Reinforcement learning for HIV treatment

Lucian Buşoniu

Consider the application of reinforcement learning (RL) to finding treatment strategies for the infection with the human immunodeficiency virus (HIV), reported by Ernst et al. (2006). Write a brief paper about this topic, integratin answers to the following questions in the logical flow of your paper. Do not restrict your reading to this single article (Ernst et al., 2006); e.g., some of the questions below may require further reading.

- How is supervised learning combined with RL in the fitted Q-iteration algorithm used in the article?

- What is the most desirable steady state for a patient whose infection is controlled using structured treatment interruptions?

- Explain in words the meaning of the cost function given in Equation (9) of the article. Assume that a treatment strategy has to be computed for a patient that is known to suffer from the side-effects of the PI and RTI drugs to a more severe extent than a typical patient. How should the cost function be modified to account for this?

- What do you think was the reason for choosing the iterative procedure to generate the data, described in Section 3.2? Would the algorithm work better, or worse, if all the data were collected in the first iteration?

- The approach in this article is essentially offline RL: data is collected in advance, and then fed into the RL algorithm to find a solution. There also exist RL algorithms for online RL, that start from scratch, with no data at all, and collect their own data while at the same time learning on the basis of this data. Give an example of such an algorithm. In online RL, a significant amount of exploration is needed, which amounts to taking many times actions at random. Do you believe that online RL would be (i) more, or (ii) less suitable than offline RL, for the treatment of HIV-infected patients? Why?

References

Ernst, D., Stan, G.-B., Gonçalves, J., and Wehenkel, L. (2006). Clinical data based optimal STI strategies for HIV: A reinforcement learning approach. In *Proceedings 2006 Machine Learning Conference of Belgium and The Netherlands (Benelearn-06)*, pages 65–72, Ghent, Belgium.