CIS 410/510 (Spring 2020): Multi-agent Systems and Game Theory

Lecture 18: Prediction Markets and Scoring Rules

Thanh H. Nguyen

Most slides are by Haifeng Xu, http://www.haifeng-xu.com/cs6501fa19/
Announcement

- Final exam
  - Time: 14:45 on Thursday, June 11, 2020
  - Place: take-home exam
  - Duration: 120 minutes

- Exam review
  - Wednesday, June 3\textsuperscript{rd}, 2020
Outline

- Scoring Rule and its Characterization
- Connection to Prediction Markets
- Manipulations in Prediction Markets
Recap: Prediction Markets

A prediction market is a financial market that is designed for event prediction via information aggregation.

- Payoffs of the traded contract are determined by outcomes of future events.

\[
\begin{align*}
\text{\$1 iff } e_1 & \quad \cdots \quad \text{\$1 iff } e_n \\
\end{align*}
\]

contracts

We design a market maker by specifying the payment for bundles of contracts.
Example: Logarithmic Market Scoring Rule (LMSR [Hanson 03, 06])

- Define value function $(q = (q_1, \ldots, q_n)$ is current sales quantity)

$$V(q) = b \log \sum_{j \in [n]} e^{q_j/b}$$

- Price function

$$p_i(q) = \frac{e^{q_i/b}}{\sum_{j \in [n]} e^{q_j/b}} = \frac{\partial V(q)}{\partial q_i}$$

- To buy $x \in \mathbb{R}^n$ amount, a buyer pays: $V(q + x) - V(q)$
  - Negative $x_i$’s mean selling contracts to MM
  - Negative payment means market maker pays the buyer
  - Market starts with $V(0) = b \log n$
Properties of LMSR

Fact. The optimal amount an expert purchases is the amount that moves the market price to her belief $\lambda$. Her expected utility of purchasing this amount is always non-negative.

- I.e., should purchase $x^*$ such that $\frac{\partial V(q+x^*)}{\partial x^*_i} = \lambda_i$
- Market efficiency

Fact. Worst case market maker loses is $b \log n$ (i.e., bounded).
Outline

- Scoring Rule and its Characterization
- Connection to Prediction Markets
- Manipulations in Prediction Markets
Consider a Simpler Setting

- We (designer) want to learn the distribution of random var $E \in [n]$
  - $E$ will be sampled in the future
- We have no samples from $E$; Instead, we have an expert/predictor who has a predicted distribution $\lambda \in \Delta_n$
- We want to incentivize the expert to truthfully report $\lambda$
Consider a Simpler Setting

- We (designer) want to learn the distribution of random var $E \in [n]$
  - $E$ will be sampled in the future
- We have no samples from $E$; Instead, we have an expert/predictor who has a predicted distribution $\lambda \in \Delta_n$
- We want to incentivize the expert to truthfully report $\lambda$

Example
- $E$ is whether UO will win NCAA title in 2020
- Expert is the UO coach

- Expert’s prediction does not need to be perfect
  - But, better than the designer who knows nothing
- Assume expert will not give you truthful info for free
Idea: “Score” Expert's Report

Will reward the expert certain amount $S(i; p)$ where:

(1) $p$ is the expert’s report (does not have to equal $\lambda$);

(2) $i \in [n]$ is the event realization

*Not like a prediction market yet, but will see later they are related*
Idea: “Score” Expert's Report

Will reward the expert certain amount $S(i; p)$ where:

1. $p$ is the expert’s report (does not have to equal $\lambda$);
2. $i \in [n]$ is the event realization

**Q:** what is the expert’s expected utility?

- Expert believes $i \sim \lambda$
- Expected utility $\mathbb{E}_{i \sim \lambda} S(i; p) = \sum_{i \in [n]} \lambda_i \cdot S(i; p) = S(\lambda; p)$

**Q:** what $S(i; p)$ function can elicit truthful report $\lambda$?

- When expert finds that $\lambda = \arg \max_{p \in \Delta_n} [\sum_{i \in [n]} \lambda_i \cdot S(i; p)]$
- Ideally, $\lambda$ is the unique maximizer
Proper Scoring Rules

**Definition.** The “scoring rule” $S(i; p)$ is [strictly] proper if truthful report $p = \lambda$ [uniquely] maximizes expected utility $S(\lambda; p)$.

- Expert is incentivized to report truthfully iff $S(e; p)$ is proper

**Observations.**

1. $S(i; p) = 0$ is a trivial proper scoring function
2. Proper scores are closed under affine transformation
   - I.e., if $S(i; p)$ is [strictly] proper, so is $\alpha \cdot S(i; p) + \beta$ for any constant $\alpha \neq 0, \beta$

- Thus, typically, strict properness is desired
Examples of Scoring Rules

Example 1 [Log Scoring Rule]

- \( S(i; p) = \log p_i \)
- \( S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i \)

- Negative, but okay – can always add a constant
- Properness requires \( \lambda = \arg \max_{p \in \Delta_n} S(\lambda; p) \)
  
  \[
  S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i
  \]

  \[
  = \sum_{i \in [n]} \lambda_i \left[ \log p_i - \log \lambda_i \right] + \sum_{i \in [n]} \lambda_i \log \lambda_i
  \]
Examples of Scoring Rules

Example 1 [Log Scoring Rule]
- $S(i; p) = \log p_i$
- $S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i$

- Negative, but okay – can always add a constant
- Properness requires $\lambda = \arg \max_{p \in \Delta_n} S(\lambda; p)$
  \[
  S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i \\
  = \sum_{i \in [n]} \lambda_i [\log p_i - \log \lambda_i] + \sum_{i \in [n]} \lambda_i \log \lambda_i \\
  = -\sum_{i \in [n]} \lambda_i \cdot \log \frac{\lambda_i}{p_i} - \text{Entrop}(\lambda)
  \]
Examples of Scoring Rules

Example 1 [Log Scoring Rule]

- $S(i; p) = \log p_i$
- $S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i$

- Negative, but okay – can always add a constant
- Properness requires $\lambda = \arg \max_{p \in \Delta_n} S(\lambda; p)$
  
  $S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i$
  
  $= \sum_{i \in [n]} \lambda_i [\log p_i - \log \lambda_i] + \sum_{i \in [n]} \lambda_i \log \lambda_i$
  
  $= -\sum_{i \in [n]} \lambda_i \cdot \log \frac{\lambda_i}{p_i} - \text{Entrop}(\lambda)$

- KL-divergence $KL(\lambda; p)$ (a.k.a. relative entropy)
  - Measures the distance between two distributions
  - Always non-negative, and equals 0 only when $p = \lambda$
Examples of Scoring Rules

Example 1 [Log Scoring Rule]

- $S(i; p) = \log p_i$
- $S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i$

- Negative, but okay – can always add a constant
- Properness requires $\lambda = \arg \max_{p \in \Delta_n} S(\lambda; p)$

\[
S(\lambda; p) = \sum_{i \in [n]} \lambda_i \cdot \log p_i \\
= \sum_{i \in [n]} \lambda_i [\log p_i - \log \lambda_i] + \sum_{i \in [n]} \lambda_i \log \lambda_i \\
= -\sum_{i \in [n]} \lambda_i \cdot \log \frac{\lambda_i}{p_i} - \text{Entrop}(\lambda)
\]

- $p$ should minimize distance $\text{KL}(\lambda; p)$, which is achieved at $p = \lambda$
- Log scoring rule is strictly proper
Examples of Scoring Rules

Example 2 [Quadratic Scoring Rule]

- $S(i; p) = 2p_i - \sum_{j \in [n]} p_j^2$
- $S(\lambda; p) = \sum_{i \in [n]} \lambda_i [2p_i - \sum_{j \in [n]} p_j^2]$

\[
S(\lambda; p) = \sum_{i \in [n]} \lambda_i [2p_i - \sum_{j \in [n]} p_j^2] \\
= \sum_{i \in [n]} 2\lambda_i p_i - (\sum_{i \in [n]} \lambda_i) \cdot \sum_{j \in [n]} p_j^2 \\
= \sum_{i \in [n]} 2\lambda_i p_i - \sum_{i \in [n]} p_i^2 \\
= -\sum_{i \in [n]} [p_i - \lambda_i]^2 + \sum_{i \in [n]} \lambda_i^2
\]

- Prediction $p$ should minimize $l_2$-distance between $p$ and $\lambda$
- $p_1 = \lambda_1$ is the unique maximizer of $S(\lambda; p)$
- Quadratic scoring rule is also strictly proper
Examples of Scoring Rules

Example 3 [Linear Scoring Rule]

- \( S(i; p) = p_i \)
- \( S(\lambda; p) = \sum_{i \in [n]} \lambda_i p_i \)

- Linear scoring rule turns out to be not proper
What $S(i; p)$ Are Proper?

**Theorem.** The scoring rule $S(i; p)$ is (strictly) proper if and only if there exists a (strictly) convex function $G: \Delta_n \to \mathbb{R}$ such that

$$S(i; p) = G(p) + \nabla G(p)(e_i - p)$$

Recall $G(p)$ is convex if for any $\alpha \in [0,1]$

$$\alpha G(p) + (1 - \alpha)G(q) \geq G(\alpha p + (1 - \alpha)q)$$
What $S(i; p)$ Are Proper?

**Theorem.** The scoring rule $S(i; p)$ is (strictly) proper if and only if there exists a (strictly) convex function $G: \Delta_n \to \mathbb{R}$ such that

$$S(i; p) = G(p) + \nabla G(p)(e_i - p)$$

Proof of “$\Leftarrow$”

$$S(\lambda; p) = \mathbb{E}_{i \sim \lambda}[G(p) + \nabla G(p)(e_i - p)]$$

$$= G(p) + \nabla G(p)(\lambda - p)$$

$$\leq G(\lambda) = S(\lambda; \lambda)$$

By convexity
What $S(i; p)$ Are Proper?

**Theorem.** The scoring rule $S(i; p)$ is (strictly) proper if and only if there exists a (strictly) convex function $G: \Delta_n \to \mathbb{R}$ such that

$$S(i; p) = G(p) + \nabla G(p)(e_i - p)$$

Proof of “⇒”

- $S(\lambda; p) = \sum_{i \in [n]} \lambda_i S(i; p)$ is a linear function of $\lambda$ for any $p$
- By properness, $S(\lambda; \lambda) = \max_{p \in \Delta_n} \sum_{i \in [n]} \lambda_i S(i; p)$, denoted as $G(\lambda)$
  - $G(\lambda)$ is convex in $\lambda$
What $S(i; p)$ Are Proper?

**Theorem.** The scoring rule $S(i; p)$ is (strictly) proper if and only if there exists a (strictly) convex function $G: \Delta_n \to \mathbb{R}$ such that

$$S(i; p) = G(p) + \nabla G(p)(e_i - p)$$

Proof of “⇒”

- $S(\lambda; p) = \sum_{i \in [n]} \lambda_i S(i; p)$ is a linear function of $\lambda$ for any $p$
- By properness, $S(\lambda; \lambda) = \max_{p \in \Delta_n} \sum_{i \in [n]} \lambda_i S(i; p)$, denoted as $G(\lambda)$
  - $G(\lambda)$ is convex in $\lambda$
- The gradient of $G(\lambda)$ is the gradient of $\sum_{i \in [n]} \lambda_i S(i; p)$ for the $p = \lambda$
  - I.e., $\nabla G(\lambda) = S(\cdot; \lambda)$
- Thus,
  $$S(i; p) = S(p; p) + [S(i; p) - S(p; p)]$$
  $$= G(p) + S(\cdot; p) \cdot [e_i - p]$$
  $$= G(p) + \nabla G(p)[e_i - p]$$
Outline

- Scoring Rule and its Characterization
- Connection to Prediction Markets
- Manipulations in Prediction Markets
What If There are Many Experts?

- One idea: elicit their predictions privately/separately

- Drawbacks
  1. May be expensive or wasteful – if experts all agree, we pay many times for the same prediction
  2. Not clear how to aggregate these predictions (average or geometric mean would not work)
  3. In fact, it may require experts’ knowledge to correctly aggregate predictions
Sequential Elicitation

- Ask experts to make predictions in sequence
- The reward for expert $k$’s prediction $p^k$ will be
  \[ S(i; p^k) - S(i; p^{k-1}) \]
  where $p^{k-1}$ is the prediction of expert $k - 1$ 
  - I.e., experts are paid based on how much they improved the prediction

**Theorem.** If $S$ is a proper scoring rule and each expert can only predict once, then each expert maximizes utility by reporting true belief given her own knowledge.

- Proof: since $S(i; p^{k-1})$ not under $k$’s control, she maximizes reward by maximizing $S(i; p^k)$
Sequential Elicitation

- Ask experts to make predictions in sequence
- The reward for expert $k$’s prediction $p^k$ will be
  \[ S(i; p^k) - S(i; p^{k-1}) \]
  where $p^{k-1}$ is the prediction of expert $k-1$
  - I.e., experts are paid based on how much they improved the prediction

**Theorem.** If $S$ is a proper scoring rule and each expert can only predict once, then each expert maximizes utility by reporting true belief given her own knowledge.

Remarks:
- $k$ may see previous reports and then update his prediction
  - Experts will aggregate predictions automatically
Sequential Elicitation

- Ask experts to make predictions in sequence
- The reward for expert $k$’s prediction $p^k$ will be
  \[ S(i; p^k) - S(i; p^{k-1}) \]
  where $p^{k-1}$ is the prediction of expert $k-1$
  - I.e., experts are paid based on how much they improved the prediction

**Theorem.** If $S$ is a proper scoring rule and each expert can only predict once, then each expert maximizes utility by reporting true belief given her own knowledge.

Remarks:
- Not true if an expert can predict for multiple times
  - She may manipulate her initial report to mislead others’ prediction so that she has opportunity to significantly improve her prediction later
Equivalence of PMs and Sequential Elicitation

- It turns out that sequential elicitation is equivalent (in incentives) to the prediction market (PM) for buying and selling contracts.

- Each expert moves the prediction to his own belief:
  - Recall in PMs, expert will buy shares until prices hit his own belief.

- Any strictly proper scoring rule can be used to implement a PM and any PM correspond to some proper scoring rules.
Equivalence of PMs and Sequential Elicitation

**Theorem (informal).** Under mild technical assumptions, efficient prediction markets are in one-to-one correspondence to sequential information elicitation using proper scoring rules.

What does it mean?

- Experts will have exactly the same incentives and receive the same return
- Market maker’s total loss is the same

Next: will *informally* argue using the LMSR and log-scoring rules
Equivalence of LMSR and Log-Scoring Rules

Recall LMSR

- Value function with current sales quantity \( q \): \( V(q) = b \log \sum_{j \in [n]} e^{q_j/b} \)
- To buy \( x \in \mathbb{R}^n \) amount, a buyer pays: \( V(q + x) - V(q) \)
- Price function (they sum up to 1)
  \[ p_i(q) = \frac{e^{q_i/b}}{\sum_{j \in [n]} e^{q_j/b}} = \frac{\partial V(q)}{\partial q_i} \]

**Fact.** The optimal amount an expert purchases is the amount that moves the market price to her belief \( \lambda \).

**Fact.** Worst case market maker loses is \( b \log n \).
Equivalence of LMSR and Log-Scoring Rules

Q1: If current market price is $p^{k-1}$, what is the optimal payoff for an expert with belief $\lambda = p^k$?

- Let $q^{k-1}$ denote the market standing corresponding to price $p^{k-1}$
- That is
  \[
  \frac{e^{q_i^{k-1}/b}}{\sum_{j\in[n]} e^{q_j^{k-1}/b}} = p_i^{k-1}
  \]

Crucial terms:
- Value function $V(q) = b \log \sum_{j\in[n]} e^{q_j/b}$
- Price function $p_i(q) = \frac{e^{q_i/b}}{\sum_{j\in[n]} e^{q_j/b}} = \frac{\partial V(q)}{\partial q_i}$
Equivalence of LMSR and Log-Scoring Rules

Q1: If current market price is $p^{k-1}$, what is the optimal payoff for an expert with belief $\lambda = p^k$?

- Let $q^{k-1}$ denote the market standing corresponding to price $p^{k-1}$
- Optimal purchase for the expert is $x^*$ such that

$$p_i(q^{k-1} + x^*) = \frac{e^{(q_i^{k-1} + x_i^*)/b}}{\sum_{j \in [n]} e^{(q_j^{k-1} + x_j^*)/b}} = p_i^k$$

and pays

$$V(q^{k-1} + x^*) - V(q^{k-1})$$

$$= b \log \sum_{j \in [n]} e^{(q_j^{k-1} + x_j^*)/b} - b \log \sum_{j \in [n]} e^{q_j^{k-1}/b}$$

Crucial terms:
- Value function $V(q) = b \log \sum_{j \in [n]} e^{q_j/b}$
- Price function $p_i(q) = \frac{e^{q_i/b}}{\sum_{j \in [n]} e^{q_j/b}} = \frac{\partial V(q)}{\partial q_i}$
Equivalence of LMSR and Log-Scoring Rules

**Q1:** If current market price is \( p^{k-1} \), what is the optimal payoff for an expert with belief \( \lambda = p^k \)?

- Let \( q^{k-1} \) denote the market standing corresponding to price \( p^{k-1} \).
- Optimal purchase for the expert is \( x^* \) such that

\[
p_i(q^{k-1} + x^*) = \frac{e^{(q_i^{k-1}+x_i^*)/b}}{\sum_{j \in [n]} e^{(q_j^{k-1}+x_j^*)/b}} = p_i^k
\]

and pays

\[
V(q^{k-1} + x^*) - V(q^{k-1}) = b \log \sum_{j \in [n]} e^{(q_j^{k-1}+x_j^*)/b} - b \log \sum_{j \in [n]} e^{q_j^{k-1}/b}
\]
Equivalence of LMSR and Log-Scoring Rules

**Q1:** If current market price is $p^{k-1}$, what is the optimal payoff for an expert with belief $\lambda = p^k$?

- Let $q^{k-1}$ denote the market standing corresponding to price $p^{k-1}$.
- Optimal purchase for the expert is $x^*$ such that

$$p_i(q^{k-1} + x^*) = \frac{e^{(q^{k-1}i + x^*_i)/b}}{\sum_{j \in [n]} e^{(q^{k-1}j + x^*_j)/b}} = p^k_i$$

and pays

$$V(q^{k-1} + x^*) - V(q^{k-1})$$

$$= b \log \sum_{j \in [n]} e^{(q^{k-1}j + x^*_j)/b} - b \log \sum_{j \in [n]} e^{q^{k-1}j/b}$$

$$= b \log \frac{e^{(q^{k-1}i + x^*_i)/b}}{p^k_i} - b \log \frac{e^{q^{k-1}i/b}}{p^k_i - 1}.$$
Equivalence of LMSR and Log-Scoring Rules

**Q1:** If current market price is \( p^{k-1} \), what is the optimal payoff for an expert with belief \( \lambda = p^k \)?

- Let \( q^{k-1} \) denote the market standing corresponding to price \( p^{k-1} \)
- Optimal purchase for the expert is \( x^\ast \) such that

\[
p_i(q^{k-1} + x^\ast) = \frac{e^{(q_i^{k-1} + x^\ast)/b}}{\sum_{j \in [n]} e^{(q_j^{k-1} + x^\ast)/b}} = p_i^k
\]

and pays

\[
V(q^{k-1} + x^\ast) - V(q^{k-1})
\]

\[
= b \log \sum_{j \in [n]} e^{(q_j^{k-1} + x^\ast)/b} - b \log \sum_{j \in [n]} e^{q_j^{k-1}/b}
\]

\[
= b \log \frac{e^{(q_i^{k-1} + x^\ast)/b}}{p_i^k} - b \log \frac{e^{q_i^{k-1}/b}}{p_i^{k-1}}
\]

\[
= x_i^\ast - b (\log p_i^k - \log p_i^{k-1})
\]

Note: this holds for any \( i \).
Equivalence of LMSR and Log-Scoring Rules

Q1: If current market price is $p^{k-1}$, what is the optimal payoff for an expert with belief $\lambda = p^k$?

- Let $q^{k-1}$ denote the market standing corresponding to price $p^{k-1}$
- Repeat our finding: expert pays $x_i^* - b(\log p_i^k - \log p_i^{k-1})$
  - $x^*$ is optimal amount for purchase
- What is the expert utility if outcome $i$ is ultimately realized?

\[ x_i^* - [x_i^* - b(\log p_i^k - \log p_i^{k-1})] \]

from contracts’ return
Equivalence of LMSR and Log-Scoring Rules

**Q1:** If current market price is $p^{k-1}$, what is the optimal payoff for an expert with belief $\lambda = p^k$?

- Let $q^{k-1}$ denote the market standing corresponding to price $p^{k-1}$
- Repeat our finding: expert pays $x_i^* - b(\log p_i^k - \log p_i^{k-1})$
  - $x^*$ is optimal amount for purchase
- What is the expert utility if outcome $i$ is ultimately realized?

\[
x_i^* - [x_i^* - b(\log p_i^k - \log p_i^{k-1})]
\]
\[
= b \cdot [\log p_i^k - \log p_i^{k-1}]
\]
\[
= b \cdot [\text{S}^{\log(i; p^k)} - \text{S}^{\log(i; p^{k-1})}]
\]
\[
= \text{payment in the sequential elicitation}
\]
\[
\text{(constant } b \text{ is a scalar)}
\]
Equivalence of LMSR and Log-Scoring Rules

**Q1:** If current market price is $p_{k-1}$, what is the optimal payoff for an expert with belief $\lambda = p_k$?

- Let $q_{k-1}$ denote the market standing corresponding to price $p_{k-1}$
- Repeat our finding: expert pays $x_i^* - b(\log p_i^k - \log p_i^{k-1})$
  - $x^*$ is optimal amount for purchase
- What is the expert utility if outcome $i$ is ultimately realized?

Expert achieves the same utility in LMSR and log-scoring-rule elicitation for any event realization.
Q2: What is the worst case loss (i.e., maximum possible payment) when using log-scoring rule in sequential info elicitation?

- Total payment – if event $i$ realized – is

$$\sum_{k=1}^{K} [\log p_i^k - \log p_i^{k-1}] = \log p_i^K - \log p_i^0 \leq 0 - \log p_i^0$$

- To avoid cases where some $p_i^0$ is too small (then we need to pay a lot), should choose $p^0 = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ as uniform distribution

- Worst-case loss is thus $\log n$ (same as LMSR, up to constant $b$)
Outline

- Scoring Rule and its Characterization
- Connection to Prediction Markets
- Manipulations in Prediction Markets
Generally, we cannot force experts to participate just once
- E.g., in prediction market, cannot force expert to just purchase once

Manipulations arise when experts can predict multiple times
- This is the case even two experts A, B and only A can predict twice
- The so-called A-B-A game (arguably the most fundamental setting with multiple-round predictions)
An Example of A-B-A Game

- Predict event $E \in \{0,1\}$; Outcome drawn uniformly at random
- Expert Alice observes a signal $A = E$
  - She exactly observes outcome
- Expert Bob also observes the outcome, i.e., signal $B = E$

Q: In A-B-A game, what should Alice predict at stage 1 and 3?

Report her true prediction at stage 1 (which is perfectly correct)
A-B-A Game: Example 2

- Alice observes signal $A \in \{0,1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0,1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

Q: what is the optimal experts’ behaviors in A-B-A game?

Market starts with initial prediction $p^0(\text{YES}) = P^0(\text{NO}) = 1/2$
A-B-A Game: Example 2

- Alice observes signal $A \in \{0,1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0,1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

Q: what is the optimal experts’ behaviors in A-B-A game?

- At stage 1, what is Alice’s probability belief of YES?
  - If Alice’s $A = 1$, then $\Pr(\text{YES}) = 0.49$
  - If Alice’s $A = 0$, then $\Pr(\text{YES}) = 0.51$

- Should Alice report this at stage 1?
  - No, her truthful report tells $B$ exactly the value of her $A$
  - Bob can then make a perfect prediction
A-B-A Game: Example 2

- Alice observes signal $A \in \{0,1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0,1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

Q: what is the optimal experts’ behaviors in A-B-A game?

- What should Alice do at stage 1 then?
  - Say nothing, or equivalently, predict $p^1 = p^0$
A-B-A Game: Example 2

- Alice observes signal $A \in \{0, 1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0, 1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

Q: what is the optimal experts’ behaviors in A-B-A game?

- What should Bob predict at stage 2?
  - Bob learns nothing from stage 1
  - So if $B = 1$, then $\Pr(\text{YES}) = 0.51$; if $B = 0$, then $\Pr(\text{YES}) = 0.49$
  - Should report truthfully based on the above belief – why?

He only has one chance to predict, and his belief is the best given his current knowledge
A-B-A Game: Example 2

- Alice observes signal $A \in \{0, 1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0, 1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

**Q:** what is the optimal experts’ behaviors in A-B-A game?

- What should Bob predict at stage 2?
  - Bob learns nothing from stage 1
  - So If $B = 1$, then $\Pr(\text{YES}) = 0.51$; if $B = 0$, then $\Pr(\text{YES}) = 0.49$
  - Should report truthfully based on the above belief – why?
  - Bob’s truthful report reveals his signal, but gains little utility
A-B-A Game: Example 2

- Alice observes signal $A \in \{0,1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0,1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

Q: what is the optimal experts’ behaviors in A-B-A game?

- What should Alice predict at stage 3?
  - She just learned Bob’s signal $B$
  - So can precisely predict “whether $A + B = 1$” now
  - Alice now moves the prediction from $\Pr(\text{YES}) = 0.51$ or 0.49 to $\Pr(\text{YES}) = 1$ or 0 ➔ receiving a lot of credits
A-B-A Game: Example 2

- Alice observes signal $A \in \{0, 1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0, 1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

Remarks
- Example shows how experts aggregate previous information and update their predictions along the way
- Manipulations arise even if a single expert can predict twice
A-B-A Game: Example 2

- Alice observes signal $A \in \{0,1\}$, and $\Pr(A = 0) = 0.51$
- Bob observes signal $B \in \{0,1\}$, and $\Pr(B = 0) = 0.49$
  - $A, B$ are independent
- They are asked to predict event $E = (\text{whether } A + B = 1)$
  - The answer is YES or NO

Remarks
- This is an issue in prediction markets, since experts can buy and sell whenever they want
- Equilibrium of PMs are still poorly understood, even for the fundamental A-B-A games
  - See a recent paper Computing Equilibria of Prediction Markets via Persuasion for state-of-the-art results
Remarks

Mechanism design for prediction tasks

- ML is one way but not the only way of making predictions
- In some settings, aggregating predictions from experts is more desirable