When is it Best to Best-Respond?

NOAM NISAN
School of Computer Science and Engineering, The Hebrew University of Jerusalem, and Google Tel Aviv.
and
MICHAEL SCHAPIRA
Computer Science Department, Princeton University
and
GREGORY VALIANT
Computer Science Department, UC Berkeley
and
AVIV ZOHAR
Microsoft Research, Silicon Valley

Often, in both computerized settings and economics settings, the prescribed behavior for participants is to repeatedly "best respond" to each others' actions. We aim to understand when such myopic "local rationality" is also "globally rational", *i.e.*, when is it best for a player, given that the others are repeatedly best-responding, to also repeatedly best-respond?

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: Economics

General Terms: Algorithms, Economics, Theory

Additional Key Words and Phrases: Mechanism Design, Best-Response Dynamics

1. MOTIVATION

The basic object of study in game theory and economics is the *equilibrium*. However, an equilibrium is a static concept which often abstracts away the question "*How is an equilibrium reached?*". A truly satisfactory answer to this question is one where each player follows a simple and natural rule of behavior yet, mysteriously, the resulting dynamics reach an equilibrium point. The (arguably) most compelling examples of such phenomena involve *repeated best-response dynamics*. Under repeated best-response dynamics, each player repeatedly selects the best (locally-optimal) response to what others are currently doing. This process goes on "for a while" until it "converges" to what must be a (pure Nash) equilibrium. Unfortunately, the convergence of repeated best-response dynamics is not guaranteed in general, and is the subject of extensive research.

Our focus is on a different question, which has received little attention so far: "*Is repeated best-response rational?*". We consider games in which repeated best-response dynamics *do* converge to an equilibrium and study the incentive properties of this process; Is it *rational* for players to *repeatedly* best-respond? Can a long-

Authors' addresses: noam@cs.huji.ac.il, ms7@cs.princeton.edu, gvaliant@eecs.berkeley.edu, avivz@microsoft.com

sighted player improve, in the long run, over this repeated *myopic* optimization? Our general aim is to identify interesting classes of games for which repeated best-response dynamics are *incentive compatible*. Intuitively, repeated best-response dynamics are incentive-compatible if, when all other players are repeatedly best-responding, a player has incentive to do the same.

2. OUR FRAMEWORK

Our first contribution is a formal framework for reasoning about incentives for repeated best-response. In our framework, each player aims to maximize a *private* utility function. We consider dynamic interactions where players take turns selecting strategies. At each (discrete) time step, a single player selects a strategy. The prescribed behavior for every player is to repeatedly choose his best response (given his private utility function) to the current strategies of the others. We stress that, in our framework, there is nothing preventing a player from deviating from this prescribed behavior; a player can, for instance, select strategies based on complex computations that take into account the entire history of interaction.

Defining incentive compatibility in our setting involves many intricacies. We opt to focus on a very general notion of incentive compatibility that, we believe, captures essentially any variant that the reader may desire. Intuitively, repeated best-response is incentive-compatible in our framework if, when all other players are repeatedly best-responding, a player has incentive to do the same, *i.e.*, no player can gain from unilaterally deviating from repeated best-response. We refer the reader to [Nisan et al. 2011b] for a detailed exposition of our framework.

2.1 When is Best-Responding Not Best?

While at first glance, it might seem that the existence of a *unique* pure Nash equilibrium (PNE) to which repeated best-response dynamics are guaranteed to converge implies the incentive-compatibility of repeated best-response dynamics, this intuition is false.

	C	D
A	2,1	0,0
B	3,0	1,2

Fig. 1. A game for which repeated best-response dynamics are not incentive compatible.

Observe that in this small game, (B, D) is the unique PNE and every sequence of best responses converges to it. Consider the scenario that the starting point is the strategy profile (A, C), and the column player does follow the prescribed behavior, *i.e.*, he repeatedly best-responds. Clearly, the row player's *local* improvement from (A, C) to (B, C) will lead to the column player moving to (B, D). Hence, the row player can do better by looking ahead, not moving from (A, C), and thus "getting stuck" at (A, C), which he strictly prefers to the unique pure Nash (B, D). Repeated best-responding is therefore *not* incentive compatible in this game which is strictly-dominance-solvable, is a potential game, and has a unique and Pareto-optimal PNE. So, what traits must a game have for best-response dynamics to be incentive compatible?

2.2 General Result and Applications

Our main results are identifying a class of games for which repeated best-response is incentive compatible and exhibiting several interesting environments that fall within this class. We formally present our class of games in [Nisan et al. 2011b]. Intuitively, games that belong to our class have the following property. When each player i considers the game after the other players have already iteratively eliminated dominated strategies, he can already tell that he can do no better than the outcome that is reached via repeated best-response.

Our main, and quite easy to prove, general theorem is that in this class of games, for every initial strategy profile and every (finite or infinite) order of player activations with "sufficiently many" rounds (where a round is a sequence of consecutive time steps in which each player gets to act at least once), repeated best-response is incentive compatible for every player.

We identify several environments of interest that fall within our class of games: (1) stable-roommates matchings (*e.g.*, interns-hospitals, correlated markets); (2) economic cost-sharing; (3) models of routing on the Internet; and (4) models of congestion control on the Internet. We prove that the desired outcome in each environment translates to a PNE. Thus, the above general result establishes that best-response dynamics are incentive compatible and converge to the desired outcome in all of the above settings.

2.3 Present and Future Research

We view this work as a first step towards a more general research agenda. While convergence to equilibrium of "locally-rational" dynamics, *e.g.*, repeated best-response, fictitious play and regret minimization, has been extensively studied, little attention has been given to the question of when such locally-rational dynamics are also "globally rational". Here, we tackle this question in the context of repeated best-response and PNE. We believe that the examination of other dynamics and other kinds of equilibria is an interesting direction for future research.

Our results for repeated best-response dynamics establish *sufficient* conditions for repeated best-response to be incentive compatible. Providing complete *characterizations* of conditions that imply incentive compatibility, *i.e.*, identifying conditions that are both sufficient *and* necessary, remains an open question.

We have presented our results for a very strong notion of incentive compatibility. We believe that considering more restrictive notions (*e.g.*, incentive compatibility in expectation) is of interest. Indeed, in [Nisan et al. 2011a] we present several such results for commerce environments.

REFERENCES

Nisan, N., Schapira, M., Valiant, G., and Zohar, A. 2011a. Best-response auctions. In *EC '11: Proceedings of the 12th symposium on Electronic Commerce*.

Nisan, N., Schapira, M., Valiant, G., and Zohar, A. 2011b. Best-response mechanisms. In *ICS '11: Proceedings of the 2nd symposium on Innovations in Computer Science*.