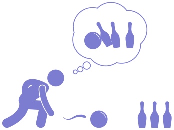
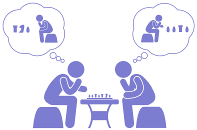
# Causality and Causal Misperception in Dynamic Games

Sungmin Park

The Ohio State University

October 12, 2024

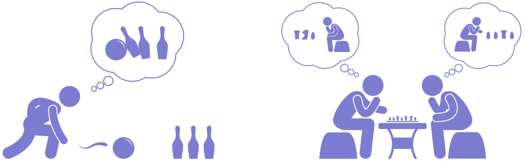




# What I do

**Question** How should we capture players' misperceptions about causality in extensive-form games?

**Answer** Let each player best respond to a belief about Nature's and other players' strategies while requiring that belief be observationally equivalent to how they actually play



**"Observation-consistent expectations" (OCE) equilibrium**

# Motivation

Limited observation of reality $\Rightarrow$ Varying perceptions of causality

People have different perceptions about

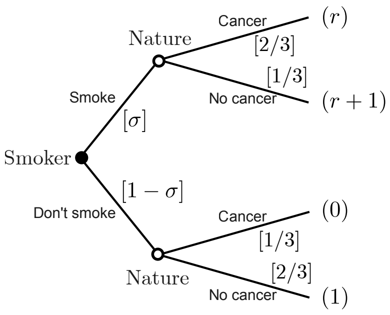
- Effects of smoking tobacco on cancer
- Effects of college degrees on labor market outcomes
- Effects of police presence on violent crimes
- Effects of social media on mental health

Subjects in lab experiments look at the same data and tell different causal narratives (Kendall and Charles, 2022)

$\Rightarrow$ **Q:** What is a useful solution concept that maintains sharp predictions while relaxing rational expectations?

# Simplest example

- Smoker chooses to smoke ($s = 1$) or not ($s = 0$).
  - If he smokes, Nature gives him cancer with prob $\pi_1 = 2/3$.
  - If not, Nature gives him cancer with prob $\pi_0 = 1/3$.
- He gets $r < \frac{1}{3}$ if he smokes and loses $1$ if he gets cancer.
- Smoker's **strategy** is the prob $\sigma \in [0, 1]$ of smoking.
- Smoker's **belief** is $\beta = (\beta_0, \beta_1)$ where $\beta_s$ is the subjective probability of getting cancer given $s$.




**Smoker's Problem**

$\Rightarrow$ Under rational expectations, one shouldn't smoke because the causal effect of smoking on cancer ($\frac{2}{3} - \frac{1}{3} = \frac{1}{3}$) is larger than the reward $r$

# Observation-consistent expectations

## Definition

Given strategy $\sigma \in [0,1]$, an **observation-consistent expectation (OCE)** is a belief $\beta \in [0,1]^2$ such that

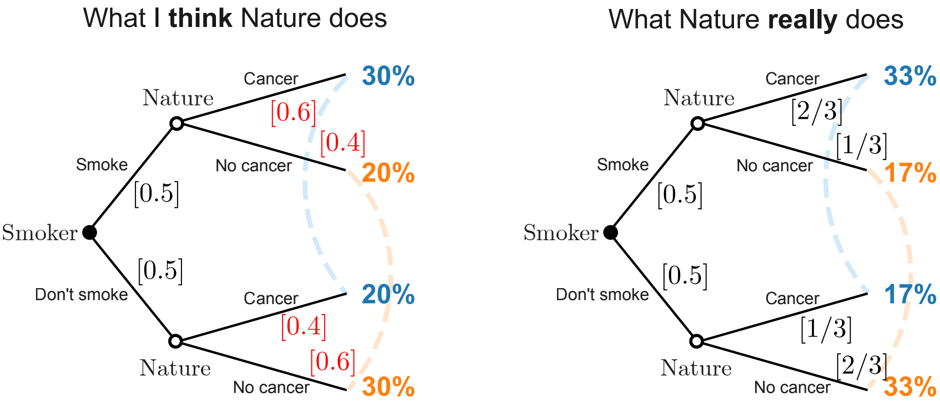
$$\underbrace{\sigma\beta_1 + (1-\sigma)\beta_0}_{\text{perceived marginal prob of cancer}} = \underbrace{\sigma \cdot \frac{2}{3} + (1-\sigma) \cdot \frac{1}{3}}_{\text{actual marginal prob of cancer}}$$

**Interpretation**

- Smoker sees a population of smokers choosing $\sigma$ overall and sees the overall rate of cancer patients, but do not know the conditional probabilities
- What the smoker thinks Nature does $(\beta_0, \beta_1)$ and what Nature really does $(\frac{1}{3}, \frac{2}{3})$ are observationally equivalent

# Illustration of an OCE

Suppose I smoke half of the time ($\sigma = 0.5$).




**Remark**: There are many OCEs.

# MaxEnt OCE

Let $\mathbf{p}(\sigma, \beta)$ be the vector of probabilities over the 4 terminal nodes.

**Definition (MaxEnt OCE)**

Given strategy $\sigma \in (0, 1)$, an OCE $\beta^* \in [0, 1]^2$ is a **maximum entropy (MaxEnt) OCE** if it satisfies

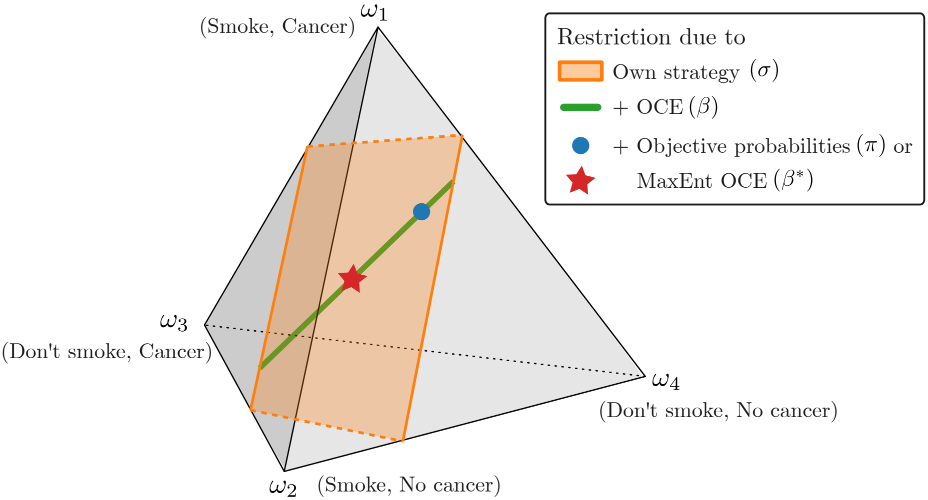
$$\beta^* \in \underset{\beta \in OCE(\sigma)}{\operatorname{argmax}} \; G(\mathbf{p}(\sigma, \beta))$$

where $G(\cdot)$ is the Shannon entropy function.

**Interpretation**

- MaxEnt OCE is the OCE with the least specific information.

# Illustration of MaxEnt OCE



# MaxEnt OCE $\Rightarrow$ correlation neglect

**Claim**

*For every $\sigma \in (0,1)$, the MaxEnt OCE $\beta^*$ satisfies*

$$\beta_0^* = \beta_1^* = (1-\sigma)\cdot\frac{1}{3} + \sigma\cdot\frac{2}{3}.$$

**Meaning** The smoker doesn't think smoking causes cancer

**Intuition** The smoker observes no evidence of dependence between smoking and cancer, so he believes in none.

**General result (Shore and Johnson, 1980; Csiszar, 1991)**
MaxEnt OCE $\Leftrightarrow$ correlation neglect, whenever agents observe only the marginal prob. distribution between two variables

# Definition of equilibrium

**Motivation**

- OCE and MaxEnt OCE take the strategy $\sigma$ as given
- Is there an equilibrium where this $\sigma$ is subjectively optimal?

## Definition

A strategy-belief pair $(\sigma, \beta)$ is an **OCE equilibrium** if

1. Given the belief $\beta$, the strategy $\sigma$ is a best response, and
2. Given the strategy $\sigma$, the belief $\beta$ is an OCE.

An OCE equilibrium $(\sigma, \beta)$ is a **MaxEnt OCE equilibrium** if some $\{(\sigma^k, \beta^k)\}_{k=1}^{\infty} \to (\sigma, \beta)$ where each $\beta^k$ is the MaxEnt OCE given $\sigma^j$.

**Question** Abbreviate the above as MOCEE (pronounced Mochi)?

# Result: OCE equilibria

## Every strategy is rationalizable with some OCE

**Claim**

*Every strategy $\sigma$ has a belief $\beta$ such that $(\sigma, \beta)$ is an OCE equilibrium.*

**Note**: Specifically, the OCE equilibria are

1. $\sigma = 0$, $\beta_0 = \frac{1}{3}$, and $\beta_1 - \beta_0 \geq r$,
2. $\sigma = 1$, $\beta_1 = \frac{2}{3}$, and $\beta_1 - \beta_0 \leq r$, and
3. $\sigma \in (0, 1)$, $\beta_0 = \sigma \cdot (\frac{2}{3} - r) + (1 - \sigma) \cdot \frac{1}{3}$, and $\beta_1 = \sigma \cdot \frac{2}{3} + (1 - \sigma)(\frac{1}{3} + r)$.

**Idea** Because there are many OCEs, there are many OCE equilibria.

# Result: MaxEnt OCE equilibrium

## A sharper prediction

**Claim**

*A strategy-belief pair $(\sigma, \beta)$ is a MaxEnt OCE equilibrium if and only if*

$$\sigma = 1 \quad \textit{and} \quad \beta_0 = \beta_1 = \frac{2}{3}.$$

**Meaning**

- Smoker keeps smoking while thinking that smoking doesn't cause cancer

**Intuition**

- MaxEnt OCE implies correlation neglect, so no other strategy is a best response.

# Generalizing the observational constraint

**Motivation**

- Correlation neglect sounds too naïve. Can we make agents more sophisticated? Yes! Give them better observation

## Definition

Given an **observation constraint matrix** $C$ and strategy $\sigma$, an **OCE** is a belief $\beta$ such that

$$C\mathbf{p}(\sigma, \beta) = C\mathbf{p}(\sigma, \pi).$$

**Examples of $C$:**

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$$

# Generalizing the approach to extensive-form games

**Finite extensive-form game with perfect recall + observational constraint**

- $N$: set of players,
- $H$: set of histories (nodes),
- $\iota$: mapping of non-terminal histories to players,
- $\pi$: probability distribution of Nature's moves,
- $\mathcal{I}$: collection of information sets,
- $u$: payoff function, and
- $C$: **observational constraint matrix**

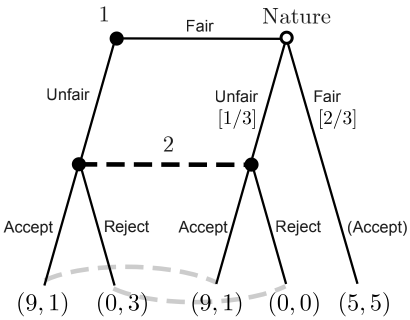
**Theorem**

*Every finite extensive-form game with perfect recall and observational constraint has a MaxEnt OCE equilibrium.*

# Example: Ultimatum-like game with causal misperception

**Manager-Worker game**

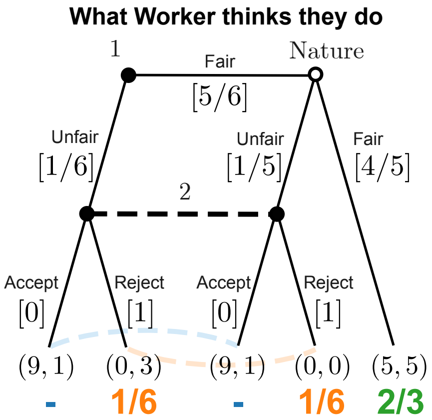
- Manager (Player 1) decides a fair or unfair bonus to Worker (Player 2)
- Even if Manager chooses a fair bonus, Nature might change it to unfair or keep it fair
- If Worker receives fair bonus, he accepts. If not, he either accepts or rejects.
    - He gets a thrill for rejecting an unfair Manager
- Worker doesn't know how likely Manager treats him unfairly in the interim or ex post

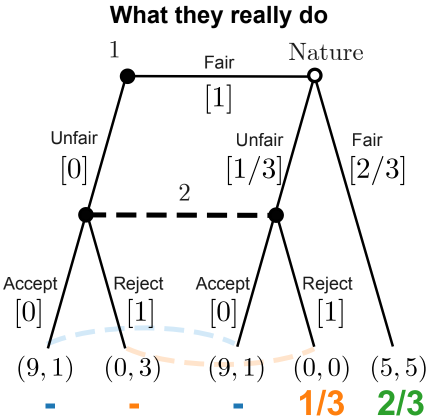



$$C = \begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix}$$

# Unique MaxEnt OCE equilibrium

## Manager always tries to be fair



# Example: A centipede game

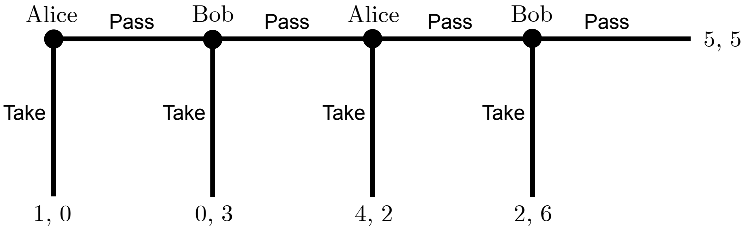



Figure: A four-node centipede game

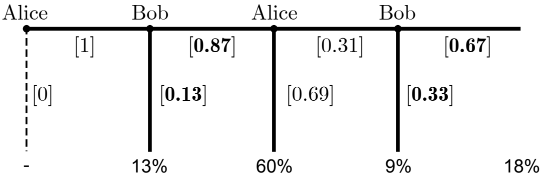
**Claim**

*Let the observational structure be $C = [0\ 1\ 2\ 3\ 4]$. There exists no MaxEnt OCE equilibrium in which Alice Takes immediately.*

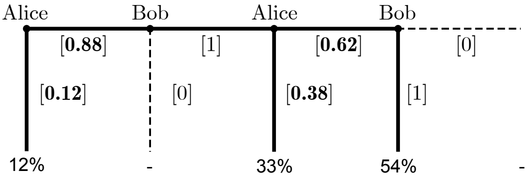
# Unique equilibrium of the centipede game

## Aligned with experimental results such as Healy (2017)

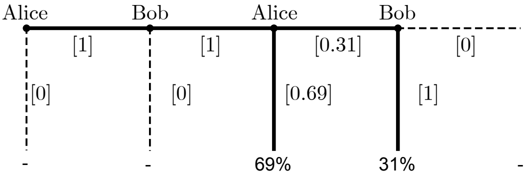
**What Alice thinks they do**

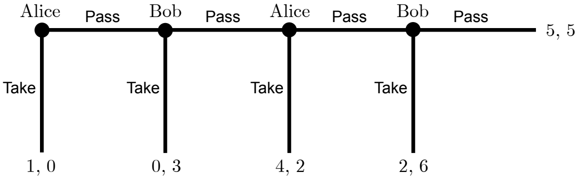



**What Bob thinks they do**




**What they really do**



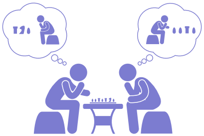



# Takeaways

Use my solution concept if you want to . . .

- allow causal misperception in a dynamic model
- let misperception arise endogenously from observational constraints, and
- want narrow predictions

## Thank you!



# Appendix

# References I


Battigalli, Pierpaolo and Danilo Guaitoli (1988) *Conjectural equilibria and rationalizability in a macroeconomic game with incomplete information*: Università Commerciale L. Bocconi.

Csiszar, Imre (1991) "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *the Annals of Statistics*, 2032–2066.

Fudenberg, Drew and David K Levine (1993) "Self-confirming equilibrium," *Econometrica: Journal of the Econometric Society*, 523–545.

Healy, Paul J (2017) "Epistemic experiments: Utilities, beliefs, and irrational play," *Unpublished manuscript, Ohio State University, Columbus, OH*.

Kendall, Chad W and Constantin Charles (2022) "Causal narratives,"Technical report.

Shore, John and Rodney Johnson (1980) "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Transactions on Information Theory*, 26 (1), 26–37.