Approximate common beliefs in rationality in dynamic games

Aleix Garcia-Galocha Jaromír Kovářík Peio Zuazo-Garin

Discussion
November 25, 2025

Enrico Mattia Salonia

Epistemic game theory in 30 seconds

Game theory = individual preferences (tastes and beliefs) + equilibrium analysis.

Equilibrium = rationality (expected utility) + common belief in rationality.

Behavioural economics introduces "behavioural heuristics" into games.

How these heuristics relate to classical preferences + equilibrium is often unclear.

Epistemic game theory takes these sums **very seriously**.

Motivation of this paper

Dynamic refinements (e.g. subgame perfection) use the chronological structure of games.

Refinements are equivalent to having specific beliefs about others' rationality.

Classical refinements often mispredict behaviour in classic experiments.

Question: how robust are classical dynamic refinements to small doubts about rationality?

Framework and concepts

Finite multistage games with observable actions.

Players hold beliefs over others' strategies.

Dynamic rationalizability: weak (initial), backward (induction), strong (forward induction).

Introduce p-belief: an event is believed with probability at least $p < 1$. (Monderer & Samet, 1989)

Define weak, backward, strong p-rationalizability.

Main results: rationalizability

For $p = 1$: p-versions coincide with classical notions.

For any $p < 1$: strong and backward p-rationalizability **collapse to weak (initial) rationalizability**.

Interpretation: even tiny doubts about higher-order rationality destroy the refinement power of dynamic reasoning.

Behavioural implications

With $p < 1$, dynamic refinements lose bite; predictions become essentially static.

Centipede game: cooperation in all but the last stage is p-rationalisable.

Finitely repeated Prisoner's Dilemma: all threshold strategies except "cooperate in the final round" are p-rationalisable.

Small doubts about rationality can rationalise non-inductive play in these games.

But I add: for any value of p you could probably rationalise many things!

Comments

Results induce a conceptual dilemma:

- Better dynamic refinements we did not develop yet? (Siniscalchi, 2022)
- Is the notion of p-belief in this paper inadequate?
- (**Pessimistic**): there is no hope.

About the definition of p-belief: it is quite weak.

If you impose p-belief at each node you have a tension with Bayesian updating.

Anything in between?

- Only on first order beliefs;
- Only locally but not on continuation play;
- Might choose how to solve trade-off with Bayesian updating.

References

Monderer, D., & Samet, D. (1989). Approximating common knowledge with common beliefs. *Games and Economic behavior*, *1*(2), 170–190.
Siniscalchi, M. (2022). Structural Rationality in Dynamic Games. *Econometrica*, *90*(5), 2437–2469.