High-level modeling and verification of cellular signaling

Natasa Miskov-Zivanov
University of Pittsburgh
Electrical and Computer Engineering
Bioengineering
Computational and Systems Biology
Pittsburgh, PA, USA
nmzivanov@pitt.edu

Paolo Zuliani
Newcastle University
Computer Science
Newcastle upon Tyne, UK
paolo.zuliani@ncl.ac.uk

Qinsi Wang, Edmund M. Clarke
Carnegie Mellon University
Computer Science
Pittsburgh, PA, USA
{qinsiw,emc)@cs.cmu.edu

James R. Faeder
University of Pittsburgh
Computational and Systems Biology
Pittsburgh, PA, USA
faeder@pitt.edu

***Abstract* — We use computational modeling and formal analysis techniques to study temporal behavior of a discrete logical model of the naïve T cell differentiation. The model is analyzed formally and automatically by performing temporal logic queries via statistical model checking. While the model can be verified and then further explored using Monte Carlo simulations, model checking allows for much more efficient analysis by testing a large set of system properties, with much smaller runtime than the one required by simulations. The results obtained using model checking provide details about relative timing of events in the system, which would otherwise be very cumbersome and time consuming to obtain through simulations only. We efficiently test a large number of properties, and confirm or reject hypotheses that were drawn from previous analysis of experimental and simulation data.**

Keywords— formal methods, Immune system, Boolean networks, T cell, Statistical model checking, Stochastic simulation.

I. Introduction

In this work, we apply temporal logic model checking to verify model correctness and explain system behavior for signaling network that controls T cell differentiation. The model used in this work couples exogenous signaling inputs to T cell phenotype decisions. The goal of our study is to identify key factors and pathways that contribute to the discrimination of the T-cell receptor (TCR) signal strength (*i.e.*, antigen dose or duration presented to TCR) by the differentiating T cell. Different T cell phenotype ratios play an important role in T-cell mediated immunity, in both autoimmune diseases and in cancer. The two primary phenotypes we consider are: 1) regulatory (Treg) cells that express the transcription factor Foxp3 but do not express the cytokine IL-2; 2) helper (Th) cells that do not express Foxp3 but do express and secrete IL-2. Control of the Treg vs. Th cell phenotype induction in periphery from naïve T cells is a promising approach to either eliminate antigen-specific Treg cells and decrease (or even reverse) immune suppression in cancer, or enhance Treg induction to prevent autoimmune diseases. Previous studies have indicated that the timing of T cell stimulation, antigen dose and the duration of antigen stimulation, strongly influence the T cell phenotype choice [5][8].

The model we are studying, adopted from [5], was developed using a discrete, logical modeling approach, and simulated using a random asynchronous approach with the BooleanNet tool [1]. Experimental observations from [8] that the induction/expansion of Foxp3+ Treg cells by low dose antigen is inversely correlated with the levels of signaling in the mTOR pathway suggest a complex interaction between cell surface receptors, signaling molecules and important transcription factors. The model in [5] captures critical signaling events, from stimulatory signals at receptors, through activation of transcription factors, to production of proteins representing different phenotypes. Figure 1 presents a simplified interaction network highlighting all system elements whose temporal behavior is studied in this work.

Model simulations described in [5] allow for recapitulating a number of experimental observations and provide new insights into the system. However, to further verify model properties and to test new properties on the model, it is usually necessary to write new parts of the simulator code, or to manually analyze a significant amount of simulation data. This approach quickly becomes tedious and error-prone. Thus, to study the system in Figure 1, we apply computational modeling approaches together with formal methods (Figure 2).

Since the underlying semantic model of the simulation tool is essentially a discrete-time Markov chain, we need to verify probabilistic (stochastic) models. The verification problem for stochastic systems amounts to computing the probability that a given temporal logic formula is satisfied by the system. One approach is to use precise numerical methods to compute exactly the probability that the formula is true. However, these methods suffer from the state explosion problem, and do not scale well to large-scale systems. Statistical model checking can be effectively used for verifying temporal logic

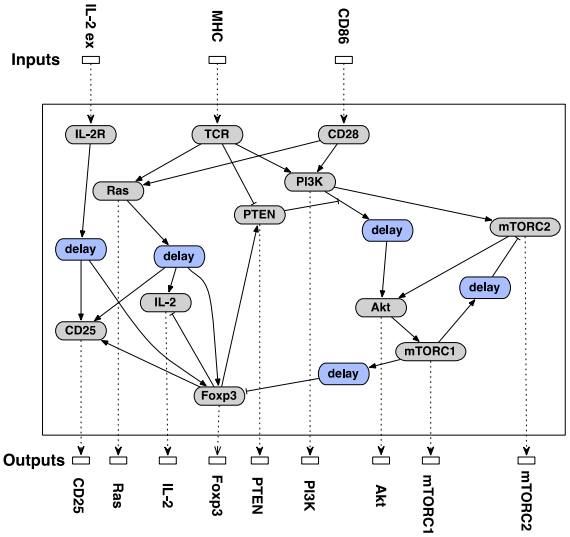


Figure 1. T cell model: inputs (antigen and co-stimulatory signal, cytokine IL-2 exogenous), interactions between key elements, and outputs of the model that are studied in this work (CD25, Ras, IL-2, Foxp3, PTEN, PI3K, Akt, mTORC1, and mTORC2).

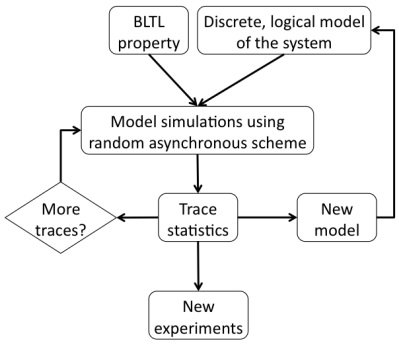


Figure 2. Analysis framework with statistical model checking: model of the system and a set of properties are used as inputs to the framework; model is simulated using random asynchronous scheme.

specifications for systems affected by the state explosion problem [11]. The technique relies on simulation, thereby avoiding a full state space search. This implies that the answer to the verification problem (*i.e.*, the probability that the property holds) is only approximate, but its accuracy can be arbitrarily constrained by the user. In return, statistical model checking is more scalable and hence more useful for large models.

1) The remainder of the paper is organized following the steps in our framework, outlined in Figure 2. In Section II, we describe the preliminaries of the discrete modeling approach that was used to develop the model in [5], as well as the characteristics of the system, the T cell differentiation control network. In Section III, we provide brief description of temporal logic, rules for creating BLTL properties, and how we apply them to the system that we are studying. In Section IV, we outline advantages and drawbacks of model simulations, present the motivation for using model checking, and describe how we apply statistical model checking. The results of applying our approach to study the T cell differentiation are presented in Section V. We conclude the paper with Section VI.

II. Network model

The model in [5] represents a generalization of the Boolean network model, and allows for modeling elements of the network as discrete (not only Boolean) variables. The model consists of 38 elements, and includes ligands outside of the cell, receptors, signaling molecules inside the cell, transcription factors, and several genes of interest. As shown with simplified interaction map in Figure 1, three signals serve as inputs to the model, and their level of activity or duration of presence defines the direction that naïve T cell takes in its response:

- Major Histocompatibility Complex (MHC), a molecule presented by antigen presenting cells (APC); MHC represents a major signal that activates T cell receptor;
- Cluster of Differentiation 86 (CD86), a molecule that binds to CD28 receptor and functions as a co-stimulatory signal;
- Exogenous Interleukin 2 (IL-2 ex).

While most of the variables in the model take values from the set {0,1}, there are several variables that can take values from the set {0,1,2}. In order to create the circuit model that includes AND, OR and Inverter logic gates, the three-valued variables are encoded using two Boolean variables. For example, T cell receptor, TCR, and one of the internal cell signaling molecules, PI3K, are each modeled using two variables, TCRlow and TCRhigh for TCR, and PI3Klow and PI3Khigh for PI3K. While TCRlow and TCRhigh are assigned values directly from the input variable MHC, PI3Klow and PI3Khigh have separate logic rules that are used to compute their next value:

PI3Klow_next = (TCRlow and CD28) or (PI3Klow and IL2_ex and IL2R)

PI3Khigh_next = (TCRhigh and CD28) or (PI3Klow and IL2_EX and IL2R).

Furthermore, delays are included in the model, through inserting Buffers (additional logic gate type used) that just propagate inputs to outputs with a delay. Although many of the elements of the model are important as indicators of the system state, we focus in our studies on the following elements:

- Forkhead box P3 Foxp3 (Foxp3), marker for Treg cells;
- CD25, alpha chain of the IL-2 receptor (IL-2Rα), expressed on the surface of the cell, after the cell is stimulated through the TCR;
- Interleukin 2 (IL-2), expressed by the cell and secreted outside of the cell, marker for Th cells;
- Mammalian Target of Rapamycin Complexes 1 and 2 (mTORC1 and mTORC2), major inhibitors of Foxp3;
- Phosphatase and Tensin Homolog (PTEN), expressed in naïve T cells as well as in Treg cells, found at low levels in Th cells.

III. Temporal logic

We encode relevant properties of the model as temporal logic formulae, which are then verified via statistical model checking. We use Bounded Linear Temporal Logic (BLTL) [4] as our specification language. BLTL is a variant of the well-known Linear Temporal Logic (LTL) [6] in which the temporal operators are endowed with a time bound. For example, a BLTL formula expressing the specification "it is not the case that in the **F**uture 10 time steps CD25 is **G**lobally activated (i.e., it equals 1) for 17 time steps" is

$$F[10]\ G[17]\ (CD25 = 1)$$

where the F[10] operator encodes "future 10 time steps", G[17] expresses "globally for 17 time steps", and CD25 is a state variable of the model. The syntax of BLTL is given by:

$$\varphi ::= y \sim v \mid (\varphi 1 \vee \varphi 2) \mid (\varphi 1 \wedge \varphi 2) \mid \neg\varphi \mid (\varphi 1\ U[t]\ \varphi 2)$$

where $\sim \in \{\geq, \leq, =\}$, $y \in SV$ (the finite set of state variables), $v \in Q$, and $t \in Q_{\geq 0}$. Formulae of the type $y \sim v$ are called atomic propositions (AP). The formula $\varphi 1\ U[t]\ \varphi 2$ holds true if and only if, within time t, $\varphi 2$ will be true and $\varphi 1$ will hold until then. Note that the operators F[t] and G[t] referenced above are easily defined in terms of the until operator: $F[t]\ \psi = True\ U[t]\ \psi$ requires ψ to hold true within time t (*True* is the atomic proposition identically true); $G[t]\ \psi = \neg F[t]\ \neg\psi$ requires ψ to hold true up to time t.

The semantics of BLTL is defined with respect to executions of the Boolean model. An execution (or trace) is a sequence $\sigma = (s_0, t_0)\ (s_1, t_1)\ \ldots$ where the s_i's are states of the model and the t_i's represent time lengths. A pair (s_i, t_i) denotes that the system transitioned to state s_{i+1} after having sojourned for t_i time units in state s_i.

In general, BLTL is defined over rational time bounds. However, in our Boolean network model time does actually belong to the (positive) integers. More specifically, in our executions all the time durations t_i's are just 1. This means that the time bound of a formula specifies exactly the length of a model simulation (*i.e.*, the number of transitions taken in the Markov chain).

IV. Network analysis methods

In the following, we describe two methods that we use for studying the developed model. First, we present the simulation approach and discuss data types that can be obtained using

simulations. Next, we describe details of statistical model checking, and how we combine simulations and model checking to gain insights about the system efficiently.

A. Model simulations

When studying the model using simulations, we apply an asynchronous simulation scheme that assumes execution of all update rules in a random order within a single round of updates. Initial values of all elements are chosen at the beginning, and in each round the update order of rules is chosen randomly. The simulation is run until the steady state is reached. Such update scheme allows for a single state to have multiple next states, and for a given initial state, each simulation results in one trajectory out of a (most often) large number of possible trajectories. The number of trajectories that can result from a

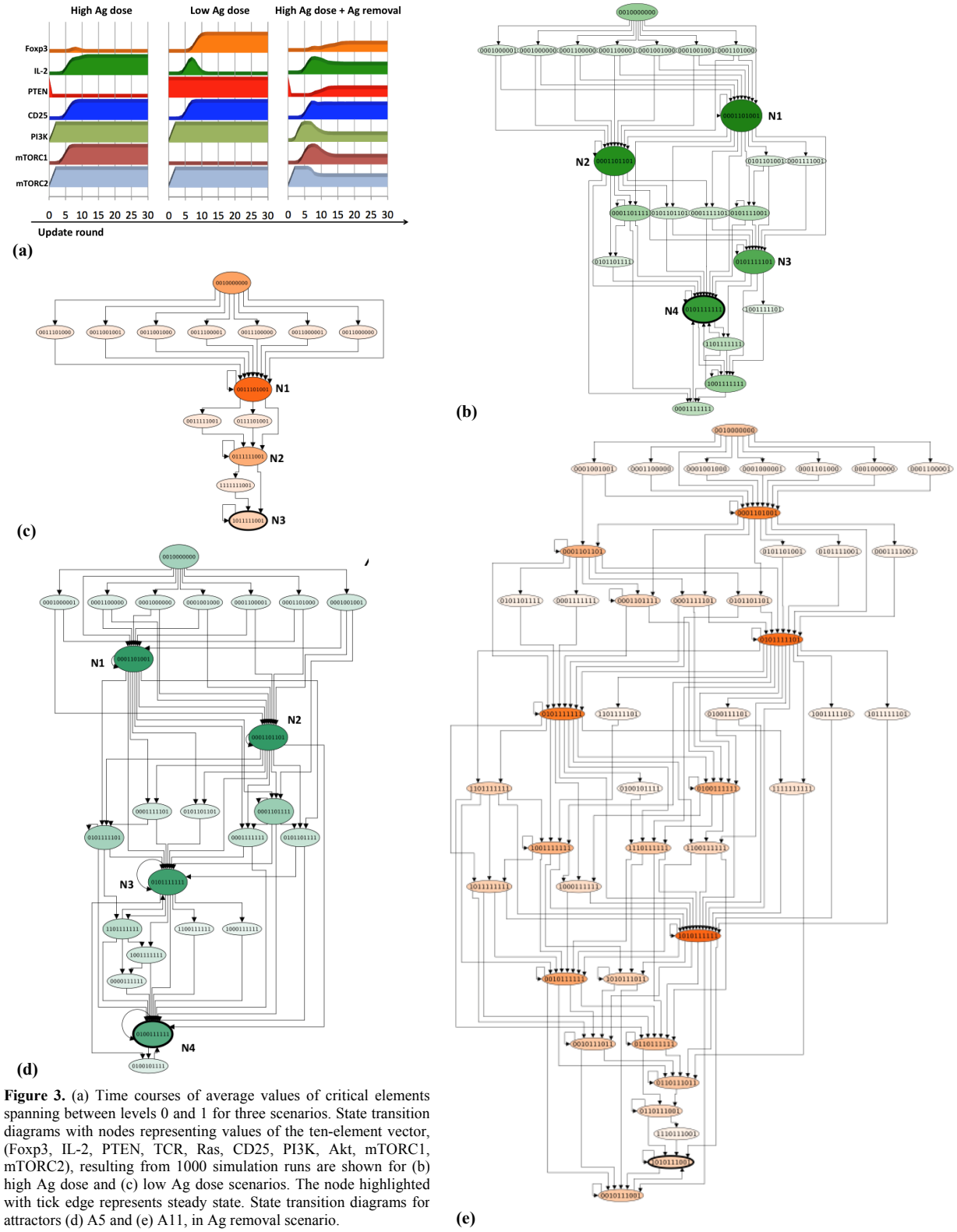


Figure 3. (a) Time courses of average values of critical elements spanning between levels 0 and 1 for three scenarios. State transition diagrams with nodes representing values of the ten-element vector, (Foxp3, IL-2, PTEN, TCR, Ras, CD25, PI3K, Akt, mTORC1, mTORC2), resulting from 1000 simulation runs are shown for (b) high Ag dose and (c) low Ag dose scenarios. The node highlighted with tick edge represents steady state. State transition diagrams for attractors (d) A5 and (e) A11, in Ag removal scenario.

single initial condition largely depends on the topology of the network and update rules.

Several scenarios with different initial conditions are of special interest for immunologists:

1. Stimulation of TCR with high level of MHC – experiments show that such stimulations result in differentiated Th cells (referred to as High antigen (Ag) dose scenario);
2. Stimulation of TCR with low level of MHC – experiments show that such stimulations result in a mixed population with a significant number of Treg cells (Low Ag dose scenario);
3. Stimulation of TCR with high level of MHC, and removal of MHC after 18 hours – experiments show that, similar to stimulations with low level of MHC, such stimulations result in a mixed population with a significant number of Treg cells (High Ag dose + Ag removal scenario).

We implemented these scenarios by setting different initial states and analyzing trajectories from initial to steady state. Steady state in which Foxp3 is at level 1 is used as a marker of differentiated Treg cells, and steady state in which IL-2 is at level 1 is used as a marker of differentiated Th cells.

Although simulations can provide an accurate estimate of average trajectory that the system transients through from initial to steady state, it is important to have means to also analyze individual simulation trajectories in an efficient manner. Often a number of queries about the system are created only after already analyzing simulation results, and these new queries require either tedious data analysis, or conducting new simulations. Moreover, model simulations apply Monte Carlo approach, and do not always find all reachable states. In some systems, like T cell differentiation, besides revealing states through which most of the trajectories transition, uncovering states that lead to trajectory divergence is crucial. Such insights can help us find factors underlying occurrence of different cell phenotypes. Only after obtaining these results we can come up with additional questions that tackle important characteristics of the system and its behavior.

To this end, formal methods such as model checking, allow for querying the system very efficiently, and for quickly obtaining information about transitions that are allowed (or not allowed) by the model configuration, as well as for understanding ordering of events and causal relationships.

B. *Statistical model checking*

Statistical model checking treats the verification problem for stochastic systems as a statistical inference problem, using randomized sampling to generate traces (or simulations) from the system model, then using model checking methods and statistical analysis on those traces. For a (closed) stochastic system and a BLTL property φ, the probability p that the system satisfies φ is well defined (but unknown in general). Now, we ask two questions: for a fixed $0<\theta<1$ we may ask whether $p\leq\theta$, or we may just want to know the value of p. In the statistical model checking approach, the first question is solved via hypothesis testing methods, while the second via estimation techniques. In particular, hypothesis tests can be thought of as probabilistic decision procedures, i.e., algorithms with a yes/no reply, and which may give wrong answers. Estimation techniques instead compute (probabilistic) approximations of the unknown probability p.

V. Results

In this section we compare the results obtained using statistical model checking with results obtained using simulations. We also describe system properties that were tested using model checking only.

A. *Framework implementation*

We obtain model simulation traces using the BooleanNet [1] simulation tool. We have combined BooleanNet with a parallel statistical model checker, such that verification of BLTL properties of logical models can be performed efficiently and automatically on a multi-core system. In particular, each core runs a BooleanNet simulation and checks its trace with respect to the BLTL property. A designated core carries out the required statistical computations. We have used the OpenMP API for shared-memory parallel programming in C++.

We have implemented the following hypothesis tests:

- Bayes Factor sequential hypothesis test [4][12];
- Sequential Probability Ratio Test [10];
- Bayes Factor sequential hypothesis test with indifference region.

With respect to estimation techniques, we have implemented:

- Bayesian sequential estimation [12];
- Chernoff-Hoeffding bound [3].

When running multiple Monte Carlo simulations in parallel, one has to make sure: a) to avoid any bias in the collection of the simulations' result, and b) that the simulations are actually independent, since most statistical techniques assume independency of the samples.

Unbiasedness can be enforced by using barrier synchronization after each simulation and trace checking. In particular, on a system with N cores, our tool launches N parallel BooleanNet simulations of the model and trace checking. After the barrier synchronization, the core with id 0 computes on all the samples obtained so far the tests requested by the user. Note that multiple tests can be specified, so the entire process stops when all the tests are done.

B. *Analysis of fixed stimulation outcome*

Model simulation results for the three simulation scenarios listed above were obtained using BooleanNet [5], and are shown in Figure 3(a). These results present the behavior of key elements in the model averaged across 1000 simulation trajectories. We created state transition graphs (STGs) from simulated state trajectories. In the following, we describe the hypotheses about important system properties that we draw from STGs, and results of testing these hypothesis using model checking.

1) *High stimulation level*

When naïve T cells are stimulated with high Ag dose, they differentiate into Th cells and have high levels of IL-2. Although differentiated Th cells do not express Foxp3, there is a brief transient period shortly after stimulation during which some cells express Foxp3 Figure 3(a)(left) [5]. Averaged simulation results show that the peak magnitude of Foxp3 transient is approximately 10%. However, it is not clear from these averaged time courses, whether this means that 10% of all trajectories exhibit such Foxp3 behavior, or this occurs in most of the trajectories, but at different rounds, thus leading to lower average peak.

Figure 3(b) outlines the STG created from 1000 simulation runs for the high-stimulation scenario. Each node in this STG represents a value of the ten-element vector, (Foxp3, IL-2, PTEN, TCR, Ras, CD25, PI3K, Akt, mTORC1, mTORC2). Most frequent nodes observed in simulation trajectories are highlighted in the STG and marked with names N1-N4. As can be seen from the STG in Figure 3(b), 1000 simulation trajectories are relatively uniform, with overall 23 different nodes and 71 edges. Furthermore, all trajectories lead to the same steady state. The same conclusion can also be drawn from

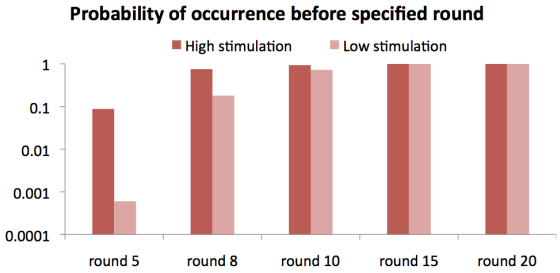


Figure 4. Probability of occurrence of High-stimulation and Low-stimulation attractors by round 5,8,10,15, or 20, estimated using statistical model checker.

average element time courses in Figure 3(a)(left), since each element has either average value 0 or value 1 in steady state.

In order to further investigate whether there are additional transient or steady states that can be reached in this scenario, we use model checking. We also use model checking to get more specific information about the transient behavior of Foxp3.

We define properties P1-P4 to test the probability of Foxp3 occurrence in case of high stimulation scenario. As our results show (Table I, P1), the probability that Foxp3 transiently increases to level 1 is about 24%, confirming our hypothesis that more than 10% of trajectories exhibit this transient Foxp3 behavior. We also see from properties P2-P3 that the transient is more probable to last only one round. Finally, the tests (P4) show that the high stimulation scenario indeed leads to a unique steady state.

2) Low stimulation level

When naive T cells are stimulated with low Ag dose, they can differentiate into Treg cells expressing Foxp3. Similarly, model simulations that mimic the low-stimulation scenario result in steady state with Foxp3 at level 1 (Figure 3(a)(middle)). Model simulation results show that the behavior of IL-2 gene expression early after stimulation is similar for both low- and high-stimulation scenarios.

This is not so straightforward to measure in experiments as IL-2 is measured outside of cells, where it is consumed quickly after being expressed and secreted. It is not clear from averaged simulation trajectories (Figure 3(a)(middle)) whether IL-2 reaches value 1 in all trajectories, but at different update rounds, or it reaches value 1 in 80% trajectories only. To test this, we consider the property P5 (Table I). Statistical model checking shows that the probability that this property holds is close to 1.

We have also computed the probability that IL-2 remains at level 0 until its inhibitor, Foxp3, becomes 1. This property:

$$(IL2 = 0)\ U[15]\ (FOXP3 = 1)$$

is returned as a low-probability event. In other words, our model predicts initial increase in IL-2, irrespective of antigen dose scenario, and the criticality of variations in other element values for phenotype decision.

Similar to the high-stimulation scenario, we created an STG for 1000 simulated trajectories in this scenario, shown in Figure 3(c). This STG is even more uniform than the one obtained for the high-stimulation scenario (Figure 3(b)), and it has 14 nodes and 26 edges. Interestingly, trajectories in this scenario initially vary, but then they all pass through a "hub" node, with following element values: Foxp3=0, IL-2=0, PTEN = 1, TCR=1, Ras=1, CD25=0, PI3K=1, Akt=1, mTORC1=0 and mTORC2=1. Furthermore, the time to reach steady state is different for these two scenarios, as discussed in [5]. Using model checking, we compare probabilities of reaching steady state at different rounds, as shown in Figure 4. The results show that Th cells obtain their fate much sooner than Treg cells. This points to the fact that those cells that take longer to differentiate due to lower stimulation will have a larger chance to become Treg cells.

C. Analysis of varying stimulation outcome

Another observation from experiments is that removal of antigen at 18 hours after stimulation results in a mixed population of Treg and Th cells [7]. Studies of the model have indicated that early events and relative timing of the Foxp3-activating and Foxp3-inhibiting paths play crucial role in differentiation [5]. With model checking, we were able to carry further and more efficient studies of early behavior in these elements.

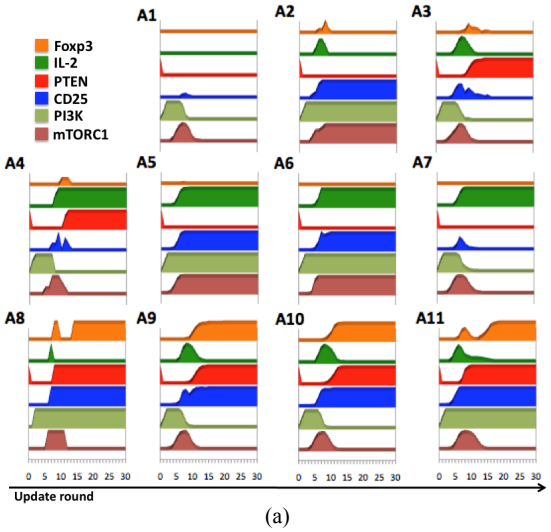
1) Distribution of attractors

The Ag removal scenario results in a mixed population of cells with different phenotypes. As shown in [5], depending on the time (simulation round) of stimulation removal, the presence and frequency of phenotypes vary. Eleven different phenotypes (attractors) were observed in those simulations, namely A1-A11. Antigen removal at round 6 (scenario T6), leads to a mixed population that carries all eleven phenotypes, while other scenarios, with earlier or later removal, often lead to a less diverse population. Time courses of averaged transient and steady state values of key model elements (Foxp3, IL-2, PTEN, CD25, PI3K, mTORC1) for scenario T6 are shown in Figure 5(a). These time courses emphasize differences in timing of elements across attractors. Hypotheses about critical events leading to differentiation can be drawn from analyzing these time courses and can be tested using model checking.

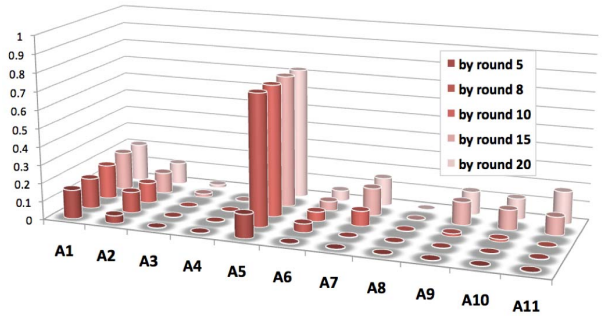
Using model checking, we compute a probability for each of these eleven attractors to be observed by a specific time point. Figure 5(b) shows estimated probabilities to reach steady state by round 5, 8, 10, 15, or 20, for all eleven attractors. Besides attractor A1, which represents cells that were not activated, the attractor that has highest probability of being reached early is A5 – attractor representing Th phenotype. All Th-like phenotypes (A5-A7) reach steady state by round 8, while Treg-like phenotypes (A9-A11) establish much later, by round 15-20. It is also interesting to note that attractors A9-A11 have similar probabilities of being reached in these later stages, while A5 is much more prevalent among Th-like phenotypes. This suggests that the variability present in trajectories leading to Treg cells, which causes delayed fate decision, remains present in steady state. This property is emphasized by existence of several different attractors with similar probabilities of occurrence.

2) Frequent state analysis

We show STGs obtained in Ag removal scenario for A5, the Th-like attractor that is closest to the high-dose attractor (Figure 3(d)), and A11, the Treg-like attractor that is closest to the low-dose attractor (Figure 3(e)). Compared to STGs for high and low dose scenarios, these two STGs have a larger number of nodes and connections. STG of attractor A5 is derived from 374 trajectories, it contains 24 nodes and 76 edges, while STG of attractor A11 is derived from 175 trajectories, and it contains 43 nodes and 139 edges. This once again emphasizes the relationship between system's transient behavior and steady states: conditions and initial states that result in more uniform trajectories also lead faster to steady states, in contrast to conditions and initial states that lead to large variability in trajectories and thus, delayed decision making.



(a)

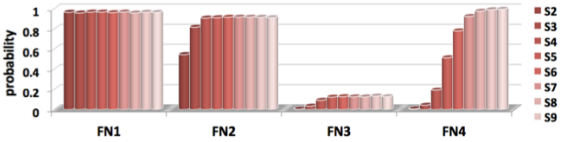


(b)

		High Ag dose + Ag removal at round 6											No removal	
Attractor		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	HD	LD
Attractor size		40	6	17	3	374	13	127	1	118	126	175	1000	1000
Number of nodes in STG		32	15	46	19	24	16	58	11	74	47	43	23	14
Number of edges in STG		76	29	110	28	76	34	206	15	250	151	139	71	26
Frequent nodes in STG	N1	78	78	74	78	78	78	74	78	74	73	78	78	177
	N2	74	73	508	73	73	73	508	73	469	476	422	73	337
	N3	28	426	798	256	426	426	426	789	483	426	426	425	849
	N4	3	42	255		469	490	511	785	799	800	810	426	
	N5	0										814		

(c)

Frequent node #	Node ID	Element values						
		Foxp3	IL-2	PTEN	CD25	Akt	mTORC1	mTORC2
FN1	73	0	0	0	0	1	0	1
FN2	74	0	0	0	0	1	1	1
FN3	78	0	0	0	0	0	0	1
FN4	426	0	1	0	1	1	1	1



(d)

Figure 5. Scenario in which high stimulation is not continuous but is instead removed after several rounds: (a) Average element values, obtained from [5] for all observed attractors (A1-A11) when stimulation is removed at round 6. (b) Probability of occurrence of each attractor by round 5,8,10,15, or 20, estimated using statistical model checker. (c) Comparison of STGs of 11 attractors resulting from antigen removal, high antigen dose attractor and low dose attractor. Several attractors have a large frequency of occurrence but very uniform STG (HD, LD, A5), several attractors have a characteristic to result from trajectories that vary a lot (A3, A7, A2, A10, A11), several attractors have a small frequency of occurrence, but often have varying trajectories (A1, A2, A3, A4, A8). (d) The most frequent nodes across all STGs, their IDs, and values of several elements (Foxp3, IL-2, PTEN, CD25, Akt, mTORC1, mTORC2) in the nodes are shown (top). Presence of frequent nodes in simulation trajectories for scenarios S2,..,S10, when antigen is removed at round i (i = 2,..,10), respectively (bottom).

Analysis of STGs for all observed attractors in both fixed and varying input simulations highlights a set of frequent nodes on all trajectories. Interestingly, as can be seen in Figure 5(c), some of these nodes are present in many of the trajectories, irrespective of the reached steady state (attractor). We highlight the most frequent nodes in Figure 5(c), and show key element values for those nodes in Figure 5(d)(top). Node FN1 occurs irrespective of attractor or trajectory. Node FN2 occurs mainly in trajectories that lead to activation. Node FN3 occurs early in trajectories and although it occurs less often, it is observed in a variety of trajectories leading to different attractors. Node FN4 is only present in trajectories leading to activation, and it occurs closer to steady state of Th phenotype, as a marker of this phenotype. However, it also transiently occurs in trajectories leading to Treg phenotype. This highlights the importance of transient behavior of other elements in the network. Values of other elements are necessary to further characterize this state in different attractors.

3) *Importance of timing in differentiation*

Testing a set of properties allows for uncovering events that lead to different attractors. In Table I, we present a set of properties (P9-P35) that we tested using statistical model checking, and results obtained from those tests. We also include the CPU time used for checking those properties.

Properties P9-P13 test the interplay between mTORC1, mTORC2 and CD25 in the model. Model checking computes a very high probability (close to 1) that mTORC1 and mTORC2 will both be at level 1 by round 7. Testing these properties also shows that early events in CD25, mTORC1 and mTORC2 are good predictors of the mixed population, as most of the results show close to 50% successes. In other words, although the tested properties would return more uniform behavior in other scenarios, in the case of antigen removal, we see more variability between possible trajectories.

Properties P14-P17 are related to the behavior of PTEN, an element found to be critical for the Treg phenotype [5]. Model checking confirms the hypotheses that mTORC1 can be transiently active in trajectories that lead to sustained PTEN (P14), that PTEN is necessary for sustained Foxp3 expression (P15), that about 50% of mixed population will have both Foxp3 and PTEN at level 0 (P14), and that there are trajectories in which Foxp3 is transiently expressed but do not have PTEN (and consequently, Foxp3) present in steady state (P17).

Properties P18-P29 test different conditions that are hypothesized to be important in deciding cell fate. These hypotheses are drawn from analyzing time courses in Figure 5(a). Properties P18-P19 are related to attractor A3, which occurs 1.7% of the time within 1000 simulations, and is marked as non-active phenotype. The P16-P17 tests show that Foxp3 occurs transiently in this attractor, and that presence of exogenous IL-2 does not change the outcome for this attractor. Properties P20-P21 are related to the A4 attractor, which occurs rarely within 1000 simulations, only about 0.6% of the time,

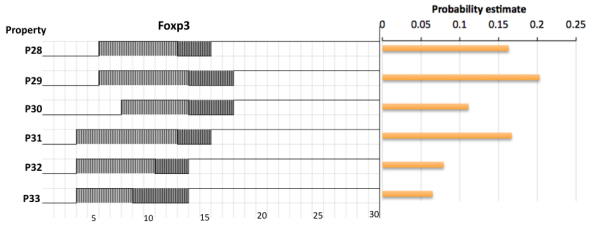


Figure 6. Analysis of Foxp3 behavior.

and although it has IL-2 and PTEN active in steady state, most of the other elements are inactive, thus it can be marked as another non-active phenotype. P20-P21 tests show that such phenotype can exhibit transient behavior of Foxp3 early, however, presence of PTEN is not sufficient to keep Foxp3.

Properties P22-P25 refer to attractors A5 and A6. The only difference between these two attractors is in element Ras, which is active in A5, but not active in A6. However, this seems to make a significant difference in occurrence rate of these two attractors. We tried to find the main cause of this difference using model checking. However, our hypotheses about sources of differentiation are rejected by testing properties P22-P25. Although attractor A5 has the highest occurrence in simulations, the properties P22-P23 do not match well the events leading to this attractor. One of the two hypotheses about the system does not seem to be correct: either Foxp3 has to occur transiently in trajectories leading to this attractor, or PI3K does not transient back to 0 once activated. Further studies are needed to resolve this. Similarly, transient behavior of PI3K does not seem the be the cause of different activation of Ras that leads to two phenotypes, A5 and A6, as initially hypothesized.

On the other hand, the property P26 does match well the behavior observed in trajectories leading to attractor A7. Foxp3 is not activated at all in this attractor, while mTORC1 is active early, opposite to CD25. Late activation of CD25 is most probably the reason for this attractor to become another non-active phenotype (only IL-2 present in steady state).

Attractors A9 and A10 have Foxp3 at level 1 in steady state, and defer only in Ras activation. These two attractors are tested using properties P27 and P28, respectively. The results show that both of them require early activation of Foxp3. Finally, testing P29 confirms the hypothesis that IL-2 and CD25 can both rise early in A11 attractor. Consequently, signaling through IL-2 receptor is activated early in those cases, establishing PI3K signaling. Interestingly, this is not sufficient to keep mTORC1 pathway active, most probably due to early activation of PTEN.

Since Foxp3 is a marker for Treg cells, and it has been found that its early transient behavior is important in differentiation, we test this behavior using properties P30-P35. We want to answer the following questions using those properties: Does early transient in Foxp3 always lead to the same attractor? What is the probability of Foxp3 transient starting at different time points and lasting different periods of time? The properties are presented as time courses in Figure 6, and estimated probabilities are plotted next to them. Studying these properties suggests that the Foxp3 increase to level 1 has a larger probability of starting at round 5 or later, that Foxp3 can very briefly decrease to level 0 after that, and increase back and stay at level 1.

VI. Conclusion

Model checking is an efficient approach for verifying and studying models of cell signaling networks, as it allows for answering a variety of questions about the system. Instead of manually analyzing simulation trajectories and large output files, one creates properties that can be automatically verified. Our goal in this work was to develop a fast, scalable method that is applicable to high-level models of cell signaling networks, and which can uncover events following varying stimulation conditions and leading to different cell phenotypes. Here, we used T cell differentiation model to demonstrate our framework, however, the framework is applicable to much larger models implementing complicated big mechanism networks of cellular signaling. Using the T cell model, we were able to find the most probable set of events, as well as temporal behavior of model elements resulting in mixed population of cells and in specific phenotypes. We tested a large number of properties, confirmed or rejected existing hypotheses, and specified a number of directions for future experiments and system studies.

Acknowledgement

This work is supported in part by DARPA Big Mechanism award W911NF-14-1-0422.

References

[1] Albert, I., Thakar, J., Li, S., Zhang, R., and Albert, R. 2008. Boolean network simulations for life scientists, Source Code for Biology and Medicine, vol. 3, no. 1, p. 16.

[2] Gong, H. Zuliani, P., Komuravelli, A., Faeder, J.R., and Clarke, E.M. 2010. Analysis and verification of the HMGB1 signaling pathway, BMC Bioinformatics, vol. 11, no. 7.

[3] Hoeffding, W. 1963. Probability Inequalities for Sums of Bounded Random Variables, Journal of the American Statistical Association 58 (301): 13–30.

[4] Jha, S.K., Clarke, E.M., Langmead, C.J., Legay, A., Platzer, A., and Zuliani, P. A Bayesian approach to Model Checking biological systems, in CMSB, ser. LNCS, vol. 5688, 2009, pp. 218–234.

[5] Miskov-Zivanov, N., Turner, M. S., Kane, L. P., Morel, P. A. and Faeder, J. R. 2013. Model Predicts Duration of T Cell Stimulation is a Critical Determinant of Cell Fate and Plasticity, under submission.

[6] Pnueli, A. 1977. The temporal logic of programs, in FOCS. IEEE, pp. 46–57.

[7] S. Sauer et al., T cell receptor signaling controls Foxp3 expression via PI3K, Akt, and mTOR. Proc Natl Acad Sci U S A 105, 7797 (Jun 3, 2008).

[8] Turner, M.S., Kane, L.P. and Morel, P.A. 2009. Dominant Role of Antigen Dose in CD4+ Foxp3+ Regulatory T Cell Induction and Expansion, The Journal of Immunology, vol. 183, no. 8, pp. 4895–4903.

[9] Vardi, M.Y. 1985. Automatic Verification of Probabilistic Concurrent Finite-State Programs, in FOCS, pp. 327-338.

[10] Wald, A. 1945. Sequential tests of statistical hypotheses. Annals of Mathematical Statistics, 16(2), 117–186.

[11] Younes, H.L.S., Simmons, R.G. 2002. Probabilistic verification of discrete event systems using acceptance sampling, in CAV, ser. LNCS, vol. 2404, pp. 223–235.

[12] Zuliani, P., Platzer, A. and Clarke, E.M. 2010. Bayesian statistical model checking with application to Stateflow/Simulink verification, in HSCC, pp. 243–252.

Table I. List of tested properties, probability estimate, count of trajectories resulting in a positive answer, sample size, and elapsed CPU time of a model checker on a 32-core system.

#	Property	Prob. estimate	Success count	Sample size	Elapsed time [s]
	***Scenario**: High stimulation dose*				
P1†	F^{29} (FOXP3 == 1), F^{10} (FOXP3 == 1 & F^{19} (FOXP3 == 0))	0.237494	2857	12032	120
P2*	F^{10} G^{2} (FOXP3 == 1)	0.0415313	10970	264160	2704
P3†	F^{10} G^{1} (FOXP3 == 1)	0.119089	830	6976	73
P4†	F^{20} G^{9} (FOXP3 == 0 & IL2 == 1 & PTEN == 0 & CD25 == 1 & PI3K == 1 & MTORC1 == 1 & MTORC2 == 1)	0.996124	256	256	2
	***Scenario**: Low stimulation dose*				
P5	F^{29} (IL2 == 1), F^{10} (IL2 == 1 & F^{19} (IL2 == 0))	0.996124	256	256	3
P6	F^{10} G^{2} (IL2 == 1)	0.781024	8873	11360	115
P7	F^{10} G^{1} (IL2 == 1)	0.979681	1349	1376	14
P8	F^{20} G^{9} (FOXP3 == 1 & IL2 == 0 & PTEN == 1 & CD25 == 1 & PI3K == 1 & MTORC1 == 0 & MTORC2 == 1)	0.996124	256	256	2
	***Scenario**: High stimulation dose + stimulation removal at round 6*				
P7†	G^{7} ~(MTORC1 == 1 & MTORC2 == 1)	0.019567	46	2400	38
P8†	F^{7} (MTORC1 == 1 & MTORC2 == 1)	0.982159	2201	2240	34
P9†	F^{10} (MTORC1 == 1 & MTORC2 == 1 & CD25 == 0 & (F^{18} (CD25 == 1)))	0.600977	15616	25984	407
P10†	F^{28} (MTORC1 == 1 & MTORC2 == 1 & CD25 == 0 & (F^{1} (CD25 == 1)))	0.590649	15461	26176	405
P11†	F^{10} (MTORC1 == 1 & MTORC2 == 1 & CD25 == 0 & (F^{1} (G^{17} (CD25 == 1))))	0.405376	10585	26112	404
P12†	F^{10} (MTORC1 == 1) & F^{15} G^{10} (PTEN == 1)	0.197865	3409	17232	175
P13§	F^{25} G^{4} (FOXP3 == 1 & PTEN == 0)	2.893e-05	0	34560	350
P14¶	F^{2} G^{26} (FOXP3 == 0 & PTEN == 0)	0.550633	608	1104	11
P15†	F^{2} G^{25} (PTEN == 0) & F^{10} (FOXP3 == 1)	0.036126	143	3984	41
P16*	F^{9} (MTORC1 == 1) & F^{10} (IL2 == 1) & F^{12} (IL2_EX == 1) & F^{13} (FOXP3 == 1) & F^{15} (PTEN == 1) & F^{20} G^{9} (FOXP3 == 0 & IL2 == 0 & MTORC1 == 0 & PTEN == 1 & CD25 == 0 & MTORC2 == 0)	0.009301	569	61280	622
P17*	F^{9} (MTORC1 == 1) & F^{10} (IL2 == 1) & F^{13} (FOXP3 == 1) & F^{15} (PTEN == 1) & F^{20} G^{9} (FOXP3 == 0 & IL2 == 0 & MTORC1 == 0 & PTEN == 1 & CD25 == 0 & MTORC2 == 0)	0.009067	541	59776	610
P18*	F^{9} (MTORC1 == 1) & F^{10} (IL2 == 1) & F^{12} (IL2_EX == 1) & F^{13} (FOXP3 == 1) & F^{15} (PTEN == 1) & F^{20} G^{9} (FOXP3 == 0 & IL2 == 1 & MTORC1 == 0 & PTEN == 1 & CD25 == 0 & MTORC2 == 0)	0.001202	10	9152	92
P19*	F^{9} (MTORC1 == 1) & F^{10} (IL2 == 1) & F^{13} (FOXP3 == 1) & F^{15} (PTEN == 1) & F^{20} G^{9} (FOXP3 == 0 & IL2 == 1 & MTORC1 == 0 & PTEN == 1 & CD25 == 0 & MTORC2 == 0)	0.000858	4	5824	60
P20*	G^{29} (FOXP3 == 0) & F^{5} (MTORC1 == 1) & F^{3} (PI3K == 1 & F^{4} (PI3K == 0)) & F^{20} G^{9} (IL2 == 1 & MTORC1 == 1 & PTEN == 0 & CD25 == 1 & MTORC2 == 1 & PI3K == 1 & RAS == 1)	0.000434	0	2304	25
P21*	G^{29} (FOXP3 == 0) & F^{5} (MTORC1 == 1) & F^{6} (CD25 == 1) & F^{3} (PI3K == 1 & F^{4} (PI3K == 0)) & F^{20} G^{9} (IL2 == 1 & MTORC1 == 1 & PTEN == 0 & CD25 == 1 & MTORC2 == 1 & PI3K == 1 & RAS == 1)	0.000434	0	2304	24
P22*	G^{29} (FOXP3 == 0) & F^{5} (MTORC1 == 1) & F^{3} (PI3K == 1 & F^{4} (PI3K == 0)) & F^{20} G^{9} (IL2 == 1 & MTORC1 == 1 & PTEN == 0 & CD25 == 1 & MTORC2 == 1 & PI3K == 1 & RAS == 0)	0.000434	0	2304	24
P23*	G^{29} (FOXP3 == 0) & F^{3} (MTORC1 == 1) & F^{6} (CD25 == 1) & F^{3} (PI3K == 1 & F^{4} (PI3K == 0)) & F^{20} G^{9} (IL2 == 1 & MTORC1 == 1 & PTEN == 0 & CD25 == 1 & MTORC2 == 1 & PI3K == 1 & RAS == 0)	0.000434	0	2304	24
P24*	G^{29} (FOXP3 == 0) & F^{5} (MTORC1 == 1) & G^{5} (CD25 == 0) & F^{20} G^{9} (IL2 == 1 & MTORC1 == 0 & PTEN == 0 & CD25 == 0 & MTORC2 == 0 & PI3K == 0 & RAS == 0)	0.080004	39072	488384	4967
P25*	G^{11} (FOXP3 == 0) & F^{15} G^{14} (FOXP3 == 1 & IL2 == 0 & MTORC1 == 0 & PTEN == 1 & CD25 == 1 & MTORC2 == 0 & PI3K == 1 & RAS == 0)	0.000434	0	2304	23
P26*	G^{11} (FOXP3 == 0) & F^{15} G^{14} (FOXP3 == 1 & IL2 == 0 & MTORC1 == 0 & PTEN == 1 & CD25 == 1 & MTORC2 == 0 & PI3K == 1 & RAS == 1)	0.000434	0	2304	24
P27*	F^{4} (IL2 == 1) & F^{4} (CD25 == 1) & F^{20} G^{9} (FOXP3 == 1 & IL2 == 0 & MTORC1 == 0 & PTEN == 1 & CD25 == 1 & MTORC2 == 1 & PI3K == 1 & RAS == 1)	0.020230	2661	131584	1338
P28†	(G^{5} (FOXP3 == 0)) & (F^{12} (FOXP3 == 1 & F^{3} (FOXP3 == 0))) & G^{29} ((step < 16 | FOXP3 == 1))	0.16306	1476	9056	95
P29†	(G^{5} (FOXP3 == 0)) & (F^{13} (FOXP3 == 1 & F^{4} (FOXP3 == 0))) & G^{29} (step < 18 | FOXP3 == 1)	0.202901	2181	10752	107
P30†	(G^{7} (FOXP3 == 0)) & (F^{13} (FOXP3 == 1 & F^{4} (FOXP3 == 0))) & G^{29} (step < 18 | FOXP3 == 1)	0.111162	732	6592	68
P31†	(G^{3} (FOXP3 == 0)) & (F^{12} (FOXP3 == 1 & F^{3} (FOXP3 == 0))) & G^{29} (step < 16 | FOXP3 == 1)	0.166919	1543	9248	93
P32†	(G^{3} (FOXP3 == 0)) & (F^{10} (FOXP3 == 1 & F^{3} (FOXP3 == 0))) & G^{29} (step < 14 | FOXP3 == 1)	0.079243	38363	484128	4933
P33†	(G^{3} (FOXP3 == 0)) & (F^{8} (FOXP3 == 1 & F^{5} (FOXP3 == 0))) & G^{29} (step < 14 | FOXP3 == 1)	0.065006	26216	403296	4088

Coverage probability=0.999; Half-interval=†0.01, *0.001, §0.0001, ¶0.5.