Strategy Purification*

(Extended Abstract)

Sam Ganzfried, Tuomas Sandholm, and Kevin Waugh
Computer Science Department
Carnegie Mellon University
{sganzfri, sandholm, waugh}@cs.cmu.edu

ABSTRACT

There has been significant recent interest in computing good strategies for large games. Most prior work involves computing an approximate equilibrium strategy in a smaller abstract game, then playing this strategy in the full game. In this paper, we present a modification of this approach that works by constructing a *deterministic* strategy in the full game from the solution to the abstract game; we refer to this procedure as *purification*. We show that purification, and its generalization which we call *thresholding*, lead to significantly stronger play than the standard approach in a wide variety of experimental domains. One can view these approaches as ways of achieving robustness against one's own lossy abstraction.

Categories and Subject Descriptors

I.2 [**Computing Methodologies**]: Artificial Intelligence

General Terms

Algorithms, Economics

Keywords

Game theory

1. INTRODUCTION

Significant work has been done in recent years on computing game-theory-based strategies in large games; this work typically follows a three-step approach. In the first step, an *abstraction algorithm* is run on the original game G to construct a smaller game G' which is strategically similar to G [1, 3]. Next, an *equilibrium-finding algorithm* is run on G' to compute an ϵ-equilibrium σ' [2, 7]. Finally, a *reverse mapping* is applied to σ' to compute an approximate equilibrium σ in the full game G [4, 5]. Almost all prior work has used the trivial reverse mapping, in which σ is the straightforward projection of σ' into G. In other words, once the abstract game is solved, its solution is just played directly in the full game. In this paper, we show that applying a nontrivial reverse mapping can lead to significant performance improvements — even in games where the trivial mapping is possible.

*This material is based upon work supported by the National Science Foundation under IIS grants 0905390 and 0964579. We also acknowledge Intel Corporation and IBM for their machine gifts.

Cite as: Strategy Purification (Extended Abstract), Sam Ganzfried, Tuomas Sandholm, and Kevin Waugh, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1111-1112.

2. THRESHOLDING AND PURIFICATION

Let τ be a mixed strategy for a player in a strategic-form game, and let $S = \arg\max_j \tau_j$, where j ranges over all of the player's pure strategies. Then we define the *purification* $\text{pur}(\tau)$ of τ as follows:

$$\text{pur}(\tau)_j = \begin{cases} 0 & : \quad j \notin S \\ \frac{1}{|S|} & : \quad j \in S \end{cases}$$

Informally, this says that if τ plays a single pure strategy with highest probability, then the purification will play that strategy with probability 1. If there is a tie between several pure strategies of the maximum probability played under τ, then the purification will randomize equally between all maximal such strategies. Thus the purification will usually be a pure strategy, and will only be a mixed strategy in degenerate special cases when several pure strategies are played with identical probabilities.

Purification can sometimes seem quite extreme; for example, if τ plays action a with probability 0.51 and action b with probability 0.49, then $\text{pur}(\tau)$ will never play b. Maybe we would like to be a bit more conservative, and only set a probability to 0 if it is below some threshold ϵ, then normalize the probabilities. We refer to this new algorithm as *thresholding*. One intuitive interpretation of thresholding is that actions with probability below ϵ were just given positive probability due to noise from the abstraction (or because an anytime equilibrium-finding algorithm had not yet taken those probabilities all the way to zero), and really should not be played in the full game.

3. RANDOM MATRIX GAMES

The first set of experiments we conducted to demonstrate the power of purification was on random matrix games. We studied two-player zero-sum games with three actions per player and payoffs for the row player drawn uniformly at random from [0,1]. The payoffs for the column player are 1 minus the row player's payoff, so for each strategy profile the payoffs sum to 1.

We repeatedly generated random games and analyzed them using the following procedure. First, we computed an equilibrium of the full 3×3 game Σ; denote this strategy profile by σ_F. Next, we constructed an abstraction Σ' of Σ by ignoring the final row and column of Σ and computed an equilibrium σ_A of Σ'. We then compared $u_1(\sigma_A, \sigma_F)$ to $u_1(\text{pur}(\sigma_A), \sigma_F)$.

Our experiments conclude at the 95% confidence level that purification improves performance over the standard abstraction approach; the average payoff for purification was 0.449 while that of abstraction was 0.447[1]. These results are very surprising, since the abstractions we used were completely random and hence quite naïve.

4. LEDUC HOLD'EM

Leduc Hold'em is a small poker game that has been previously used to evaluate imperfect-information game-playing techniques [6]. It is large enough that abstraction has a nontrivial impact, but unlike larger games of interest it is small enough that equilibrium solutions in the full game can be quickly computed.

To evaluate the effects of purification in Leduc Hold'em, we compared the performance of the 24 abstract equilibrium strategies from [6] against a single equilibrium opponent. We observed that purification improved the performance of the abstract equilibrium in all but five cases. In many cases this improvement was quite substantial. For example, prior to purification the best abstract equilibrium strategy lost at 43.8 millibets per hand (mb/h); but after purification, 14 of the 24 strategies performed better than this strategy, the best of which lost at only 1.86 mb/h. The strategy that benefitted the most from purification increased its winnings by 68%. In the instances where purification did not help, we observed that at least one of the players used the worst abstraction in our selection – one that does not look at its initial card.

From these experiments, we conclude that purification tends to improve the performance of an abstract equilibrium strategy against an unadaptive equilibrium opponent in Leduc Hold'em. Experiments on thresholding had similar results, but interestingly we observed that all the strategies that were improved by purification obtained their maximum performance when completely purified.

5. TEXAS HOLD'EM

In the 2010 AAAI computer poker competition, the CMU team (Ganzfried, Gilpin, and Sandholm) submitted bots that used both purification and thresholding in the two-player no-limit Texas Hold'em division. Both bots use the same abstraction and equilibrium-finding algorithms; they differ only in their reverse-mapping algorithms. Tartanian4-IRO (IRO) uses thresholding with a threshold of 0.15, while Tartanian4-TBR (TBR) uses purification.

The two-player no-limit competition consisted of two sub-competitions with different scoring rules. In the *instant-runoff* scoring rule, each pair of entrants plays against each other, and the bot with the worst head-to-head record is eliminated. This procedure is continued until only a single bot remains. The other scoring rule is known as *total bankroll*. In this competition, all entrants play against each other and are ranked in order of their total profits. While both scoring metrics serve important purposes, the total bankroll competition is considered by many to be more realistic, as in many real-world multiagent settings the goal of agents is to maximize total payoffs against a variety of opponents.

We submitted IRO to the instant-runoff competition and TBR to the total bankroll competition; the bots finished third and first respectively. Although the bots were scored only with respect to the specific scoring rule and bots submitted to that scoring rule, all bots were actually played against each other, enabling us to compare the performances of IRO and TBR.

One observation is that TBR actually beat IRO when they played head-to-head (at a rate of 80 milli big blinds per hand). Furthermore, TBR performed better than IRO against every single opponent except for one. Even in the few matches that the bots lost, TBR lost at a lower rate than IRO. Thus, even though TBR uses less randomization and is perhaps more exploitable in the full game, the opponents submitted to the competition were either not trying or not able to find successful exploitations. Additionally, TBR would have still won the total bankroll competition even if IRO were also submitted.

These results show that purification can in fact yield a big gain over thresholding (with a lower threshold) even against a wide variety of realistic opponents in very large games.

6. CONCLUSION

We presented two new reverse-mapping algorithms for large games: purification and thresholding. Both of these algorithms consistently improve performance over a wide variety of domains, including random matrix games, Leduc Hold'em, and Texas Hold'em; in fact, purification seems to outperform thresholding in practice.

7. REFERENCES

[1] D. Billings, N. Burch, A. Davidson, R. Holte, J. Schaeffer, T. Schauenberg, and D. Szafron. Approximating game-theoretic optimal strategies for full-scale poker. *IJCAI*, 2003.

[2] A. Gilpin, S. Hoda, J. Peña, and T. Sandholm. Gradient-based algorithms for finding Nash equilibria in extensive form games. *WINE*, 2007. Extended version in *Math. of OR*, 2010.

[3] A. Gilpin and T. Sandholm. A competitive Texas Hold'em poker player via automated abstraction and real-time equilibrium computation. *AAAI*, 2006.

[4] A. Gilpin, T. Sandholm, and T. B. Sørensen. A heads-up no-limit Texas Hold'em poker player: Discretized betting models and automatically generated equilibrium-finding programs. *AAMAS*, 2008.

[5] D. Schnizlein, M. Bowling, and D. Szafron. Probabilistic state translation in extensive games with large action sets. *IJCAI*, 2009.

[6] K. Waugh, D. Schnizlein, M. Bowling, and D. Szafron. Abstraction pathologies in extensive games. *AAMAS*, 2009.

[7] M. Zinkevich, M. Bowling, M. Johanson, and C. Piccione. Regret minimization in games with incomplete information. *NIPS*, 2007.

[1]In order to decrease the number of samples required to obtain statistical significance, we ignored games Σ for which the abstraction Σ' contained a pure strategy equilibrium, as purification and abstraction perform identically.