Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

PREDICTING COOPERATION AND DESIGNING INSTITUTIONS: AN INTEGRATION OF BEHAVIORAL DATA, MACHINE LEARNING, AND SIMULATION

John Jacob Nay*

Vanderbilt University
Nashville, TN, USA

ABSTRACT

Empirical game theory experiments attempt to estimate causal effects of institutional factors on behavioral outcomes by systematically varying the rules of the game with human participants motivated by financial incentives. I developed a computational simulation analog of empirical game experiments that facilitates investigating institutional design questions. Given the full control the artificial laboratory affords, simulated experiments can more reliably implement experimental designs. I compiled a large database of decisions from a variety of repeated social dilemma experiments, developed a statistical model that predicted individual-level decisions in a held-out test dataset with 90% accuracy, and implemented the model in agent-based simulations where I apply constrained optimization techniques to designing games – and by theoretical extension, institutions – that maximize cooperation levels. This presentation describes the methodology, preliminary findings, and future applications to applied simulation models as part of ongoing multi-disciplinary projects studying decision-making under social and environmental uncertainty.

1 DATA

I gathered existing data from experiments conducted with human subjects playing many variations on repeated prisoner's dilemma games with real financial incentives (Bereby-Meyer and Roth 2006, Duffy and Ochs 2009, Kunreuther, Silvasi, Bradlow, and Small 2009, Dal Bo and Frechette 2011, Fudenberg, Rand, and Dreber 2012) and created standardized features of the games and behavior across the datasets. This process resulted in a database of over 140,000 individual decisions with consistent measurements on over 60 behavioral-economics-inspired features for each observation. The features represent three categories: (1) the history of an interaction between a player and her opponents (e.g. the player cooperated the past two rounds and her opponent defected both rounds); (2) repeated game strategies for a player that interaction (e.g. the action that she would take if she followed a "tit-for-tat" strategy); and (3) the rules of the game (e.g., the payoff associated with cooperating when her opponent defects).

2 STATISTICAL MODEL

The statistical goal is to learn an approximation to a stochastic function mapping these features into cooperation decisions of person i at time t in game g that minimizes prediction error on unseen data, i.e. to find the discriminate function $f(\theta)$ that minimizes the expected value of $L(Cooperate_{itg}, f(\vec{Game}_g, \vec{History}_{1:t-1}, \theta))$ over the population joint distribution of the features and cooperation decisions, where $L()$ is a loss function specific to the learning algorithm, $Game_g$ are the rules of the game, and $History_{1:t-1}$ is i's and the other players' decisions and payoffs from the beginning of the game until $t-1$. Support-vector machines and neural networks have a comparative advantage in extracting complex combinations of features, while binary-split

*I would like to thank Eugene Vorobeychik and Jonathan Gilligan for feedback and insights, and to acknowledge financial support from U.S. National Science Foundation Grant No. NSF-EAR 1204685.

978-1-4799-7486-3/14/$31.00 ©2014 IEEE

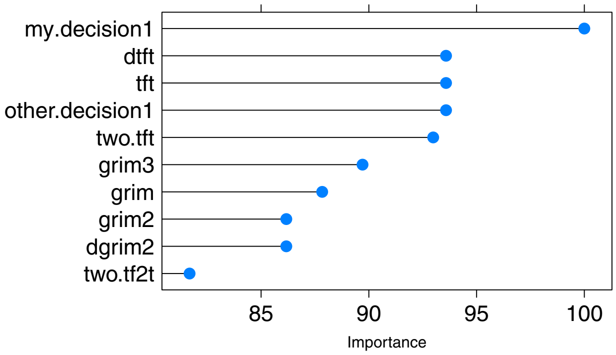

Figure 1: 10 most important predictors for a random forest model with just under 90% out-of-sample accuracy. Mydecision1 and otherdecision1 are the decisions taken in the previous period. The rest are repeated game strategies. The only strategies in the top 10 predictor list were variants of the famous tit-for-tat (tft) reciprocal strategy, and the grim strategy, where after one defection of your partner you always defect.

tree-based models have a comparative advantage in handling noisy feature sets (Hastie, Tibshirani, and Friedman 2009). Random forests can also provide useful feature importance rankings (see Figure 1). I compare these alternatives[1] and determine that an ensemble of neural network models has the highest predictive accuracy, predicting individual-level decisions in a held-out test dataset with 90% accuracy.

3 SIMULATION MODEL

I took an agent-based approach to empirical modeling by modeling the game situation as the outcome of the behavior of multiple agents making decentralized, inductive decisions with data input unique to each agent. In order to use the empirical models of individual adaptive behavior to simulate outcomes of new experimental designs I employ agent-based model simulations. I am using genetic algorithms to search game parameter space of the simulation model for games, $\vec{Game}_g$, that maximize average cooperation levels, $\bar{C}$. Only a few configurations of $\vec{Game}_g$ are tested in the data due to feasibility, but simulation allows us to search much larger game parameter space using the highly predictive models and to control "social context," e.g. by using null opponents that always cooperate. The simulation results, intended to show the size and direction of the effects of the rules of the game on cooperation levels, will be used to inform hypotheses to empirically test in observational settings related to social dilemmas and the environment.

REFERENCES

Bereby-Meyer, Y., and A. E. Roth. 2006. "The speed of learning in noisy games: partial reinforcement and the sustainability of cooperation". *American Economic Review* 96 (4): 1029–1042.

Dal Bo, P., and G. R. Frechette. 2011, February. "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence". *American Economic Review* 101 (1): 411–429.

Duffy, J., and J. Ochs. 2009, July. "Cooperative behavior and the frequency of social interaction". *Games and Economic Behavior* 66 (2): 785–812.

Fudenberg, D., D. G. Rand, and A. Dreber. 2012, April. "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World". *American Economic Review* 102 (2): 720–749.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. 2nd ed. 2009. Corr. 7th printing 2013 edition ed. Springer.

Kunreuther, H., G. Silvasi, E. T. Bradlow, and D. Small. 2009. "Bayesian analysis of deterministic and stochastic prisoners dilemma games". *Judgment and Decision Making* 4 (5): 363–384.

[1] I use a Gini index as a measure of prediction error when building the trees: $G = \sum_{k=1}^{K} \hat{p}_{mk}(1-\hat{p}_{mk})$, where $\hat{p}_{mk}$ is the proportion of training observations in the mth region of feature space from the kth class of the outcome. I randomly sample observations, grow a tree on each sample, and then use the majority vote when making predictions. Randomness is also introduced by forcing each tree to consider different randomly selected sets of features at each step. I also employ a bootstrap aggregating procedure with feed-forward neural network models with single hidden layers. Finally, I use a support vector machine with a radial basis kernel function, $K(x,x') = exp(-\gamma||x-x'||^2)$ (Hastie, Tibshirani, and Friedman 2009).