Sensorimotor Games: Dynamical Models and Experiments

Benjamin Chasnov, Lillian Ratliff, Samuel Burden, Electrical & Computer Engineering, UW, Seattle, WA.
{bchasnov, ratlifflj, sburden}@uw.edu

COMPUTATIONAL NEUROSCIENCE CENTER

UNIVERSITY of WASHINGTON

Model: Learning is an optimization process

- Reward: $R : \Theta \to \mathbb{R}$
- Variable: $\theta \in \Theta$

Optimality: $R(\theta^*) \geq R(\theta), \theta \in \Theta_{x^*}$ (locally)

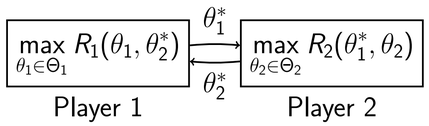
$$\max_{\theta \in \Theta} R(\theta)$$

Single objective

Game (two players):

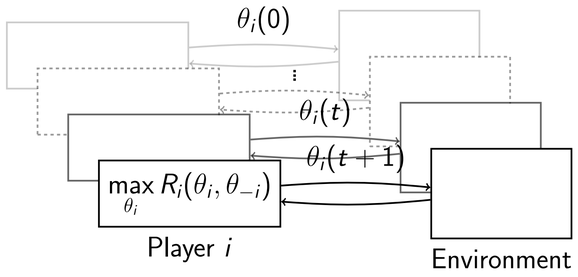
- Game: $\mathcal{G} = (R_1, R_2)$.
- Reward: $R_i : \Theta_1 \times \Theta_2 \to \mathbb{R}, \ i \in [1, 2]$
- Variables: $(\theta_1, \theta_2) \in \Theta_1 \times \Theta_2$.

Optimality:
$R_1(\theta_1^*, \theta_2^*) \geq R_1(\theta_1, \theta_2^*), \ \theta_1 \in \Theta_1,$
$R_2(\theta_1^*, \theta_2^*) \geq R_2(\theta_1^*, \theta_2), \ \theta_2 \in \Theta_2$ (locally).



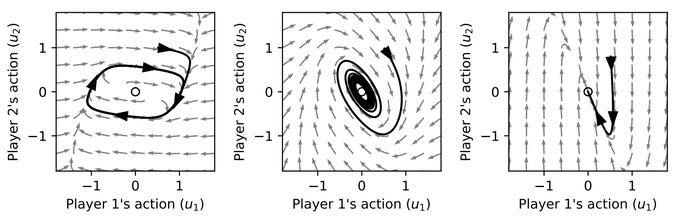
Dynamics of learning process

Observe learning $\theta_i(0), \cdots, \theta_i(t)$ over time...



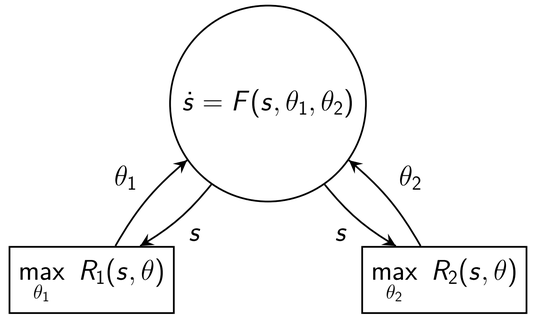
Prediction: periodic orbits and spurious attractors

- Vector field of $\omega \equiv (D_1 R_1, D_2 R_2)$ with rewards
- $R_1(\theta_1, \theta_2) = \frac{1}{4}\theta_1^4 - \frac{1}{2}\theta_1^2 - \theta_1\theta_2$
- $R_2(\theta_1, \theta_2) = \frac{1}{2}\theta_2^2 + 2\theta_1\theta_2$



A Dynamical Systems Perspective

A coupled set of dynamics describes (multi-agent) learning in environments.



- parameters $\theta = (\theta_1, \theta_2)$
- shared state s evolves via dynamics F.

Theory: Stability and Convergence Guarantees

- Asymptotic convergence: as $t \to \infty$, does $u(t) \to u^*$?
- Finite-time bounds: for $t \geq T$, what is $\|u(t) - u^*\|$?
- Regret/no-regret learning: "best action *in hindsight*."

Solution Concepts: Differential Notions

Twice-continuously differentiable objective $c \in \mathcal{C}^2$.
Local optimality of $u^* = (\theta_1^*, \theta_2^*)$ [1]:
First player's condition

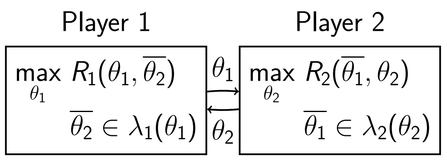
$$D_1 R_1(\theta_1^*, \theta_2^*) = 0, \ D_1^2 R_1(\theta_1^*, \theta_2^*) < 0$$

Second player's condition

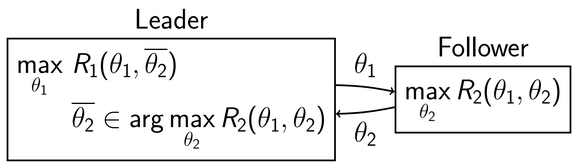
$$D_2 R_2(\theta_1^*, \theta_2^*) = 0, \ D_2^2 R_1(\theta_1^*, \theta_2^*) < 0$$

Novel Framework: Bounded Rationality

Players form *internal models* about others [2]:



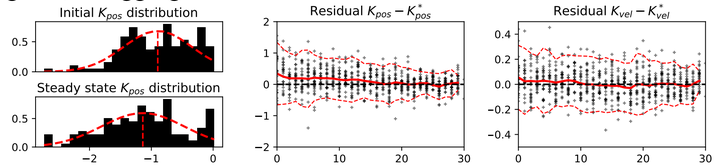
For example, a Stackelberg game can be represented as



Experiments: Learning to Integrate

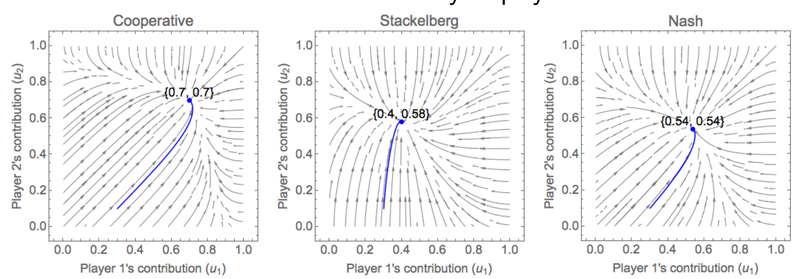
Operating a double integrator model $\dot{s} = (A + B\theta_t)s$. How do subjects learn θ_t over time? [2]

We observe subjects' learning curve for feedback and feedforward gains, in aggregate.



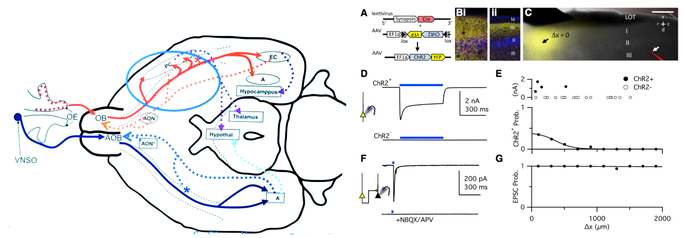
Experiments: Computational Predictions

- Individual contribution to public goods.
- Utilities: $-R_i(\theta_i, \theta_{-i}) = (I_i - \theta_i)^{\alpha_i}(\theta_i + \theta_{-i})^{1-\alpha_i}$
- Internal models dictate the hiarchy of play.



Experiments: Recurrent Circuitry in Piriform Cortex

Future work: modeling neurons in the piriform cortex "voting" for consensus on odors. [3]



References

[1] Ratliff, Burden, Sastry. *Characterization and computation of local nash equilibria in continuous games.* 2013.

[2] Chasnov, Yamagami, Parsa, Ratliff, Burden. *Experiments with sensorimotor games in dynamic human/machine interaction.* 2019.

[3] Franks, Russo, Sosulski, Mulligan, Siegelbaum, Axel. *Recurrent Circuitry Dynamically Shapesthe Activation of Piriform Cortex.* 2011.