simulates the virus reactivation by means of a series of injections of virus particles in the system at given times, while the Treg and Teff modules encode the activation of the T cells, the an-

nihilation of the virus by the Teff action, the control mechanism of the Treg over the Teff. The NK module describes the killing of self-reactive Teff and Treg cells respectively, due to NK cells. The IL2 module is focused on the IL2 role. IL2 is consumed by the Treg, Teff and NK functions and it is produced by the Teff activation. The ODC module describes instead the ODC behaviour, characterized particularly by the damage caused by Teff cells on ODC cells. Indeed, when the myelin level reaches the lowest value, an irreversible damage occurs and a remyelination of the neurons is no more possible. Finally, the DAC module encodes the drug administration and its pharmacokinetics inhibition of the expansion of Treg and Teff.

The model is characterized by four color classes: $PosX$, $PosY$, and $PosZ$ representing the coordinates of the position of a molecule in a 3D cubic grid; Mye encoding the myelination levels of ODC. Mye is divided into five static subclasses ranging from $Lmin$ (no myelination) to $Lmax$ (full myelination). Then, all the places except the ODC and EffectorMemory are characterized by the color domain defined as $Pos3D = PosX \times PosY \times PosZ$, i.e. the three-dimensional Cartesian product of the three coordinates color classes. Instead, the ODC place is characterized by the three coordinates plus the myelination levels, so that its color domain is $Pos3D \times Mye$. Finally, the EffectorMemory place has neutral color domain. Moreover, we assume that the EBV, Teff, Treg and DAC cells are able to move in all the cubic cells of the grid. Practically, the EBVs move uniformly in all the cells, the Teff cells move with higher probability towards a location in which there is higher concentration of EBV, and Treg and DAC cells move with higher probability towards a location in which there is higher concentration of Teff cells. Hereafter, the notation of the color combinations $\langle p_x, p_y, p_z \rangle$ and $\langle q_x, q_y, q_z \rangle$, representing the location coordinates, is simplified to $\langle \mathbf{p} \rangle$ and $\langle \mathbf{q} \rangle$, respectively. In particular, we define $x_{CellType_{\langle \mathbf{p} \rangle}}$ as the number of $CellType$ in the location $\langle \mathbf{p} \rangle$ at a specific time point. Hence, the movement functions can be defined as follows. The transition ***EBVMovement*** simulates the movements of EBV cells from point (with coordinates represented by the color combination) $\langle \mathbf{p} \rangle$ to point $\langle \mathbf{q} \rangle$. The speed of this movement (the rate of transition EBVMovement) is uniform in all directions and is captured in the following formula by assuming

that the probability to move is equally distributed among all the grid cells.

$$f_{\langle EBVMovement,\mathbf{p},\mathbf{q}\rangle}(\hat{x}(\nu),\nu) = r_{moves} * p^{EBV}_{\langle\mathbf{q}\rangle} * x_{EBV_{\langle\mathbf{p}\rangle}}$$

where r_{moves} is a coefficient that we set equal to 0.1 in our numerical experiments. Differently, the transition ***TeffMovement*** simulates the movement of Teff cells from point $\langle\mathbf{p}\rangle$ to point $\langle\mathbf{q}\rangle$, and its speed is inversely related to the number of EBV cells in $\langle\mathbf{p}\rangle$ (since more the virus in $\langle\mathbf{p}\rangle$ less the Teff cells are tempted to leave the position) and depends on the number of EBV in $\langle\mathbf{q}\rangle$ (a greater number of EBV cells leads to a higher probability to move into that location). This is captured by the following formula

$$f_{\langle TeffMovement,\mathbf{p},\mathbf{q}\rangle}(\hat{x}(\nu),\nu) = r_{moves} * (exp(-\frac{x_{EBV_{\langle\mathbf{p}\rangle}}}{C_{EBV}})) * p^{Teff}_{\langle\mathbf{q}\rangle} * x_{Teff_{\langle\mathbf{p}\rangle}}$$

where r_{moves} is again set equal to 0.1; the second term of the function, defined as $exp(-\frac{x_{EBV_{\langle\mathbf{p}\rangle}}}{C_{EBV}})$, varies in the interval $[1, 0)$, simulating the decreasing of the movement velocity with respect to the number of EBV cells present in the starting point; $p^{Teff}_{\langle\mathbf{q}\rangle} = \frac{x_{EBV_{\langle\mathbf{q}\rangle}}}{EBV_{tot}}$ represents the probability to move in the cell with coordinates $\langle\mathbf{q}\rangle$ where EBV_{tot} is the total number of EBV in the grid at time ν; and C_{EBV} is an experimental constant that we set equal to 1000. All these quantities are functions of the time ν which is omitted in the formula to keep the notation simpler.
Transitions ***TregMovement*** and ***DACMovement*** represent the movements of the Treg and DAC cells (respectively) from point $\langle\mathbf{p}\rangle$ to point $\langle\mathbf{q}\rangle$. Similarly to what explained for transition TeffMovement, their speeds are inversely related to the number of Teff and T (= Treg+Teff) cells in $\langle\mathbf{p}\rangle$ and depend on the number of Teffs and Ts in $\langle\mathbf{q}\rangle$. The detailed expressions of the formulas that encode the firing rate dependencies for these two transitions, with the information regarding all the transitions and the files exploited thorough this study, are freely available at https://github.com/qBioTurin/ESSNandRRMS /tree/master/DeterministicModel/Multidimensional.

We exploited the GreatSPN tool [1] to develop the ESSN model Fig. 1b), in particular the corresponding system of ODEs is automatically generated from the ESSN model using the C/C++ mod-

ule PN2ODE embedded in GreatSPN. In details, the ODEs system is defined by one equation for each pair $\langle p, c\rangle$, where p is a place and $c \in cd(p)$ is a color which encodes a coordinate in the 3D grid, plus the myelination level (the latter only in place ODC). Differently, the system of SODE has been generated with the SNespression tool (http://di.unito.it/∼depierro/SNexpression) and integrated with the definition of the functions for the *general* transitions and the initial marking. In this case there is one equation for each pair $\langle p, \hat{c}\rangle$, where p is a place and $\hat{c}$ corresponds to a subset of $cd(p)$ with equivalent behavior. The SODE for a given $\hat{c}$ is representative of all ODEs in the *unfolded* system for all $c \in \hat{c}$. The identification of the equivalence classes in $cd(p)$ and of the generation of the representative SODE is completely automatized.

4 Results

In this work we studied the RRMS considering a tissue portion explicitly modeled through a cubic grid consisting of 27 cubic cells (Fig. 1a)). To achieve this, we defined the color classes $PosX = \{x_1, x_2, x_3\}$, $PosY = \{y_1, y_2, y_3\}$ and $PosZ = \{z_1, z_2, z_3\}$. For all the simulations, we assumed 500 ODC with level L_{max} of neuronal myelinization, 1687 resting Teff cells, 63 resting Treg cells, 375 NK cells and 1000 IL2 molecules, and zero cells in the other places (for more details see [11]).

This model is equivalent to a system of 433 ODEs, but with few assumptions it is possible to derive the corresponding reduced ODEs system including only the symbolic equations. In details, let us define the set of all the 27 location coordinates as $\mathbf{P} = \{\langle p_x, p_y, p_z\rangle, p_x \in PosX, p_y \in PosY, p_z \in PosZ\}$. Then, we consider three disjoint subsets of $\mathbf{P}$, namely $P1$, $P2$, $P3$; the first two correspond to grouping the EBV and DAC injection locations, respectively, while the third *P3* groups all the remaining locations. For simplicity, and to maintain the symmetries into the system as well, the EBV and DAC injection locations do not change over the simulation time and do not overlap. Given this, it is possible to derive the symbolic ODEs (SODEs) system characterized by 49 equations only. Indeed, the 433 original equations can be partitioned into 49 groups of similar equations. Each group is expressed in the reduced model by one repre-

sentative equation. The grouping derives from the observation that the behaviors of the modeled elements do not depend on their actual positions, but only on the presence of the EBV and/or DAC cells. When the grid size grows the number of groups does not change, as long as the number of locations where different quantities of EBV and/or DAC cells appear is fixed; instead the size of each group of equivalent ODEs increases with the grid size. A further reduction is represented by the number of terms in each SODE, representative of each group of ODEs, with respect to the number of terms appearing in the ODEs in the equivalence class. This reduction is due to the factorization obtained by exploiting symmetries. Other examples are reported in Table 1, where the R file dimension and the number of differential equations of the complete and reduced models are compared considering different cubic grid dimensions, from $3 \times 3 \times 3$ to $5 \times 5 \times 5$. It is easy to see that an increasing number of locations is associated with an increase in the number of ODEs and of the R file containing them, while the SODE system does not change. Note that when the $5 \times 5 \times 5$ grid is considered, the ODEs generation procedure fails because it exceeds the available memory.
The advantage can also be observed from the point of view of the simulation time, we obtained a speed up from 8.927205 hours to 12.76043 secs on an Intel Xeon processor @ 2GHz, by using one core. Note that the simulation was performed considering $3 \times 3 \times 3$ cubic grid, one year interval and assuming EBV injections at regular times (every two months), and each injection introduces into the system 10000 EBV copies.

Number of locations	R File dimension ODEs / SODEs	Number of ODEs / SODEs
27 (3x3x3)	0.43 MiB / 0.023 MiB	433 / 49
64 (4x4x4)	5.0 MiB / 0.023 MiB	1025 / 49
125 (5x5x5)	Out of memory / 0.023 MiB	2001 / 49

Table 1: Comparing the ODE and SODE system, varying the cubic grid dimension.

A possible evolution of the system is shown in Fig. 2, where the red circles represent the location in which EBV is injected. For each plot, the three rows represent the z-planes and the columns refer to the time points in which the injections are done. Fixing the time point and the z-plane, the corresponding 3×3 square reports the number of ODCs damaged into the nine grid cells obtained varying

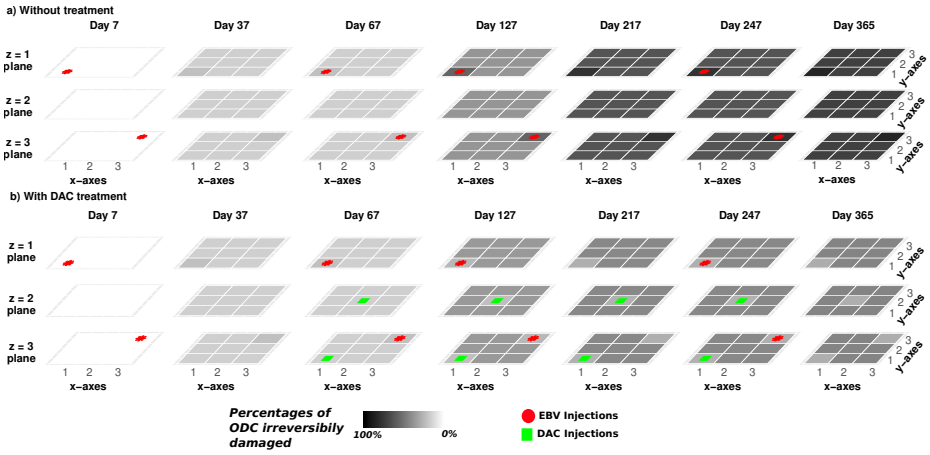

Fig. 2: Percentages of ODCs irreversibly damaged a) without and b) with DAC treatment.

the x and y coordinates. As expected, the panel A of Fig. 2 shows the progressive accumulation of ODC irreversibly damaged until day 365. Instead, in panel B of Fig. 2 is reported the results of the simulation of the DAC effect. In details, every month after two months of simulation, two injections are simulated (green squares) introducing 300 DAC copies for each administration. These results agree with those proposed in [11] since the number of irreversibly damaged ODCs decreases in the case with DAC administration with respect to the case in which no drug is injected. With DAC the percentage of irreversibly damaged ODCs ranges from 28% to 45%, while with no DAC the number of irreversibly damaged ODCs is between 70% and 85%.

5 Conclusion

In this work we extended the model presented in [11] including the spatial coordinates of all entities in a cubic tissue portion. This gives the opportunity to model more realistic scenarios, where different quantities of virus enter into the system from different directions.

Moreover, we described how the intrinsic symmetries of the derived ESSN model may be automatically exploited to reduce the complexity of the analysis step. This allows us to study models which are independent from the grid size, while with the classical approach

it is hard to generate the ODEs system corresponding to the model with a $5 \times 5 \times 5$ grid.
As further work, we will focus our experiments on the dosage of DAC and on the representation of the DAC pharmacokinetics in order to simulate the up taking of DAC by the body, its biotransformation and the distribution of DAC and its metabolites in the tissues.

References

1. Amparore, E.G., Balbo, G., Beccuti, M., Donatelli, S., Franceschinis, G.: 30 years of GreatSPN. In: Principles of Performance and Reliability Modeling and Evaluation, pp. 227–254. Springer (2016)
2. Beccuti, M., Capra, L., De Pierro, M., Franceschinis, G., Follia, L., Pernice, S.: A tool for the automatic derivation of symbolic ODE from symmetric net models. In: Proc. of 27th IEEE Int. Symp. on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, MASCOTS 2019, Rennes, France, October 21-25, 2019. pp. 36–48 (2019)
3. Beccuti, M., Fornari, C., Franceschinis, G., Halawani, S., Ba-Rukab, O., Ahmad, A., Balbo, G.: From symmetric nets to differential equations exploiting model symmetries. Comput. J. 58(1), 23–39 (2015)
4. Bielekova, B.: Daclizumab Therapy for Multiple Sclerosis. Cold Spring Harb Perspect Med 9(5) (May 2019)
5. Capra, L., De Pierro, M., Franceschinis, G.: Computing structural properties of symmetric nets. In: Proc. of 12th Int. Conf. on Quantitative Evaluation of Systems, QEST 2015, Madrid, Spain, September 1-3, 2015. LNCS, vol. 9259, pp. 125–140. Springer (2015)
6. Chiola, G., Dutheillet, C., Franceschinis, G., Haddad, S.: Stochastic well-formed coloured nets for symmetric modelling applications. IEEE Tran. Comput. 42(11), 1343–1360 (1993)
7. Dutta, R., Trapp, B.: Mechanisms of Neuronal Dysfunction and Degeneration in Multiple Sclerosis. Prog. Neurobiol. 93(1), 1–12 (2011)
8. Guan, Y., Jakimovski, D., Ramanathan, M., Weinstock-Guttman, B., Zivadinov, R.: The role of Epstein-Barr virus in multiple sclerosis: from molecular pathophysiology to in vivo imaging. Neural Regen Res 14(3), 373–386 (Mar 2019)
9. Kurtz, T.G.: Solutions of ordinary differential equations as limits of pure jump Markov processes. J. Appl. Probab. 1(7), 49–58 (1970)
10. Pernice, S., Follia, L., Balbo, G., Milanesi, L., Sartini, G., Totis, N., Lió, P., Merelli, I., Cordero, F., Beccuti, M.: Integrating Petri nets and flux balance methods in computational biology models: A methodological and computational practice. Fundamenta Informaticae 171(1-4), 367–392 (2019)
11. Pernice, S., Pennisi, M., Romano, G., Maglione, A., Cutrupi, S., Pappalardo, F., Balbo, G., Beccuti, M., Cordero, F., Calogero, R.: A computational approach based on the colored Petri net formalism for studying multiple sclerosis. BMC Bioinformatics 20 (2019)
12. Sospedra, M., Martin, R.: Immunology of Multiple Sclerosis. Semin Neurol 36(2), 115–127 (Apr 2016)
13. Virtanen, J., Jacobson, S.: Viruses and Multiple Sclerosis. CNS Neurol Disord Drug Targets 11(5), 528–544 (2012)