Steering Evolution and Biological Adaptation Strategically:

Computational Game Theory and Opponent Exploitation for Treatment Planning, Drug Design, and Synthetic Biology*

Tuomas Sandholm
Computer Science Department
Carnegie Mellon University

Living organisms adapt to challenges through evolution and adaptation. These survival mechanisms have proven to be a key difficulty in developing therapies, since the challenged organisms develop resistance.

It would be desirable to harness evolution/adaptation for therapeutic, technological, and scientific goals. I propose steering them *strategically* using computational game theory and opponent exploitation techniques. A sequential contingency plan for steering evolution/adaptation is constructed computationally for the setting at hand. For example, for therapeutics, I propose modeling this as a (zero-sum) imperfect-information game between a treater and a disease, with potentially both sequential and simultaneous moves.

Game-theoretic solution. Solving the game for Nash equilibrium (or its refinements) provides an optimal treatment plan assuming the disease plays optimally. The scalability of algorithms for imperfect-information games has increased by orders of magnitude over the last ten years [12, 13]—largely driven by the Annual Computer Poker Competition. The leading approach involves running an abstraction algorithm to construct a smaller, strategically similar game, then computing a (near-)equilibrium of the abstract game, and then mapping the computed strategies to the original game. The game-theoretic approach does not need a probabilistic model of disease behavior and can thus substitute for statistical knowledge at states-action-pairs where such knowledge is missing: it assumes the opponent behaves in the worst possible way for us. This is safe in the zero-sum setting: if the opponent actually does not behave in this way, that can only benefit us.

Opponent exploitation. The game-theoretic approach may be overly conservative: the disease may not play optimally. I propose that opponent exploitation techniques be used to take advantage of the disease's suboptimal play. An opponent model predicts what the opponent would do—perhaps probabilistically—at various points in the game. Then, there are many approaches to opponent exploitation [12, 11, 15, 4, 16]. For example, one can start by playing game theoretically and then adjust play toward exploiting the opponent at points of the game that have been frequently visited so there is good statistical information about the opponent's play there [2]. Or, one can compute an ϵ-*safe best response* (or approximation thereof), *i.e.*, a strategy that exploits our model of the opponent maximally subject to the constraint that even against a worst-case opponent it will do at most ϵ worse than a game-theoretic strategy [6, 5].[1] As a third example, one can compute a set of strategies and then use learning—in simulation or live—to determine which of the strategies perform well against the opponent [1]. As a fourth example, if one trusts the opponent model enough, one can abandon game-theoretic safety completely and compute a best-response strategy (or an approximation thereof) to the opponent model. This can be computationally complex in large games and with lots of randomness in the game. To find solutions for this setting, techniques from stochastic optimization can be leveraged, such as *trajectory-based optimization* (*e.g.*, based on sample trajectories of possible futures) and *policy gradient techniques*. At the other extreme, one can assume no prior knowledge of the opponent and yet require that one's opponent exploitation performs at least as well in expectation as a game-theoretic equilibrium strategy. Perhaps surprisingly, it turns out that it is indeed possible to exploit an opponent more than any game-theoretic equilibrium strategy can, while still having this safety property [3]. Intuitively, if we can measure—or at least bound from below—how much value the opponent has gifted to us through suboptimal moves, we can use that value to bankroll our risky exploitation while guaranteeing safety overall.

*Patent pending, 2012.

[1] Typically ϵ is measured in expectation, but could potentially alternatively be measured in the worst case or in some other sense.

Biological opponents have a distinct weakness that we can further exploit. ***Evolution and adaptation are myopic***: they do not look ahead in the game tree. We can thus steer the disease to evolve into a ***trap*** where it can be easily attacked so that it is destroyed or becomes less powerful. More generally, the task is to compute a *strategy* for the treater in the game that causes the myopic opponent to obtain low utility [9].

Benefits. Algorithms can often solve games better than humans, so there is potential for better treatment plans. Most medical treatment today is myopic: the treater tries to take an action that improves the patient's health immediately. This puts the treater at the same disadvantage that the opponent has! The planning I propose may myopically make the patient worse in preparation for later, effective treatments. For example, the sequential treatment may cause a virus population in an individual to evolve to a population that can be effectively treated. (Some multistage plans are in use already, such as radiosensitization and chemosensitization, but those plans are manually crafted, and much shorter/simpler than the ones contemplated here.)

Because the proposed planning is automated, it is dramatically faster and requires fewer human resources. Thus custom plans can be generated for more specific population segments—even individuals.

The speed also enables the user of the system to conduct what-if analyses (sensitivity analysis) to test how the system-generated plan would change under different assumptions about the game and opponent.

There is also potential to direct medical research as a form of active learning guided by the proposed framework. The most valuable knowledge to generate is that which enhances value of treatment plans most.

Applications

Battling disease in an individual. Consider treating an HIV patient. At each point in the game, the actions the treater can take may include treatments (such as which drug/cocktail to use) and tests. The actions the adversary (HIV) can take may include evolving the virus pool within the patient, making the patient worse or better in any number of ways, *etc*. There is HIV data about the typical mutation sites and the typical mutations at each site. There is also data on treatment outcomes. In the game model, utilities can be based on the patient's health and projected health (including side effects), how virulent/contagious a state the disease is in, how easily attackable the disease is in its current state, the cost of treatment and other costs to the patient/treater/insurance so far, projected costs, *etc*. The output is a strategy, that is, a contingent plan.

Battling disease at the molecular level: Drug design. Consider designing drugs or cocktails for treating an HIV patient or segment of patients. HIV's actions at any point in the game include the likely mutations in the likely mutation sites. The treater's actions at any point in the game include which cocktail (what amounts of each drug) to use. Or, the actions can be to choose a cocktail of some existing drugs and some *de novo* drugs designed via the proposed approach. So, actions can include selection from a vast, or even infinite, space of potential drugs. It is not atypical for game models to be solvable even if the action space is infinite [12, 8]. The actions of the treater can also include tests on the patient and virus population therein. A model can be used to predict how well each of the potential drugs in the potential cocktails would bind to each mutation at each site (*e.g.*, [7, 10]). The utility of the disease can then be the sum over binding sites of the predicted binding energy at the site. The output is again a strategy.

Battling disease in a population. Consider an epidemic. Its actions at any point in the game include spread of the various strands—possibly including mutations—to various parts of the population. The actions of the treater at any point in the game include which drug or cocktail to use in each part of the population. The actions can also include quarantining, *etc*. In principle, the actions could also include selection from an unrestricted drug design space. The actions of the treater can also include conducting tests on patients from various parts of the population and testing aspects of the epidemic within such patients. The treater's utility can be based on deaths and other costs such as hospitalizations. The output is again a strategy.

Cell repurposing. The approach has applications beyond battling diseases. For example, one could apply it to repurpose cells. Could one evolve, say, a blood cell into a liver cell, or even grow an organ or limb? In fact, I am currently proposing to start a project to test the approach with a biology collaborator to steer the adaptation of a patient's own T cell population to better tackle cancer or autoimmune diabetes.

Synthetic biology. The ideas apply beyond medical treatment as well. For instance, could one evolve bacteria into ones that eat toxins—such as oil spills—without introducing foreign genetic material?

Tackling questions in natural science. The approach also enables one to formalize and potentially answer fundamental questions in science. For example, can a certain kind of cell be transformed into a certain other kind of cell by evolutionary pressures using a given set of manipulations? How much more power do multi-stage treatment plans offer? Does there exist a strategy (that uses only a given set of available manipulations and tests) that will destroy a given diverse (*e.g.*, cancer) cell population in a way that leaves no persistors? What is inherently impossible to achieve via evolution?

For further information about this vision and related research, see my AAAI-15 Senior Member Track Blue Skies paper [14].

References

[1] Nolan Bard, Michael Johanson, Neil Burch, and Michael Bowling. Online implicit agent modelling. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2013.

[2] Sam Ganzfried and Tuomas Sandholm. Game theory-based opponent modeling in large imperfect-information games. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2011.

[3] Sam Ganzfried and Tuomas Sandholm. Safe opponent exploitation. *ACM Transaction on Economics and Computation (TEAC)*, 2014. Best of EC-12 special issue.

[4] Bret Hoehn, Finnegan Southey, Robert C. Holte, and Valeriy Bulitko. Effective short-term opponent exploitation in simplified poker. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 783–788, July 2005.

[5] Michael Johanson and Michael Bowling. Data biased robust counter strategies. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2009.

[6] Michael Johanson, Martin Zinkevich, and Michael Bowling. Computing robust counter-strategies. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2007.

[7] Hetunandan Kamichetty. Application: Games of molecular conflict. In *Structured Probabilistic Models of Proteins across Spatial and Fitness Landscapes*, Carnegie Mellon University, School of Computer Science PhD thesis (CMU-CS-11-116), chapter 9, pages 121–127. March 2011.

[8] Christian Kroer and Tuomas Sandholm. Discretization of continuous action spaces in extensive-form games. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2015.

[9] Christian Kroer and Tuomas Sandholm. Limited lookahead in incomplete-information games. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.

[10] Christopher Langmead and Hetunandan Kamichetty. Resistance-proof drug design. Provisional US patent application 61/464149, 2011.

[11] Peter McCracken and Michael Bowling. Safe strategies for agent modelling in games. In *AAAI Fall Symposium on Artificial Multi-agent Learning*, 2004.

[12] Tuomas Sandholm. The state of solving large incomplete-information games, and application to poker. *AI Magazine*, pages 13–32, Winter 2010. Special issue on Algorithmic Game Theory.

[13] Tuomas Sandholm. Abstraction for solving large incomplete-information games. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2015. Senior Member Track.

[14] Tuomas Sandholm. Steering evolution strategically: Computational game theory and opponent exploitation for treatment planning, drug design, and synthetic biology. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2015. Senior Member Track.

[15] Finnegan Southey, Michael Bowling, Bryce Larson, Carmelo Piccione, Neil Burch, Darse Billings, and Chris Rayner. Bayes' bluff: Opponent modelling in poker. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 550–558, July 2005.

[16] Nathan Sturtevant, Martin Zinkevich, and Michael Bowling. Prob-max^n: Opponent modeling in n-player games. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 1057–1063, Boston, MA, 2006.