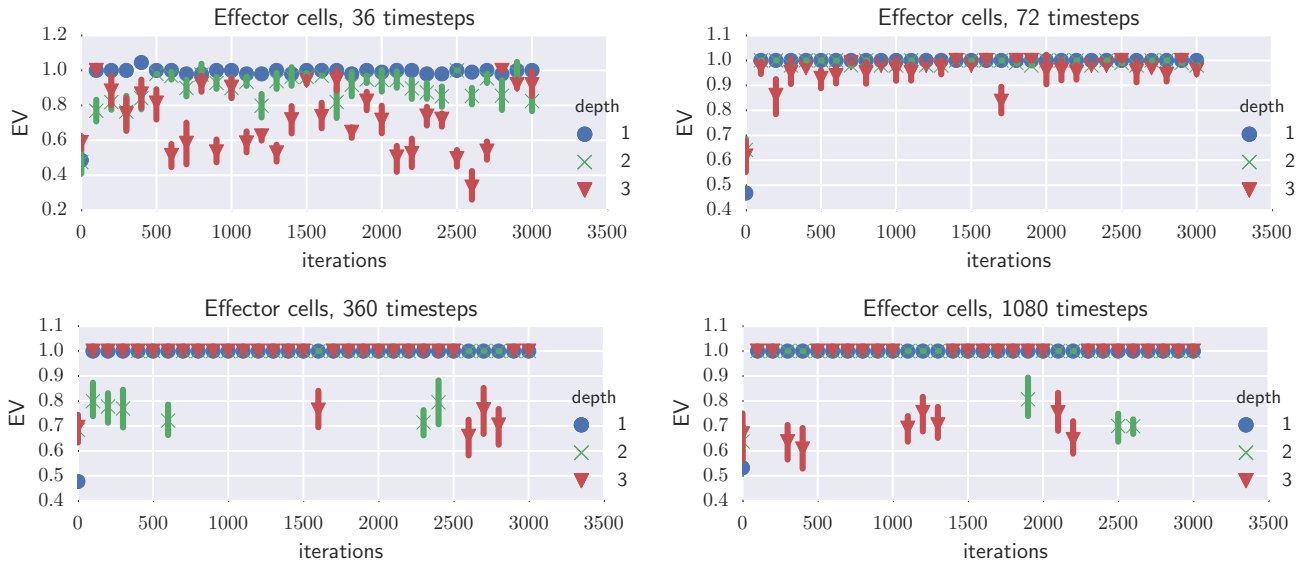

Figure 3: Average utility as a function of the number of UCT iterations (nodes touched in the tree) for effector T cell production. Top left: 36 total time steps, top right: 72 total time steps, bottom left: 360 total time steps, bottom right: 1080 total time steps.

Observed state	Action to apply
1-step plans	
All four *BioNetGen* simulation steps	
Initial state (depth 0):	TCR=1
2-step plans	
36 *BioNetGen* sim. steps	
Initial state (depth 0):	TCR=1, TGFβ=1, IL-2=0
Depth 1, FOXP3=0, IL-2=0:	nothing
Depth 1, FOXP3=0, IL-2=1:	TCR=0 , CD28=0
Depth 1, FOXP3=1, IL-2=0:	CD28=0
Depth 1, FOXP3=1, IL-2=1:	TCR=0 , CD28=0 , IL-2=1
2-step plans	
72 *BioNetGen* sim. steps	
Initial state (depth 0):	TCR=1, TGFβ=0 , IL-2=0
Depth 1, FOXP3=0, IL-2=0:	nothing
Depth 1, FOXP3=0, IL-2=1:	TCR=0, CD28=0, TGFβ=1
Depth 1, FOXP3=1, IL-2=0:	CD28=0, TGFβ=1, IL-2=1
2-step plans	
360 *BioNetGen* sim. steps	
Initial state (depth 0):	TGFβ=1, IL-2=0
Depth 1, FOXP3=0, IL-2=1:	TCR=1
2-step plans	
1080 *BioNetGen* sim. steps	
Initial state (depth 0):	TGFβ=1
Depth 1, FOXP3=0, IL-2=1:	TCR=1, IL-2=0

Table 1: Generated plans for regulatory cell development. For depth 2, omitted rows denote configurations that are never reached. In the "action to apply" column, TCR, 1/0 denotes high/low respectively; for other variables 1/0 denotes activate/inhibit, respectively.

and computational effort, and the latter also affects testing costs. With the ability to do those and other experiments first *in silico*, significant time and cost savings can be obtained by reducing the needed *in vitro* and *in vivo* experimentation.

Future work also involves using our multi-step steering approach for steering other biological entities beyond T cells, such as steering the evolution of bacteria or viruses into states where they can be effectively tackled, steering cancer cell populations to states where they can be destroyed without leaving persistors, or, in synthetic biology, steering bacteria into states where they perform useful tasks (such as consuming oil spills) without introducing foreign genetic material into the bacteria, which is costly and risky.

Finally, there are opportunities for better performance by considering other, more sophisticated, planning algorithms. For example, algorithms such as POMCP [Silver and Veness, 2010] or DESPOT [Somani *et al.*, 2013] could allow better scalability, and thereby enable analysis of larger models.

Acknowledgements This material is based on work supported by the National Science Foundation under grants IIS-1320620 and IIS-1546752, and by the ARO under award W911NF-16-1-0061. We thank Penelope A. Morel, Jose-Juan Tapia, James R. Faeder, and Natasa Miskov-Zivanov for providing immunology expertise, and for discussions about the biological model and simulation. We also thank the anynomous reviewers and SPC members for useful critiques and interesting new directions.

References

[Adams *et al.*, 2004] B Adams, H Banks, H-D Kwon, and H Tran. Dynamic multidrug therapies for HIV: Optimal and STI control approaches. *Mathematical Biosciences and Engineering*, 1:223–241, 2004.