Three opportunities for IP DT

- **An IP theory of *Forward Induction* in sequential games.**

Methodological Conjecture:

IP affords a novel opportunity to formalize *signals* in sequential games as part of a theory of *Forward Induction*.

We consider *strategic* games, i.e. normal-form strategies for sequential games. That is, each player is required to specify a plan for action at each point in a game tree where that player is called upon to choose an action.

To illustrate the theme highlighted here, it is sufficient to consider 2-person games without *chance* nodes.

At each point in the game tree, either both players choose (simultaneously), or a designated player makes a choice.

Subgame Perfection **requires that a strategy is acceptable** ***if and only if*** **it yields an acceptable strategy in each sub-game within the larger game.**

First, we review an old objection to Subgame Perfection.

Consider the (old) "BoS" 1-stage, simultaneous coordination game.

Row **and** ***Column*** **players are simultaneously choosing, individually, to which concert venue to go in order to share an evening together.**

If they attend different concerts, the evening is ruined, equally, for each.

If they attend the same concert, their preferences depend upon which concert they attend, as <u>Row</u> likes ***Country*** **and <u>Column</u> likes** ***Classical*****.**

Concert A: **Bruch violin concerto, played by Itzhak Perlman**
Concert B: **Dolly Parton performs her favorites.**

Values are in cardinal utilities, with Row's payoffs first in each pair.

	A_C	B_C
A_R	1,2	0,0
B_R	0,0	2,1

Outcomes are quantified by their expected utility for the player in question.

The game admits 3 Nash equilibria:

$< A_R, A_C >$ **yields outcomes (1,2)**

$< B_R, B_C >$ **yields outcomes (2,1)**

$< (1/3)A_R \oplus (2/3)B_R, (2/3)A_C \oplus (1/3)B_C >$ **yields outcomes (2/3, 2/3)**

Each player has a security-maximizer. The mixed strategies

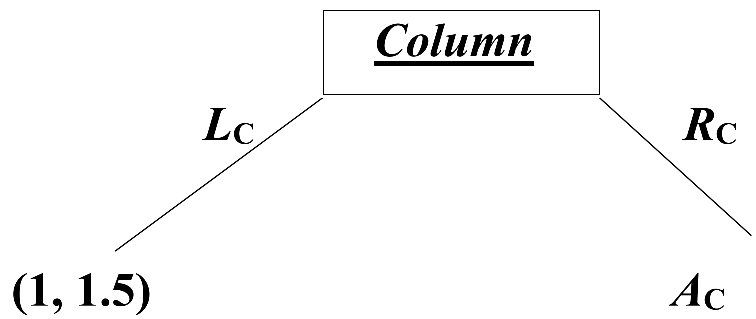
$$< (2/3)A_R \oplus (1/3)B_R, (1/3)A_C \oplus (2/3)B_C >$$

secures each player an outcome of 2/3, independent of the other player's strategy.

But, this pair does not form a Nash equilibrium.

Consider the 2-stage game where <u>Column</u> moves first and chooses between an outcome pair L_C: (1, 1.5) and the Concert game, above:

A: Perlman plays Bruch. B: Parton sings her favorites.



	A_C	B_C
A_R	1,2	0,0
B_R	0,0	2,1

A subgame perfect Nash equilibrium is: $< B_R; <L_C, B_C> >$.

However, this reasoning fails to respect the total evidence available to <u>Row</u> associated with playing the Concert subgame.

Row's Total Evidence at the choice point of the Concert subgame is not merely that they are playing the Concert subgame.

The total evidence is that Column has selected that subgame for them to play!

- **Recall the difference between Row player conditioning on an event**

  ***C*: We play the Concert Game**

and conditioning on an epistemic *signal*.

  ***D*: I know we play the Concert Game because Column showed me that choice.**

If Row believes Column is even mildly rational, Row models Column as follows:

Having chosen option R_C,

Column has an expectation of at least 1.5 – else Column would choose L_C.

So, Column has adopted the strategy $\langle R_C, A_C \rangle$

So, I (= Row) should play $\langle A_R \rangle$.

Then the only subgame perfect Nash equilibrium that respects Total Evidence is:

$$\langle A_R; \langle \mathrm{R}_C, A_C \rangle \rangle.$$

Forward Induction reasoning within IP DT is powerful!

Two players, *Sidney* and *Isaac*, are about to play this 2-stage sequential game

Sidney

Lecture — *Concert*

A: Chomsky lecture

B: Ellsberg lecture

C: Perlman plays Bruch

D: Dolly sings her favorites

Lecture (columns: *Sidney*; rows: *Isaac*)

	A_S	B_S
A_I	**1.5, 4**	**0, 0**
B_I	**0, 0**	**4, 1.5**

Concert (columns: *Sidney*; rows: *Isaac*)

	C_S	D_S
C_I	**1, 2**	**0, 0**
D_I	**0, 0**	**2, 1**

Nash in subgames

< A_I, A_S> yields (1.5, 4)

< B_I, B_S> yields (4, 1.5)

< (3/11)A_I⊕(8/11)B_I, (8/11)A_S⊕(3/11)B_S > yields (12/11, 12/11)

< C_I, C_S> yields (1,2)

< D_I, D_S> yields (2,1)

< (1/3)A_I ⊕ (2/3)B_I, (2/3)A_S ⊕ (1/3)B_S > yields (2/3, 2/3)

Security in subgames

< (8/11)A_I ⊕ (3/11)B_I, (3/11)A_S ⊕ (8/11)B_S > secures (12/11, 12/11) in *Lecture*

< (2/3)C_I ⊕ (1/3)D_I, (1/3)C_S ⊕ (2/3)D_S > secures (2/3, 2/3) in *Concert*

Sidney: *"*Isaac, suppose I choose to go to the *Concert*. What will you do?"

Isaac **mumbles to himself:** ***Well, choosing the Lecture gives Sid an E-admissible option with a security of 12/11. Hmm…***

Isaac: "Then, Sid, I guess we're going to hear Perlman play the Bruch."

Sidney: "Very good. So, let's go to the *Lecture*!

Isaac mumbles to himself*: *Well, rejecting Concert means that Sid expects at least 2 by going to the Lecture.

Isaac: "Then, Sid, I see I'm stuck going to hear Chomsky."

Sidney: "But at least you'll enjoy that more than you would the Bruch!"

***Note*: The application IP to model *Forward Induction* in this sequential game conforms to the *Methodological Conjecture*.**

It uses both the *Expectation* and *Security* components of IP-reasoning in a recursive (two-stage) application of Forward Induction.

First, fixing an expectation of 2 for Sidney in hypothetically choosing the Concert sub-game by rejecting the Lecture sub-game.

Second, using this as a lower bound to signal an expectation of 4 in the Lecture sub-game by rejecting the Concert game.

I do not know of a theory of "precise" expectations that reconstructs this iterated account of forward induction.

Opportunity #2: IP Forecasting

We begin with a general questionnaire.

Each of the following 25 assertions is factual, either true or false.

Next to each assertion, offer your personal probability that it is true,

where probability 1 is "certainly" true and probability 0 is "certainly" false.

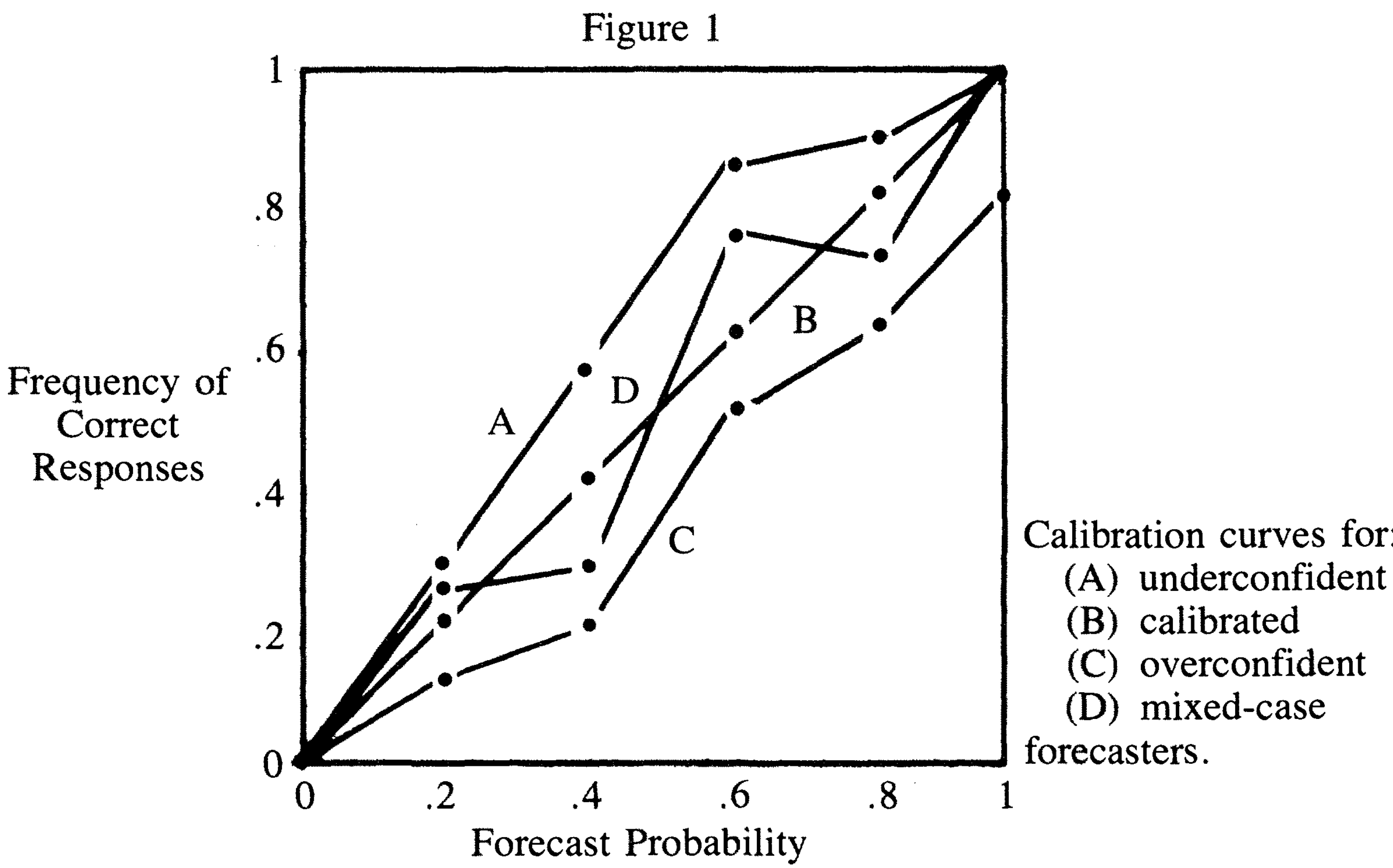
Please use probability numbers in units of 1/20, e.g., .15, .30, .65, etc.

We will be considering how to assess expertise in probabilistic forecasting, and your answers will serve as a sample of such expert forecasts!

1. **The Amazon River is longer than the Congo River.**
2. **The Congo River is longer than the Missouri River.**
3. **The Missouri River is longer than the Niger River.**
4. **The Niger River is longer than the Mississippi River.**
5. **The Mississippi River is longer than the Volga River.**
6. **At its deepest point the Pacific Ocean is deeper than the Atlantic Ocean.**
7. **At its deepest point the Atlantic Ocean is deeper than the Indian Ocean.**

8. **At its deepest point the Indian Ocean is deeper than the Artic Ocean.**
9. **At its deepest point the Artic Ocean is deeper than the Mediterrian Sea.**
10. **In area, the Sahara Desert is larger than the Gobi Desert.**
11. **In area, the Gobi Desert is larger than the Libyan Desert.**
12. **In area, the Libyan Desert is larger than the Kalahari Desert.**
13. **In area, the Kalahari Desert is larger than the Arabian Desert.**
14. **In area, the Arabian Desert is larger than the Painted Desert.**
15. **Blaise Pascal was born before Gottfried Leibnitz was born.**
16. **Gottfried Leibnitz was born before George Berkeley was born.**
17. **George Berkeley was born before David Hume was born.**
18. **David Hume was born before Emmanuel Kant was born.**
19. **Emmanuel Kant was born before Jeremy Bentham was born.**
20. **Jeremy Bentham was born before Georg Hegel was born.**
21. **The length of the Earth's equator is greater than the length of its Meridian.**
22. **The Sun has greater density than does liquid water.**
23. **Syracuse, N.Y. has a greater average annual snowfall than Juneau, Alaska.**
24. **Juneau, Alaska has a greater average annual snowfall than Flagstaff, Arizona.**
25. **Flagstaff, Arizona has a greater average annual snowfall than Buffalo, N.Y.**

Calibration curves for probabilistic forecasting



- **What do you think will be your calibration curve for the survey with which we began?**

- **For which probability forecasts will you be best calibrated? Worst calibrated?**

How will your forecasts be affected by feedback?

I will now reveal the truth values of each prime numbered question:

Q2 is _, Q3 is _, Q5 is _, Q7 is _, Q11 is _, Q13 is _, Q17 is _, Q19 is _, Q23 is _.

- **Reset your probability forecasts for the remaining 16 questions given this news.**
- **Do you think your revised 16 forecasts will have better calibration than the first forecasts you gave for these 16 forecasts? Why??**
- **Now I will reveal the truth values for the remaining assertions. Please check your calibration.**
- **Suppose that, before final calibration scores are awarded, you are asked to include forecasts for the following 15 additional assertions.**

26. George Washington’s famous white horse was dark brown in color.

27. George W. Bush has his picture on the US $1 bill.

28. Las Vegas is the capital of the USA.

29. Hollywood is the capital of the USA.

30. The Pittsburgh Pirates, who lost more games than they won each year between 1993 and 2012, had the best won/loss record in Major League baseball over those 20 years.

31. It will snow at least 1 meter in Bristol today (August 16, 2022).

32. ….

….

40. Senator Ted Cruz is the most popular living American, especially among registered Democrats.

- **IS CALIBRATION RELEVANT TO EXPERTISE IN FORECASTING?**

Regret-like Scoring Rules for Probabilistic Forecasting – best outcome is 0 loss.

In each of the following, multiple forecasts get the sum of the individual scores.

1. "0-1" loss for non-probabilistic forecasts. (T/F exams)

You are penalized 1 utile if a forecast differs from the realized value of the random variable, 0 if the forecast agrees with the realized value.

In what follows, X is a random quantity, i.e. X(ω) is a real number for each ω ∈ Ω. For forecasting events, note that the event E can be identified with its indicator function,

$$\chi_E \quad \begin{cases} = & 1 \text{ if E obtains} \\ = & 0 \text{ if E fails to obtain.} \end{cases}$$

***Problem*:**

- **As an expected utility maximizer, if your personal probability for event E is P(E) what should you offer as your forecast Q(E) of E subject to 0-1 loss?**

- **What should you announce as your forecast for the simple random variable X?**

2. "Mean Deviation" loss.

For random quantity X, you are penalized $|X - F(X)|$ for your forecast $F(X)$.

***Problem*: As an expected utility maximizer, what is your forecast for X subject to mean-deviation loss?**

***Answer*: Forecast the median of YOUR distribution for *X*.**

3. "Squared-error" loss -- Brier score.

For random quantity X, your penalty is $[X - F(X)]^2$ utiles for your forecast F(X).

- ***Class Problem*: *As an expected utility maximizer, what is your best forecast for event A subject to squared error loss?***

Suppose $I_A(\omega)$ is an indicator function for the event A.

$$I_A(\omega) = 1 \text{ if } \omega \in A \text{ and } I_A(\omega) = 0 \text{ if } \omega \notin A.$$

Under squared-error loss, what forecast maximizes your expected utility – that is, minimizes your expected (squared-error) loss?

Incentive compatible elicitation

Question: **When YOU are playing the role of the *Bookie* in de Finetti's betting game, what are the incentives for YOU to announce YOUR *fair prices*?**

***Coherence*$_1$: de Finetti's (1937) the 0-sum *Pricing Game*.**

The players in the *Pricing Game*:

- **The *Bookie* – who, for each random variable X in χ announces a *prevision* (a *fair price*), $P(X)$, for buying/selling units of X.**
- **The *Gambler* – who may make finitely many (non-trivial) contracts with the *Bookie* at the *Bookie*'s announced prices.**

For an individual contract, the *Gambler* fixes a real number γ_X, which determines the contract on X, as follows.

In state ω, the contract has an *outcome* to the *Bookie* (and opposite outcome to the *Gambler*) of $\gamma_X[X(\omega) - P(X)] = O_\omega(X, P(X), \gamma_X)$.

When $\gamma_X > 0$, the *Bookie* buys γ_X-many units of X from the *Gambler*.
When $\gamma_X < 0$, the *Bookie* sells $|\gamma_X|$-many units of X to the *Gambler*.

The *Gambler* may choose finitely many non-zero ($\gamma_X \neq 0$) contracts.

The *Bookie*'s net *outcome* in state ω is the sum of the payoffs from the finitely many non-zero contracts: $\sum_{X\in\chi} O_{\omega}(X, P(X), \gamma_X) = O(\omega)$.

***Coherence*$_1$: The *Bookie*'s previsions $\{P(X): X \in \chi\}$ are *coherent*$_1$ provided that there is no strategy for the *Gambler* that results in a sure (uniform) net loss for the *Bookie*.**

$$\neg\exists(\{\gamma_{X1}, \ldots, \gamma_{Xk}\}, \varepsilon > 0), \forall\omega\in\Omega \quad \sum_{X\in\chi} O_{\omega}(X, P(X), \gamma_X) \leq -\varepsilon.$$

Otherwise, the *Bookie*'s previsions are *incoherent*$_1$.

The net outcome *O* is just another random variable.

The *Bookie*'s *coherent*$_1$ previsions do not allow the *Gambler* contracts where the *Bookie*'s net-payoff is uniformly dominated by *Abstaining*.

	ω_1	ω_2	ω_3	…	ω_n	…
O	$O(\omega_1)$	$O(\omega_2)$	$O(\omega_3)$	…	$O(\omega_n)$	…
Abstain	0	0	0	…	0	…

BUT what fair prices should YOU, the *Bookie*, announce, when you know the *Gambler*?

What are the incentives for elicitation with de Finetti's *Pricing Game*?

- **Do coherent, fair odds reveal a person's degrees of belief?**

Example **5: Betting against an "expert."**

The *Bookie* has to price the indicator *A* for event A, but believes that the *Gambler* already knows which of {A, A^c} obtains.

If the *Bookie* announces a prevision $0 < P(A) < 1$, then the *Bookie* anticipates that the *Gambler* will choose γ_A so that *Gambler* wins and *Bookie* loses: $\gamma_A < 0$ if A obtains, and $\gamma_A > 0$ if A^c obtains.
Then, though the *Bookie* loses for sure, she/he is not *incoherent*$_1$.

If p_A is the *Bookie*'s "*straightforward*" fair-price (her/his credence) the *Bookie* plays *strategically* and announces:

$$P(A) = 1 \text{ if } p_A > .5$$

$$P(A) = 0 \text{ if } p_A < .5$$

$$\text{either } P(A) = 1 \text{ or } P(A) = 0 \text{ if } p_A = .5.$$

Then *Bookie* assigns a subjective probability, $max\{ p_A , (1-p_A) \} \geq .5$ to breaking-even, rather than losing for sure.

- **Bold play is optimal in an unfavorable game!**

A different approach to elicitation based on squared-error loss.

There is only the one player in the *Forecasting Game*, the *Forecaster*.

- **The *Forecaster* – who, for random variable X in χ announces a real-valued *forecast* $F(X)$, subject to a squared-error loss outcome.**

In state ω, the *Forecaster* is penalized $-[X(\omega) - F(X)]^2 = O_\omega(X, F(X))$.

The *Forecaster*'s net score in state ω from forecasting finitely variables $\{F(X_i): i = 1, \ldots, k\}$ is the sum of the k-many individual losses

$$\sum_{i=1}^{k} O_\omega(X, F(X_i)) = \sum_{i=1}^{k} -[X_i(\omega) - F(X_i)]^2 = O(\omega).$$

***Coherence*$_2$: The *Forecaster*'s forecasts $\{F(X): X \in \chi\}$ are *coherent*$_2$ provided that there is no finite set of variables, $\{X_1, \ldots, X_k\}$ and set of rival forecasts $\{F'(X_1), \ldots, F'(X_k)\}$ that yields a uniform smaller net loss for the *Forecaster* in each state.**

$$\neg\exists(\{F'(X_1), \ldots, F'(X_k)\}, \varepsilon > 0), \forall\omega\in\Omega$$

$$\sum_{i=1}^{k} -[X_i(\omega) - F(X_i)]^2 \leq \sum_{i=1}^{k} -[X_i(\omega) - F'(X_i)]^2 - \varepsilon.$$

Otherwise, the *Forecaster*'s forecasts are *incoherent*$_1$.

The *Forecaster*'s *coherent*$_2$ previsions do not allow rival forecasts that uniformly dominate in Brier Score (i.e., squared-error).

	ω_1	ω_2	ω_3	…	ω_n	…
O	$O(\omega_1)$	$O(\omega_2)$	$O(\omega_3)$	…	$O(\omega_n)$	…
O'	$O'(\omega_1)$	$O'(\omega_1)$	$O'(\omega_1)$	…	$O'(\omega_1)$	…

Theorem **(de Finetti, 1974):**

A set of previsions $\{P(X)\}$ is *coherent*$_1$.

if and only if

The same *forecasts* $\{F(X): F(X) = P(X)\}$ are coherent$_2$.

if and only if

There exists a (finitely additive) probability P such that these quantities are the P-Expected values of the corresponding variables

$$E_P[X] = F(X) = P(X).$$

Corollary**: When the variables are 0-1 indicator functions for events, A, then de Finetti's theorem asserts:**

Coherent prices/forecasts must agree with the values of a (finitely additive) probability distribution over these same events. Otherwise, they are incoherent.

Opportunity #2**: We have seen how 1-sided prices serves as a basis for IP theory.**

- **How to use proper scoring rules (e.g Brier score) as a foundation for IP forecasting?**

Opportunity #3:

Three Variations on Dominance

we admit (countably) infinite partitions, $\Omega = \{\omega_1, \omega_2, \ldots\}$

	ω_1	ω_2	…	ω_k	…
Act_1	o_{11}	o_{12}	…	o_{1k}	…
Act_2	o_{21}	o_{22}	…	o_{2k}	…

YOU strictly prefer Act_2 over Act_1 in a pairwise choice between them, if

Uniform Dominance: There exists reward o^*, strictly preferred to status quo. Each o_{2j} is strictly preferred to the composite outcome o_{1j} "+" o^*.

Simple Dominance:
Each o_{2j} is strictly preferred to o_{1j}.

Weak Dominance: Each o_{2j} is weakly preferred to o_{1j}, and for some j is strictly preferred.

- **If Act_2 uniformly dominates Act_1, then it simply dominates.**
- **And if it simply dominates, then it weakly dominates.**

Today's presentation emphasized (de Finetti's) coherence: uniform dominance.

- **What Decision Theories incorporate the other dominance principles?**
- **Can an adequate IP DT be developed also to respect, e.g., weak-dominance?**

***Example 6*: A two-sided coin is flipped. It lands either Heads or Tails.**

The decision maker strictly prefers the indicator for Heads $\{H\}$ over the indicator for Tails $\{T\}$.

But, for each real number $c > 1$, the decision maker strictly prefers cT over H.

There are no (real-valued) fair prices q_H and q_T that satisfy these preferences.

What about non-standard prices, where q_H is infinitesimally larger than q_T, and $q_T + q_T = 1$?

The same idea can be used to create infinitesimal prices for events, $\varepsilon > 0$, whose nearest standard real number is 0.

Use these infinitesimals to represent preferences that satisfy weak-dominance.

***Question*: How to develop IP incorporating this theme?**

***Answer(s)*: Under construction!**