
Computing Best-Response Strategies in Infinite Games of Incomplete Information*

2004-5-24 21:40

Daniel M. Reeves Michael P. Wellman

University of Michigan Artificial Intelligence Lab.

1101 Beal Avenue, Ann Arbor, MI 48109-2110 USA

<http://ai.eecs.umich.edu/people/{dreeves,wellman}/>
{dreeves,wellman}@umich.edu

Abstract

We describe an algorithm for computing best-response strategies in a class of two-player infinite games of incomplete information, defined by payoffs piecewise linear in agents' types and actions, conditional on linear comparisons of agents' actions. We show that this class includes many well-known games including a variety of auctions and a novel allocation game. In some cases, the best-response algorithm can be iterated to compute Bayes-Nash equilibria. We demonstrate the efficacy of our approach on existing and new games.

1 Introduction

Game theory was founded by von Neumann and Morgenstern (1947) to study situations in which multiple agents (players) interact in order to each maximize an objective (payoff) function determined not only by their own actions but also the actions of other players. In a game of incomplete information, the payoffs also depend on information that is private to the individual agents. This information is known as an agent's *type*. We consider one-shot games—games in which an agent chooses a single action based only on knowledge of its own type, the payoff function, and the distribution from which types are drawn. We assume that the available actions, payoff function, and type distributions are common knowledge.¹ A strategy in this context is a mapping from the set of types to the set of actions. The case where these sets are finite (and especially when there are no types—complete information) has been well studied in the computational game theory literature. In Sections 2 and 6 we discuss existing algorithms and software

tools for solving finite games. But for incomplete information games with a continuum of actions available to the agents—known as infinite games—we know of no available algorithms, though many particular infinite games of incomplete information have been solved in the literature.

We have so far left unsaid what we mean by solving a game. For the case of a single agent, a solution is a policy that maximizes the agent's expected utility (payoff). In the case of multiple agents, Nash (1951) proposed a solution concept now known as the *Nash equilibrium*, and proved that for finite games (as long as agents can play mixed strategies—i.e., randomize among actions) such an equilibrium always exists. A Nash equilibrium is a profile of strategies such that each strategy is a best response to the rest of the profile. That is, each agent maximizes its own utility given the strategies played by the others.²

This definition of Nash equilibrium invites an obvious algorithm for finding one: start with a profile of seed strategies and iteratively compute best-response profiles until a fixed-point is reached. This process (when it converges) yields a profile which is a best response to itself and thus a Nash equilibrium.

The notion of Nash equilibrium needs to be generalized slightly for the case of incomplete information games. Another seminal game theorist, Harsanyi (1967), introduced the concept of agent types and used it to define a *Bayesian game*. A Bayesian game is specified by a set of types T , a set of strategies S , a probability distribution F over types, and a payoff function P . Harsanyi defines a *Bayes-Nash equilibrium* (sometimes known as a Bayesian equilibrium) as the simple Nash equilibrium of the non-Bayesian game with set of strategies being the set of functions from T to S , and the payoff function being the expectation of P with respect to F .

*The latest version of this paper, along with a Mathematica implementation of our algorithm, is available at <http://ai.eecs.umich.edu/people/dreeves/equilibria>.

¹A fact is common knowledge (Fagin et al., 1995) if everyone knows it, everyone knows that everyone knows it, ad infinitum.

²The limitations of Nash equilibrium as a solution concept have been well studied in the literature, particularly the problem of what agents will do in the face of multiple equilibria (van Damme, 1983) yet finding Nash equilibria remains fundamental to the analysis of games.

2 Finite Game Approximations

GAMBIT (McKelvey et al., 1992) is a software package incorporating several algorithms for solving finite games (McKelvey and McLennan, 1996). The original 1992 implementation of GAMBIT was limited to normal form games.³ GALA (Koller and Pfeffer, 1997) introduced constructs for schematic specification of extensive-form games, with a solver exploiting recent algorithmic advances (Koller et al., 1996). GAMBIT subsequently incorporated these GALA features, and currently stands as the standard in general finite game solvers.

We employ a standard first-price sealed-bid auction (FPSB) to compare our approach for the full infinite game to a discretized version amenable to finite game solvers. Consider a discretization of (FPSB) with two players, nine types, and nine actions. Both players have a randomly (uniform) determined valuation (t) from the set $\{0, \dots, 8\}$ and the available actions (a) are to bid an integer in the set $\{0, \dots, 8\}$. Let $[p, a; a']$ denote the mixed strategy of performing action a with probability p , and a' with probability $1 - p$. (There happened never to be mixed strategies with more than two actions.) GAMBIT solves this game exactly, finding the following Bayes-Nash equilibrium:⁴

$$\begin{array}{l} t: \quad 0 \ 1 \ 2 \ 3 \ 4 \quad 5 \quad 6 \ 7 \quad 8 \\ a(t): 0 \ 0 \ 1 \ 1 \ 2 \ [.455, 2; 3] \ 3 \ 3 \ [.727, 3; 4] \end{array}$$

This result is indeed close to the unique Bayes-Nash equilibrium, $a(t) = t/2$, of the corresponding infinite game (see Section 5.1). However, it varies asymmetrically, which implies that the discretized version represents a qualitatively different game.

We observe a similar asymmetric divergence in another discrete approximation. Consider the FPSB over a continuous type range, distributed uniformly, but with actions restricted to the form $a(t) = wt$, with w a choice parameter. (This is a very conservative comparison case in that it eliminates all but a class of strategies that includes the known

³Normal form, also known as strategic form, lists explicit payoffs for every combination (profile) of agent strategies. In contrast, *extensive form* is a more compact representation for games of imperfect information in which payoffs are given for sequences of actions but only implicitly for combinations of agent strategies (mappings from private information to actions).

⁴The calculation took 90 minutes of cpu time and 17MB of memory on a machine with four 450MHz Pentium 2 processors and 2.5GB RAM, running Linux kernel 2.4.18. When 2 additional types and actions are added to the discretization, a similar equilibrium results, requiring 23 hours of cpu time and 34MB of memory. GAMBIT's algorithm (Koller et al., 1996) is worst-case exponential in the size of the game tree which is itself size $O(n^4)$ in the size of the type/action spaces. Based on this complexity and our timing results, we conclude that we have reached the limit of what GAMBIT's algorithm can compute.

equilibrium.) For the continuous action range $w \in [0, 1]$, $w = 1/2$ constitutes a unique equilibrium. When w is limited to discrete values on a grid, however, additional equilibria emerge. For example, when only increments of 0.05 are allowed, both $w = 0.45$ and $w = 0.50$ represent symmetric equilibrium policies.

The point of these examples is that whereas solving finite approximations to an infinite game can be instructive, it can also produce misleading results. In Section 5.1 we show how our method immediately finds the solution to the full infinite FPSB game.

3 Infinite Games and Bayes-Nash Equilibria

We consider a class of two-player games, defined by a payoff structure that is analytically restrictive, yet captures many well-known games of interest. Let t denote the subject agent's type and a its action, and t' and a' the type and action of the other agent. We assume that types are scalars drawn from piecewise-uniform probability distributions, and payoff functions take the following form:

$$u(t, a, t', a') = \begin{cases} \theta_1 t + \rho_1 a + \theta'_1 t' + \rho'_1 a' + \phi_1 & \text{if } -\infty < a + \alpha a' < \beta_2 \\ \theta_2 t + \rho_2 a + \theta'_2 t' + \rho'_2 a' + \phi_2 & \text{if } \beta_2 \leq a + \alpha a' \leq \beta_3 \\ \dots & \\ \theta_I t + \rho_I a + \theta'_I t' + \rho'_I a' + \phi_I & \text{if } \beta_I \leq a + \alpha a' \leq +\infty \end{cases} \quad (1)$$

Our class comprises games with payoffs that are linear⁵ functions from own type and action, conditional on a linear comparison between own and other agent action. The form is parametrized by α , β_i , θ_i , ρ_i , θ'_i , ρ'_i , and ϕ_i , where $i \in \{1, \dots, I\}$ indexes the comparison case. (We define $\beta_1 \equiv -\infty$ and $\beta_{I+1} \equiv +\infty$ for notational convenience in our algorithm description below.) The β_i, β_{i+1} regions alternate between open and closed intervals as this is without loss of generality for arbitrary specification of boundary types ('<' vs. '<=') on the regions⁶ and in particular allows the implementation of common tie-breaking rules for sealed-bid auctions.

This parameterized payoff function captures many known mechanisms. Table 1 shows the parameter settings for several such games.

Given a game description in this form, we search for Bayes-Nash equilibria through a straightforward iterative process. Starting with a seed strategy profile (typically based on a myopic or naive strategy such as truthful bidding), we repeatedly compute best-response profiles until reaching a

⁵Functions with constant terms are technically *affine* rather than linear but we ignore that distinction from here on.

⁶For example, to specify u_1 in $(-\infty, a]$, u_2 in (a, b) , u_3 in (b, c) , u_4 at c , and u_5 in (c, ∞) , translate to the alternating open/closed specification: u_1 in $(-\infty, a - \epsilon)$, u_1 in $[a - \epsilon, a]$, u_2 in (a, b) , u_2 in $[b, b]$, u_3 in (b, c) , u_4 in $[c, c]$, u_5 in (c, ∞) .

Game	$\bar{\theta}$	$\bar{\rho}$	$\bar{\theta}'$	$\bar{\rho}'$	$\bar{\phi}$	$\bar{\beta}$	α
FPSB Auction	0, 1/2, 1	0, -1/2, -1	0, 0, 0	0, 0, 0	0	0, 0	-1
Vickrey Auction (2nd Price)	0, 1/2, 1	0, 0, 0	0, 0, 0	0, -1/2, -1	0	0, 0	-1
Vicious Vickrey Auction	$0, \frac{1-k}{2}, 1-k$	$k, k/2, 0$	$-k, -k/2, 0$	$0, \frac{k-1}{2}, k-1$	0	0, 0	-1
Supply Chain Game	-1, -1, 0	1, 1, 0	0, 0, 0	0, 0, 0	0	v, v	1
Bargaining Game -Seller	-1, -1, 0	$1-k, 1-k, 0$	0, 0, 0	$k, k, 0$	0	0, 0	-1
-Buyer	0, 1, 1	0, -k, -k	0, 0, 0	$0, 1-k, 1-k$	0	0, 0	-1
All-Pay Auction	0, 1/2, 1	-1, -1, -1	0, 0, 0	0, 0, 0	0	0, 0	-1
Voluntary Participation Game	0, 1	0, -1/2	0, 0	0, 1/2	$0, -C/2$	C	1
Shared-Good Auction	0, 1/2, 1	$0, -1/4, -1/2$	0, 0, 0	$1/2, 1/4, 0$	0	0, 0	-1

Table 1: Various mechanisms as special cases of the parameterized payoff function in Equation 1. Note that the bargaining game, being asymmetric, is described by two payoff functions. These games are discussed in Section 5.

fixed-point or cycle. A strategy profile that is a best response to itself is, by definition, a Bayes-Nash equilibrium. We show that this process is effective at finding equilibria for certain games in our class. For all games in our class, the best-response algorithm can be used to verify candidate equilibria or ϵ -equilibria found by alternate means.

Our method considers only pure strategies. Although mixed strategies are generally required for infinite as well as finite games, there are broad classes of infinite games for which pure-strategy equilibria are known to exist. For example, Debreu (1952) shows that equilibria in pure strategies exist for infinite games of complete information with action spaces that are compact, convex subsets of a Euclidean space \mathbb{R}^n , and payoffs that are continuous and quasiconcave in the actions. Athey (2001) proves the existence of pure-strategy Nash equilibria for games of incomplete information satisfying a property called the single-crossing condition (SCC). These results encompass many familiar games of economic relevance, including auction games such as FPSB. Our class includes games violating SCC, for which search in the space of pure strategies may not be sufficient. Nevertheless, an ability to compute best responses for the broadest possible games is useful in itself.

The best-response algorithm takes as input a piecewise linear strategy with K pieces ($K - 1$ piece boundaries),

$$s(t) = \begin{cases} m_1 t + b_1 & \text{if } -\infty < t \leq c_2 \\ m_2 t + b_2 & \text{if } c_2 < t \leq c_3 \\ \dots & \\ m_{K-1} t + b_{K-1} & \text{if } c_{K-1} < t \leq c_K \\ m_K t + b_K & \text{if } c_K < t \leq +\infty, \end{cases} \quad (2)$$

represented by the vectors \vec{c} , \vec{m} , and \vec{b} . The piecewise-linear strategy class is sufficiently flexible to approximate any strategy, although of course the complexity of the strategy or quality of the approximation suffers as nonlinearity increases.

A two-player game is *symmetric* if both players face the same payoff function. For symmetric games we start with a single seed strategy, to be repeatedly replaced with the

strategy that responds best to it.⁷ Most of the examples presented in this paper are symmetric and have symmetric pure equilibria. For asymmetric games, we start with a pair of seed strategies, on every iteration computing a best response to each to get the new pair.

4 Existence and Computation of Piecewise Linear Best-Response Strategies

Here we present our algorithm to compute the best response to a given strategy by way of a constructive proof that in our class of games, best responses to piecewise linear strategies are themselves piecewise linear. Intuitively, the proof proceeds by first deriving an algebraic expression for expected utility against the given strategy in terms of the payoff parameters, the distribution parameters, the opponent strategy parameters, own type, and own action. By appropriate partitioning of the action space, the expected utility is expressed as a piecewise polynomial in the agent's action. We then show that the action maximizing that expression (the best response) is a piecewise linear expression of the agent's type. Finally, we establish a bound for the number of pieces in the best-response strategy.

Theorem 1 *Given a payoff function with I regions as in Equation 1, an opponent type distribution with cdf F that is piecewise uniform with J pieces and $J - 1$ piece boundaries $\{d_2, \dots, d_J\}$, and a piecewise linear strategy function with K pieces as in Equation 2, the best-response strategy is itself a piecewise linear function with no more than $2(I - 1)(J + K - 2)$ piece boundaries.*

⁷It can be shown that symmetric games must have symmetric equilibria, although there are some symmetric games with only asymmetric *pure* equilibria (Cheng et al., 2004). Iterating from a single strategy (equivalently, a symmetric profile) does not limit the search to symmetric equilibria since a cycle of length two (given our restriction to two-player games, but regardless of whether the game is symmetric) constitutes an asymmetric equilibrium.

Proof. Finding the best response strategy means maximizing expected utility over the other agent's type distribution. Let T be the random variable denoting the other agent's type.

First, redefine $s(t)$ to include additional redundant boundary points $\{d_2, \dots, d_J\}$ so there are now $J + K - 2$ boundary points of $s(t)$, $\{c_2, \dots, c_{J+K-1}\}$, and

$$s(t) = \begin{cases} m_1 t + b_1 & \text{if } -\infty < t \leq c_2 \\ \dots & \\ m_{J+K-1} t + b_{J+K-1} & \text{if } c_{J+K-1} < t \leq +\infty. \end{cases}$$

We now express the expected utility, factored over the pieces of $s()$ and $u()$, as

$$\begin{aligned} EU(t, a) &= E_T[u(t, a, T, s(T))] = \\ &\sum_{i=1}^I \sum_{j=1}^{J+K-1} E[(\theta_i t + \rho_i a + \theta'_i T + \rho'_i(m_j T + b_j) + \phi_i | \\ &\quad c_j < T \leq c_{j+1}, \beta_i \check{<} a + \alpha(m_j T + b_j) \check{<} \beta_{i+1}] \\ &\quad \cdot \Pr(c_j < T \leq c_{j+1}, \beta_i \check{<} a + \alpha(m_j T + b_j) \check{<} \beta_{i+1}). \end{aligned}$$

(We use the notation " $x_i \check{<} y$ " to denote $x_i < y$ if i is odd and $x_i \leq y$ if i is even.)

If $\alpha m_j = 0$ then the summand reduces to

$$\begin{cases} (\theta_i t + \rho_i a + (\theta'_i + \rho'_i m_j) \frac{c_j + c_{j+1}}{2} + \rho'_i b_j + \phi_i) \\ \quad \cdot (F(c_{j+1}) - F(c_j)) & \text{if } \beta_i - \alpha b_j \check{<} a \check{<} \beta_{i+1} - \alpha b_j \\ 0 & \text{otherwise.} \end{cases}$$

(The derivation of the cdf $F()$ of a piecewise uniform distribution is in Appendix A.)

For the case of $\alpha m_j \neq 0$, first define $x_{ij}(a) \equiv (\beta_i - \alpha b_j - a)/(\alpha m_j)$ and $y_{ij}(a) \equiv (\beta_{i+1} - \alpha b_j - a)/(\alpha m_j)$ with x and y swapped if $\alpha m_j < 0$. We also introduce $mm(a, b, x) \equiv \min(b, \max(a, x))$.

We consider first the probability term in the summand, rewriting it as

$$\begin{aligned} p_{ij}(a) &\equiv \Pr(\beta_i \check{<} a + \alpha \cdot (m_j T + b_j) \check{<} \beta_{i+1} \ \& \ c_j < T \leq c_{j+1}) \\ &= \Pr(x_{ij}(a) \check{<} T \check{<} y_{ij}(a) \ \& \ c_j < T \leq c_{j+1}) \\ &= F(mm(c_j, c_{j+1}, y_{ij}(a))) - F(mm(c_j, c_{j+1}, x_{ij}(a))). \end{aligned}$$

For the expectation term in the summand, we first define

$$\begin{aligned} \bar{x}y_{ij}(a) &\equiv E[T \mid mm(c_j, c_{j+1}, x_{ij}(a)) \check{<} T \\ &\quad \check{<} mm(c_j, c_{j+1}, y_{ij}(a))] \\ &= \frac{mm(c_j, c_{j+1}, x_{ij}(a)) + mm(c_j, c_{j+1}, y_{ij}(a))}{2}. \end{aligned}$$

We can now express the expected utility, $EU(t, a)$, as

$$\sum_{i=1}^I \sum_{j=1}^{J+K-1} (\theta_i t + \rho_i a + (\theta'_i + \rho'_i m_j) \bar{x}y_{ij}(a) + \rho'_i b_j + \phi_i) \cdot p_{ij}(a). \quad (3)$$

This expression is a piecewise second degree polynomial in a and simply linear in t . Treating it as a function of a , parameterized by t , we can find the boundaries for the polynomial pieces (which will be expressions of t). This is done by setting the arguments of the maxes and mins equal and solving for a , yielding the following four action boundaries for each region $\{\beta_i, \beta_{i+1}\}$ in $u()$ and each region $\{c_j, c_{j+1}\}$ in $s()$:

$$\begin{aligned} c_{j+1} = y_{ij}(a) &\Rightarrow a = \beta_{i+1} - \alpha \cdot (m_j c_{j+1} + b_j) \\ c_j = y_{ij}(a) &\Rightarrow a = \beta_{i+1} - \alpha \cdot (m_j c_j + b_j) \\ c_{j+1} = x_{ij}(a) &\Rightarrow a = \beta_i - \alpha \cdot (m_j c_{j+1} + b_j) \\ c_j = x_{ij}(a) &\Rightarrow a = \beta_i - \alpha \cdot (m_j c_j + b_j) \end{aligned}$$

This yields a total of at most $2(I-1)(J+K-2)$ unique action boundaries. So expected utility is now expressible as a piecewise polynomial in a (parameterized by t) with at most $2(I-1)(J+K-2) + 1$ pieces.

For arbitrary t , we can find the action a that maximizes $EU(t, a)$ by evaluating at each of the boundaries above and wherever the derivative (of each piece) with respect to a is zero. This yields up to $2(I-1)(J+K-2) + 1$ critical points, all simple linear functions of t . Call this set of candidate actions C and the corresponding set of expected utilities $EU(t, C)$. The best-response function can then be expressed, for given t , as $\arg \max_C (EU(t, C))$. This is a "piecewise max" of the linear functions in C , and so it is piecewise linear.

It remains to establish an upper bound on the resulting number of distinct ranges for t . We claim the size of C , $2(I-1)(J+K-2) + 1$, is such an upper bound. To see this, first note that the piecewise max of a set of linear functions must be convex (since, inductively, the max of a line and convex function is convex). It is now sufficient to show that at most one new t range can be added by taking the max of a linear function of t and a piecewise linear convex function of t . Suppose the opposite, that the addition of one line adds two pieces. They cannot be contiguous else they would be one piece. So there must be a piece of the convex function between the two pieces of the line. This means the convex function goes below, then above, then below the line and this violates convexity. Therefore, each line in C adds at most one piece to the piecewise max of C and therefore the piecewise linear best response to $s()$ has at most $2(I-1)(J+K-2) + 1$ pieces and thus $2(I-1)(J+K-2)$ type boundaries. \square

Our algorithm for finding a best response follows this constructive proof. Finding C takes time $O(IJK)$. To actually find the piecewise linear function, $\arg \max_C (EU(t, C))$,

we employ a brute force approach that requires $O((IJK)^2)$ time. First, we find all possible piece boundaries by taking all pairs in C , setting them equal, and solving for t . For each t range we then compute $\arg\max_C(EU(t, C))$ and merge whenever contiguous ranges have the same argmax. As the proof shows, this will yield at most $2(I-1)(J+K-2)$ type boundaries. Thus, we have shown how to find the piecewise linear best response to a piecewise linear strategy in polynomial time. The resulting function is converted to the same strategy representation $(\vec{c}, \vec{m}, \vec{b})$ that the algorithm takes as a seed for the opponent strategy.

5 Examples

Here we consider existing and new games and show that our method for finding best responses can confirm or rediscover known results as well as find previously unknown equilibria.

There are many games not analyzed here to which our approach is amenable, such as the All-Pay auction (both winner and loser pay their bids; encoded in Table 1), incomplete information versions of Cournot or Bertrand games, the War of Attrition (both winner and loser pay the second highest price), and voluntary participation games (agents choose an amount to contribute for a joint good and receive utility based on the sum of both contributions; encoded in Table 1). Our approach is not needed for incentive compatible mechanisms such as the Vickrey auction, but, reassuringly, our algorithm returns the dominant strategy of truthful bidding as a best response to any other strategy in that domain.

5.1 First-Price Sealed-Bid Auction

We consider the first-price sealed-bid auction (FPSB) with types that are drawn from $U[0, 1]$ and the following payoff function:

$$u(t, a, a') = \begin{cases} t - a & \text{if } a > a' \\ (t - a)/2 & \text{if } a = a' \\ 0 & \text{otherwise.} \end{cases}$$

In words, two agents have private valuations for a good and they submit sealed bids expressing their willingness to pay. The agent with the higher bid wins the good and pays its bid, thus receiving a surplus of its valuation minus its bid. The losing agent gets zero payoff. In the case of a tie, a winner is chosen randomly, so the expected utility is the average of the winning and losing utility.

This game can be given to our solver by setting the payoff parameters as in Table 1. The algorithm also needs a seed strategy, for which we can use the default strategy of truthful bidding (always bidding one's true valuation: $a(t) = t$ for $t \in [0, 1]$). This strategy is encoded as $\vec{c} = \langle 0, 1 \rangle$, $\vec{m} = \langle 0, 1, 0 \rangle$, and $\vec{b} = \langle 0, 0, 0 \rangle$ (see Section 3).

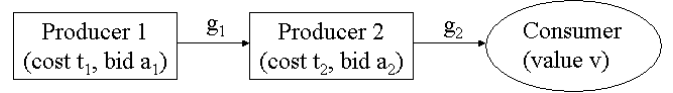


Figure 1: Supply Chain game with two producers in series.

Note that the first and last elements of \vec{m} and \vec{b} are irrelevant as they correspond to type ranges that occur with zero probability. After a single iteration (a fraction of a second of cpu time), our solver returns the strategy $a(t) = t/2$ for $t \in [0, 1]$ which is the known Bayes-Nash equilibrium for this game (McAfee and McMillan, 1987, p709). We find that in fact we reach this fixed point in one or two iterations from a variety of seed strategies—specifically, strategies $a(t) = mt$ for $m > 0$. We approach the fixed point asymptotically (within 0.001 in ten iterations) for seed strategies $a(t) = mt + b$ with $b > 0$.

5.2 Supply-Chain Game

This example derives from our previous work in mechanisms for supply chain formation (Walsh et al., 2000). Consider a supply chain with two producers in series, and one consumer (see Figure 1). Producer 1 has output g_1 and no input. Producer 2 has input g_1 and output g_2 . The consumer—which is not an agent in this model—wants good g_2 . The producer costs, t_1 and t_2 , are chosen randomly from $U[0, 1]$. A producer knows its own cost with certainty, but not the other producer's cost—only the distribution (which is common knowledge). The consumer's value, $v \geq 1$, for good g_2 is also common knowledge.

The producers place bids a_1 and a_2 . If $a_1 + a_2 \leq v$, then all agents win their bids in the auction and the surplus of producer i is $a_i - t_i$. Otherwise, all agents receive zero surplus. In other words, the two producers each ask for a portion of the available surplus, v , and get what they ask minus their costs if the sum of their bids is less than v .

Walsh et al. (2000) propose a strategy for supply-chain games defined on general graphs. In the more general setting, it is the best known strategy (for lack of any other proposed strategies in the literature). For the particular instance of Figure 1, the strategy works out to:

$$a(t) = \begin{cases} t/2 + (v/2 - 1/4) & \text{if } 0 \leq t < v - 1 \\ 3t/4 + v/4 & \text{otherwise.} \end{cases}$$

Our best-response finder proves that this strategy is not a Nash equilibrium and shows how to optimally exploit agents who are playing it.

Figure 2 shows this strategy for the game with $v = (10 - \sqrt{5})/5 \approx 1.55$ (chosen so that there is a 0.9 probability of positive available surplus) along with the best response, as determined by our algorithm and confirmed by Monte Carlo simulation.

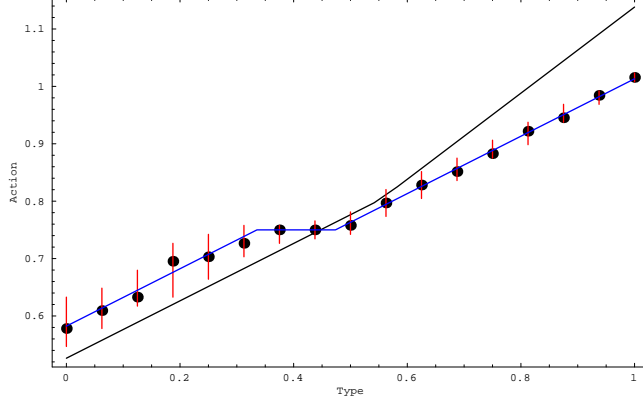


Figure 2: Hand-coded strategy for the Supply Chain game of Section 5.2, along with best response and empirical verification of best response (the error bars for the empirically estimated strategy are explained in Appendix E).

When we perform further best-response iterations it eventually falls into a cycle of period two consisting of the following strategies (where $x = 3/4$):

$$a_1(t_1) = \begin{cases} x & \text{if } t_1 \leq x \\ v & \text{otherwise} \end{cases} \quad (4)$$

$$a_2(t_2) = \begin{cases} v - x & \text{if } t_2 \leq v - x \\ v & \text{otherwise.} \end{cases} \quad (5)$$

The following theorem confirms that we have found an equilibrium, and follows an analogous result (Nash, 1953) for the similar (complete information) *Nash demand game*.

Theorem 2 *Equations 4 and 5 constitute an asymmetric Bayes-Nash equilibrium for the supply-chain game, for any $x \in [0, v]$.*

Proof. Assume producer 2 bids according to Equation 5. Since producer 1 cannot improve its chance of winning with a bid below x , and can never win with a bid above x , producer 1 effectively has the choice of winning with a bid of x or getting nothing. Producer 1 would choose to win at x precisely when $t_1 \leq x$. Hence, (4) is a best response by producer 1. By a similar argument, (5) is a best response by producer 2, if producer 1 follows Equation 4. \square

Following is a more interesting equilibrium, which our solver did *not* find but we were able to derive manually and our best-response finder confirms.

Theorem 3 *When $v \in [3/2, 3]$, the following strategy is a symmetric Bayes-Nash equilibrium for the Supply Chain game:*

$$a(t) = \begin{cases} 2/3v - 1/2 & \text{if } t \leq 2/3v - 1 \\ t/2 + v/3 & \text{otherwise.} \end{cases}$$

Appendix B contains the proof which is essentially an application of our best-response algorithm to the particular game and strategy above. When this strategy is used as the seed strategy for our solver with any particular v , the same strategy is output, thus confirming that it is a Bayes-Nash equilibrium.

5.3 Bargaining Game

The supply chain game is similar to a two-player sealed-bid double auction, or bargaining game. In this game there is a buyer with value v' and a seller with cost c' , each drawn from distributions that are common knowledge. The buyer and seller place bids and if the buyer's is greater than the seller's, they exchange the good at a price that is some linear combination of the two bids. In the supply-chain example, we can model the seller as producer 1, with $c' = t_1$. Because the consumer reports its true value, which is common knowledge, we can model the buyer as the combination of the consumer and producer 2, with $v' = v - t_2$. However, to make the double auction game isomorphic to our supply-chain example, we need to alter the game so that excess surplus is thrown away instead of shared.

The bargaining game as defined above has been well studied in the literature (Chatterjee and Samuelson, 1983; Leininger et al., 1989; Satterthwaite and Williams, 1989). We consider the special case of the bargaining game where the sale price is halfway between the buy and sell offers, and the valuations are $U[0, 1]$. The payoff function for this game is encoded in Table 1.

The following is a known equilibrium (Chatterjee and Samuelson, 1983) for a seller (1) and buyer (2):

$$a_1(t_1) = 2/3t_1 + 1/4 \\ a_2(t_2) = 2/3t_2 + 1/12$$

Our solver finds this equilibrium after several iterations (with tolerance 0.001) when seeded with truthful bidding.

5.4 Shared-Good Auction

Consider two agents who jointly own an inherently unsharable good and seek a mechanism to decide who should buy the other out and at what price. (For example, two roommates could use this mechanism to decide who gets the better bedroom.⁸) Assume that it is common knowledge that the agents' valuations (types) are drawn from $U[A, B]$. We propose the mechanism

$$u(t, a, a') = \begin{cases} t - a/2 & \text{if } a > a' \\ a'/2 & \text{otherwise} \end{cases}$$

which we chose because it has the property that if players bid their true valuations, the mechanism would allocate the

⁸Thanks to Kevin Lochner who both inspired the need for and helped define this mechanism.

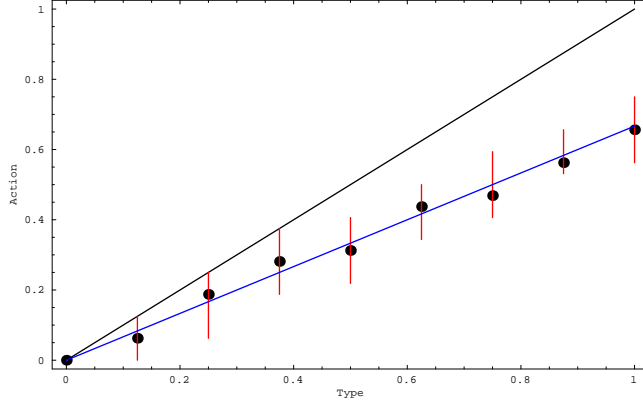


Figure 3: $a(t) = 2t/3$ is a best response to truthful bidding in the shared-good auction with $[A, B] = [0, 1]$. This strategy is in turn a best response to itself, thus confirming the equilibrium.

good to the agent who valued it most and split the surplus evenly between the two (each agent would get a payoff of $t/2$ where t is the larger of the two valuations). The following Bayes-Nash equilibrium also allocates the good to the agent who values it most, but that agent gets up to twice as much surplus as the other agent, depending on the minimum valuation A .

Theorem 4 *The following is a Bayes-Nash equilibrium for the shared-good auction game when valuations are drawn from $U[A, B]$:*

$$a(t) = \frac{2t + A}{3}.$$

Appendix C contains the proof.

Our solver finds this equilibrium exactly (for any specific $[A, B]$) in one iteration from truthful bidding. We confirm the result via simulation as shown in Figure 3.

5.5 Vicious Vickrey Auction

Brandt and Weiß (2001) introduce the following auction game:

$$u(t, a, t', a') = \begin{cases} (1-k)(t-a') & \text{if } a > a' \\ ((1-k)(t-a') - k(t'-a))/2 & \text{if } a = a' \\ -k(t'-a) & \text{otherwise.} \end{cases}$$

It is a Vickrey auction generalized by the parameter k which allows agents to be “antisocial” in the sense of getting disutility from the other agent’s utility. (This might be the case for businesses that are competitors.)

Brandt and Weiß derive an equilibrium only for a complete information version of this game. Our game solver can address the more general incomplete information setting.

Theorem 5 *The following is a Bayes-Nash equilibrium for the vicious Vickrey auction game when valuations are drawn from $U[0, 1]$:*

$$a(t) = \frac{k+t}{k+1}.$$

Appendix D contains the proof.

Our solver finds this equilibrium (for various specific values of k) within several iterations from a variety of seed strategies.

6 Related Work

The seminal works on game theory are von Neumann and Morgenstern (1947) and Nash (1951). There are several modern general texts (Aumann and Hart, 1992; Fudenberg and Tirole, 1991; Mas-Colell et al., 1995) that analyze many of the games in Section 5. Algorithms for solving finite games include the classic Lemke-Howson algorithm (Lemke and Howson, Jr., 1964) for solving bimatrix games (two-agent finite games of complete information). In addition to the algorithms discussed in connection with GAMBIT in Section 2, there has been recent work (La Mura, 2000; Kearns et al., 2001) in algorithms for computing Nash equilibria in finite games by exploiting compact representations of games. Govindan and Wilson (2003, 2002) have recently found new algorithms for searching for equilibria in normal form and extensive form games that are faster than any algorithm implemented in GAMBIT. Blum et al. (2003) have extended and implemented these algorithms in a package called GAMETRACER. Singh et al. (2004) adapt graphical-game algorithms for the incomplete information case, including a class of games with continuous type ranges and discrete actions.

The approach of finding Nash equilibria by iterated best-response, sometimes termed *best-reply dynamics*, dates back to Cournot (1838). A similar approach known as *fictitious play* was introduced by Robinson (1951) and Brown (1951) in the early days of modern game theory. Fictitious play employs a best response, not to the single strategy from the last iteration, but a composite strategy formed by mixing the strategies encountered in previous iterations according to their historical frequency. This method generally has better convergence properties than best-response, but Shapley (1964) showed that fictitious play need not converge in general. Milgrom and Roberts (1991) cast both of these iterative methods as special cases of what they term *adaptive learning* and show that in a class of games of complete information, all adaptive learning methods converge to the unique Nash equilibrium. Fudenberg and Levine (1998) provide a good general text on iterative solution methods (i.e., learning) for finite games. Hon-Snir et al. (1998) apply this approach to a particular auction game with complete information. The *relaxation algo-*

rithm (Uryasev and Rubinstein, 1994), applicable to infinite games, but only complete information games, is a generalization of best-response dynamics that has been shown to converge for some classes of games.

The literature is rife with examples of analytically computed equilibria for particular auction games. For example, Milgrom and Weber (1982) derive equilibria for first- and second-price auctions with affiliated signals. Gordy (1998) finds closed-form equilibria in certain common-value auctions given particular signal distributions.

7 Conclusion

We have presented a proof that best responses to piecewise linear strategies in a class of infinite games of incomplete information are piecewise linear. The proof is constructive and contains a polynomial-time algorithm for finding such best-responses. To our knowledge, this is the first algorithm for finding best response strategies in a broad class of infinite games of incomplete information.

For some games, this best-response algorithm can be iterated to find Bayes-Nash equilibria. It remains a goal to characterize the class of games for which iterated best-response converges. Our method confirms known equilibria from the literature (e.g., auction games such as FPSB and Vickrey), confirms an equilibrium we derive here (in the Supply Chain game), and discovers new equilibria (in the Shared Good auction and an incomplete information Vicious Vickrey auction).

Acknowledgments

Bill Walsh provided the original inspiration for this work—especially its application to the game in Section 5.2. The equilibrium result in that section and its proof is joint work with Bill. Terence Kelly provided many constructive suggestions and discussions. Kevin Lochner helped devise the mechanism in Section 5.4. Preliminary research for this paper was performed during an internship at Hewlett-Packard Labs and the following people at HP helped in the development of the first version our algorithm: Leslie Fine, Kayyut Chen, Evan Kirshenbaum, Jaap Suermondt, and Kemal Guler.

References

Susan Athey. Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69(4):861–890, 2001.

Robert J. Aumann and Sergiu Hart. *Handbook of Game Theory*, volume 1. North-Holland, 1992.

Ben Blum, Christian R. Shelton, and Daphne Koller. A continuation method for Nash equilibria in structured

games. In *Eighteenth International Joint Conference on Artificial Intelligence*, pages 757–764, Acapulco, 2003.

Felix Brandt and Gerhard Weiß. Antisocial agents and Vickrey auctions. In *Eighth International Workshop on Agent Theories, Architectures, and Languages*, volume 2333 of *Lecture Notes in Computer Science*, pages 335–347, Seattle, 2001. Springer.

G. W. Brown. Iterative solution of games by fictitious play. In T. C. Koopmans, editor, *Activity Analysis of Production and Allocation*, pages 374–376. Wiley, New York, 1951.

K. Chatterjee and W. Samuelson. Bargaining under incomplete information. *Operations Research*, 31:835–851, 1983.

Shih-Fen Cheng, Daniel M. Reeves, Yevgeniy Vorobeychik, and Michael P. Wellman. Notes on equilibria in symmetric games. Technical report, University of Michigan, 2004. In Preparation.

A. Cournot. Researches into the mathematical principles of the theory of wealth, 1838. English Edition, ed. N. Bacon (Macmillan, 1897).

Gerard Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences*, 38: 886–893, 1952.

Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.

Drew Fudenberg and David K. Levine. *The Theory of Learning in Games*. MIT Press, 1998.

Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, 1991.

Michael B. Gordy. Computationally convenient distributional assumptions for common-value auctions. *Computational Economics*, 12:61–78, 1998.

Srihari Govindan and Robert Wilson. Structure theorems for game trees. *Proceedings of the National Academy of Sciences*, 99(13):9077–9080, 2002.

Srihari Govindan and Robert Wilson. A global Newton method to compute Nash equilibria. *Journal of Economic Theory*, 110:65–86, 2003.

John Harsanyi. Games of incomplete information played by Bayesian players. *Management Science*, 14:159–182, 320–334, 486–502, 1967.

Shlomit Hon-Snir, Dov Monderer, and Aner Sela. A learning approach to auctions. *Journal of Economic Theory*, 82:65–88, 1998.

Michael Kearns, Michael L. Littman, and Satinder Singh. Graphical models for game theory. In *Seventeenth Conference on Uncertainty in Artificial Intelligence*, pages 253–260, Seattle, 2001.

- Daphne Koller, Nimrod Megiddo, and Bernhard von Stengel. Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior*, 14: 247–259, 1996.
- Daphne Koller and Avi Pfeffer. Representations and solutions for game-theoretic problems. *Artificial Intelligence*, 94(1–2):167–215, 1997.
- Pierfrancesco La Mura. Game networks. In *Sixteenth Conference on Uncertainty in Artificial Intelligence*, pages 335–342, Stanford, 2000.
- W. Leininger, P. Linhart, and R. Radner. Equilibria of the sealed-bid mechanism for bargaining with incomplete information. *Journal of Economic Theory*, 48:63–106, 1989.
- C. E. Lemke and J. T. Howson, Jr. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12(2):413–423, 1964.
- Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, New York, 1995.
- R. Preston McAfee and John McMillan. Auctions and bidding. *Journal of Economic Literature*, 25:699–738, 1987.
- Richard D. McKelvey and Andrew McLennan. Computation of equilibria in finite games. In *Handbook of Computational Economics*, volume 1. Elsevier Science, 1996.
- Richard D. McKelvey, Andrew McLennan, and Theodore Turocy. Gambit game theory analysis software and tools, 1992. <http://econweb.tamu.edu/gambit>.
- Paul Milgrom and John Roberts. Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior*, 3:82–100, 1991.
- Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, 1982.
- John Nash. Non-cooperative games. *Annals of Mathematics*, 2(54):286–295, 1951.
- John Nash. Two-person cooperative games. *Econometrica*, 21:128–140, 1953.
- Julia Bowman Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54:296–301, 1951.
- Mark A. Satterthwaite and Steven R. Williams. Bilateral trade with the sealed bid k -double auction: Existence and efficiency. *Journal of Economic Theory*, 48:107–133, 1989.
- Lloyd S. Shapley. Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, volume 5 of *Annals of Mathematical Studies*, pages 1–28. 1964.
- Satinder Singh, Vishal Soni, and Michael P. Wellman. Computing approximate Bayes-Nash equilibria in tree-games of incomplete information. In *Fifth ACM Conference on Electronic Commerce*, New York, 2004.
- S. Uryasev and R.Y. Rubinstein. On relaxation algorithms in computation of non-cooperative equilibria. In *IEEE Transactions on Automatic Control (Technical Notes)*, volume 39, 1994.
- E. van Damme. Refinements of the nash equilibrium concept. In *Lecture Notes in Economics and Mathematical Systems*. Springer, Berlin and New York, 1983.
- John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1947.
- William E. Walsh, Michael P. Wellman, and Fredrik Ygge. Combinatorial auctions for supply chain formation. In *Second ACM Conference on Electronic Commerce*, pages 260–269, 2000.

A CDF of a Piecewise Uniform Distribution

Given a piecewise uniform distribution with pdf

$$f(x) = \begin{cases} 0 & \text{if } -\infty < x \leq d_2 \\ f_2 & \text{if } d_2 < x \leq d_3 \\ \dots & \\ f_{J-1} & \text{if } d_{J-1} < x \leq d_J \\ 0 & \text{if } d_J < x \leq +\infty \end{cases}$$

it is straightforward to derive the cdf (defining $f_1 \equiv 0$, $f_J \equiv 0$, $d_1 \equiv -\infty$, and $d_{J+1} \equiv +\infty$):

$$F(x) = \begin{cases} f_i \cdot (x - d_i) + \sum_{j=1}^{i-1} f_j \cdot (d_{j+1} - d_j) & \text{if } d_i < x \leq d_{i+1}. \end{cases}$$

The expectation (not needed for our best-response algorithm) is $\sum_{i=0}^{n-1} f_i/2 \cdot (d_{i+1}^2 - d_i^2)$.

B Proof of Theorem 3

We show that the following is a symmetric Bayes-Nash equilibrium when $v \in [3/2, 3]$ (which includes the game with $v = 1.55$ analyzed in Section 5.2):

$$s(t) = \begin{cases} 2/3v - 1/2 & \text{if } t \leq 2/3v - 1 \\ t/2 + v/3 & \text{otherwise} \end{cases} \quad (6)$$

The expected utility (surplus) for agent type t bidding a

against agent type $T \sim U[0, 1]$ playing (6) is:

$$\begin{aligned} EU(t, a) &= E_T[u(t, a, T, s(T))] \\ &= (a - t) \Pr(a + s(T) \leq v) \\ &= (a - t) [\Pr(T \leq 2/3v - 1) \\ &\quad \cdot \Pr(a + 2/3v - 1/2 \leq v \mid T \leq 2/3v - 1) \\ &\quad + \Pr(T > 2/3v - 1) \\ &\quad \cdot \Pr(a + T/2 + v/3 \leq v \mid T > 2/3v - 1)]. \end{aligned}$$

We consider three cases, conditioned on the range of a .

Case 1: $a \leq 2/3v - 1/2$.

The expected surplus is:

$$\begin{aligned} EU(t, a) &= (a - t) [(2/3v - 1) + (2 - 2/3v)] \\ &= a - t. \end{aligned}$$

Since the expected surplus is monotonic in a , the optimal bid is found at the upper boundary, namely, $a = 2/3v - 1/2$, giving us $EU(t, a) = 2/3v - 1/2 - t$.

Case 2: $a \in [2/3v - 1/2, v/3 + 1/2]$.

The expected surplus is:

$$\begin{aligned} EU(t, a) &= (a - t) [(2/3v - 1) + (2/3v - 2a + 1)] \\ &= (a - t)(4/3v - 2a). \end{aligned} \quad (7)$$

Equation 7 is maximized at $a = t/2 + v/3$. With $v \in [3/2, 3]$, we need consider whether this point occurs below the lower boundary of a . This gives us two cases. If $t > 2/3v - 1$ then the maximum of (7) lies in the range of bids we consider here, and the best response is $a_1 = t/2 + v/3$, giving us $EU(t, a) = (3t - 2v)^2/18$. If $t > 2/3v - 1$ then the maximum occurs at the lower boundary, and the best response is $a = 2/3v - 1/2$, giving us $EU(t, a) = 2/3v - 1/2 - t$.

Case 3: $a > v/3 + 1/2$.

The expected surplus is always zero in this range.

The expected surplus of Case 2 is always positive, and always at least as high as Case 1, hence it must specify the best-response policy. But Case 2 gives us the bidding policy specified by (6) when $v \leq 3$.

We have shown that (6) is a best response to itself for all ranges of $v \in [3/2, 3]$, hence it is a symmetric Bayes-Nash equilibrium. \square

C Proof of Theorem 4

We show that the following is a Bayes-Nash equilibrium of the auction game from Section 5.4:

$$a(t) = \frac{2t + A}{3} \quad (8)$$

Given that the other agent employs the strategy in (8) for its type $T \sim U[0, 1]$, we show that (8) is a best response. The expected utility (surplus) for an agent playing bid a against an agent bidding according to (8) is

$$\begin{aligned} EU(t, a) &= E \left[u \left(t, a, \frac{2T + A}{3} \right) \right] \\ &= (t - a/2) \Pr(a > s(T)) \\ &\quad + E \left[1/2 \cdot \frac{2T + A}{3} \mid a < s(T) \right] \Pr(a < s(T)) \\ &= (t - a/2) \Pr\left(\frac{3a - A}{2} > T\right) \\ &\quad + E \left[\frac{2T + A}{6} \mid \frac{3a - A}{2} < T \right] \cdot \Pr\left(\frac{3a - A}{2} < T\right). \end{aligned} \quad (9)$$

We consider two cases, conditioned on the range of a .

Case 1: $a > (2B + A)/3$.

The expected surplus is:

$$EU(t, a) = t - a/2$$

which implies an optimal action at the left boundary, $a_1^* = (2B + A)/3$.

Case 2: $a < (2B + A)/3$.

In this case, both the probabilities in (9) are nonzero and the expected surplus is:

$$\begin{aligned} EU(a) &= (t - a/2) \frac{\frac{3a - A}{2} - A}{B - A} \\ &\quad + 1/6 \left(2 \cdot \frac{\frac{3a - A}{2} + B}{2} + A \right) \cdot \frac{B - \frac{3a - A}{2}}{B - A} \end{aligned}$$

which implies an optimal action where the derivative with respect to a is zero, $a_2^* = (2t + A)/3$.

Comparing the expected surpluses of the candidate action functions,

$$EU(a_2^*) - EU(a_1^*) = \frac{(B - t)^2}{2(B - A)} > 0$$

therefore $a_2^* = (2t + A)/3$ is a best response to itself and therefore a Bayes-Nash equilibrium. \square

D Proof of Theorem 5

We show that the following is a Bayes-Nash equilibrium of the Vicious Vickrey auction described in Section 5.5:

$$a(t) = \frac{k + t}{k + 1}. \quad (10)$$

Given that the other agent bids according to (10) for its type $T \sim U[0, 1]$, we show that (10) is a best response. The expected utility for bidding a against an agent playing (10) is

$$\begin{aligned}
 EU(t, a) &= E \left[-k(T - a) \mid a < \frac{T+k}{1+k} \right] \cdot \Pr(a < \frac{T+k}{1+k}) \\
 &\quad + E \left[(1-k)(t - \frac{T+k}{1+k}) \mid a > \frac{T+k}{1+k} \right] \\
 &\quad \cdot \Pr(a > \frac{T+k}{1+k}) \\
 &= -k \left(\frac{1+a(1+k)-k}{2} - a \right) (1-a(1+k)+k) \\
 &\quad + (1-k) \left(t - \frac{1/2 \cdot a(1+k) - 1/2 \cdot k + k}{1+k} \right) \\
 &\quad \cdot (a(1+k) - k)
 \end{aligned} \tag{11}$$

Tie-breaking cases can be ignored here since they occur with zero probability given the type distribution and form of the opponent strategy. We can now simply check the first-order condition of (11) to find the maximizing a , which is the best response:

$$a^* = \frac{k+t}{k+1}$$

Since (10) is a best response to itself, it is a symmetric Bayes-Nash equilibrium.

E Monte Carlo Best-Response Finder

Here we describe the simulation technique we used to sanity-check equilibria found by our analytic best-response solver. The approach takes as input an arbitrary payoff function from types and actions of all agents to real-valued payoffs. It also takes an arbitrary strategy function, which is a one-dimensional function from type to action. The strategies and payoffs are represented as arbitrary Mathematica functions. Additionally, it takes an arbitrary probability distribution from which types are drawn. From these three inputs, it computes an empirical best-response function. Figure 2 shows an example of our Monte Carlo method finding the best response to a particular strategy (see Section 5.2).

At the core of the method for empirically generating best-response strategies is a simulator that takes the given payoff function, other agent strategy, other agent type distribution, and a particular own type and action. The simulator then repeatedly samples from the other agent type distribution, computing for each sampled type the action according to the known other agent strategy. The resulting payoff is then computed for the given own type and action by evaluating the payoff function. Sample statistics for these payoffs are then recorded and the process is repeated for every

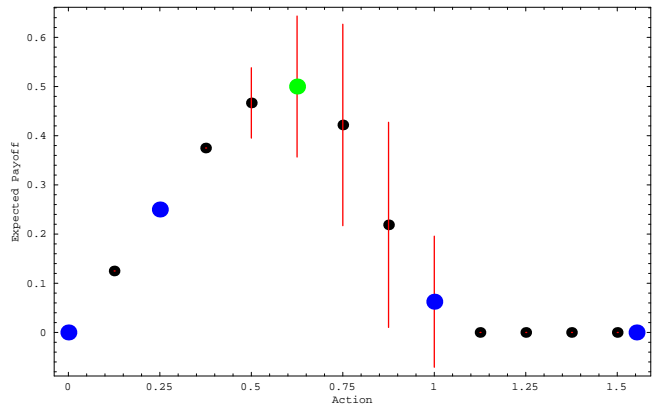


Figure 4: Empirically estimating the best action for a given type (0) against a given strategy (see Figure 2). After < 20 simulations per action the best action is determined, based on the 95% confidence intervals, to be between 0.25 and 1. The maximum likelihood best action is 0.62 (after 100 total samples) and the actual best action is 0.582.

combination of own type and action, down to a specified granularity. However, certain shortcuts are taken to avoid needless simulation. First, when simulating different actions for a given type, a confidence bound is continually computed using the sampling statistics for the various actions sampled so far. Further simulation is limited to the range of actions within the confidence bound (designated in Figure 4 by the two larger dots at $a = 0.25$ and $a = 1$). The confidence bounds are determined by performing mean difference tests on pairs of expected payoff sample statistics and considering any actions that fail at the 95% level to be within the confidence bound for best action. It is these confidence bounds that the error bars represent in Figure 2. The amount of simulation for each type-action pair is dynamically determined based on the confidence intervals for the expected payoffs. And within the confidence bound of possible actions, the confidence intervals are compared to prioritize the computation for different actions.

Figure 4 shows the Monte Carlo best-response method sampling possible actions for a specific type. It dynamically limits the actions that it samples from by comparing expected payoffs and finding an interval likely to contain the best action.