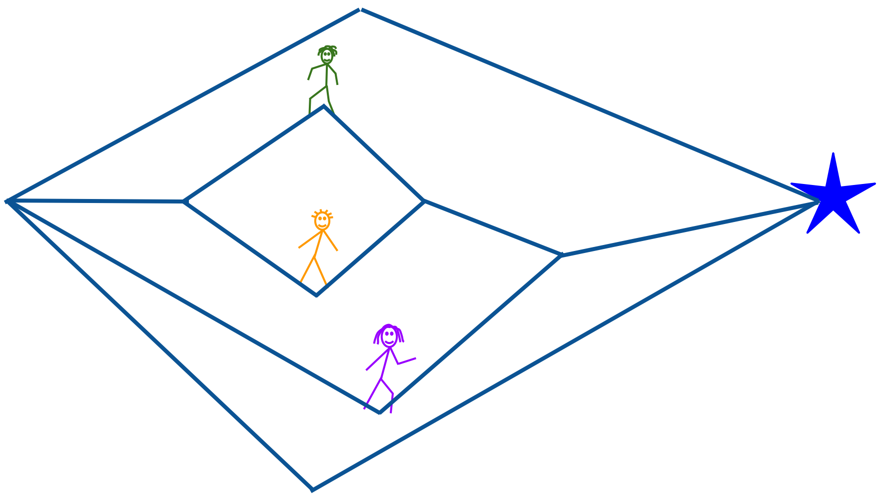
Informational Substitutes
for Prediction and Play

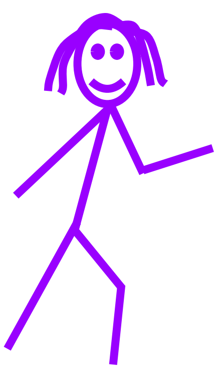


Yiling Chen
Bo Waggoner

Harvard EconCS
March 2016

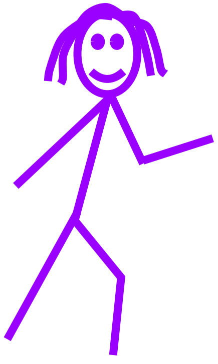
Outline:

1. Develop definitions of informational substitutes
2. A useful tool and some equivalent definitions
3. (How) is information aggregated in prediction markets?
4. How to acquire information under constraints?



Outline:

1. Develop definitions of informational substitutes
2. A useful tool and some equivalent definitions
3. (How) is information aggregated in prediction markets?
4. How to acquire information under constraints?



Intuitive outline for definitions

1. What is the "value" of information?
   → its usefulness in helping make *good decisions*

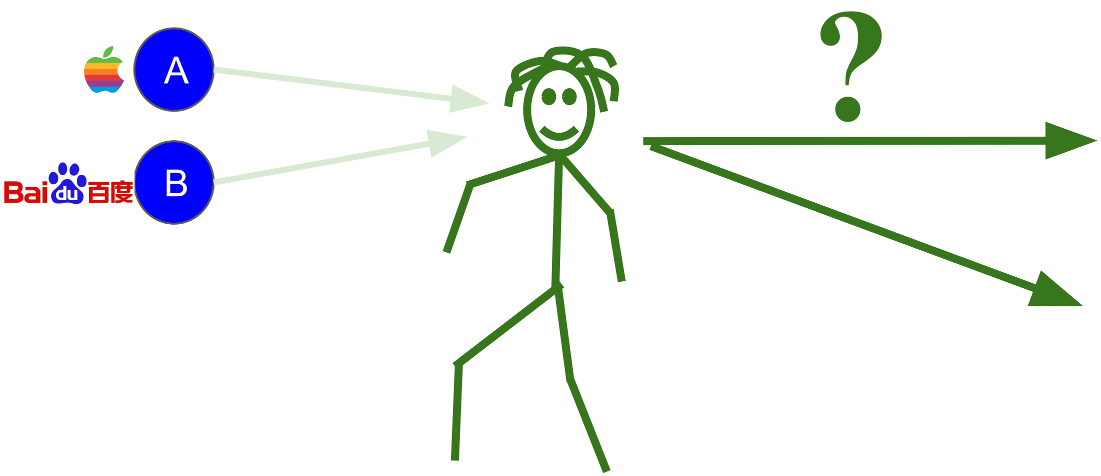
2. When are two signals *substitutes* for a particular decision problem?
   → when the marginal value of B *decreases*
   if we learn (about) A

Quick example

Signals: stock prices of Apple and Baidu

Decision problem 1: Invest in a tech index fund (y/n)?
→ A and B are *substitutes*.

Decision problem 2: Invest in Apple or invest in Baidu?
→ A and B are *complements*.

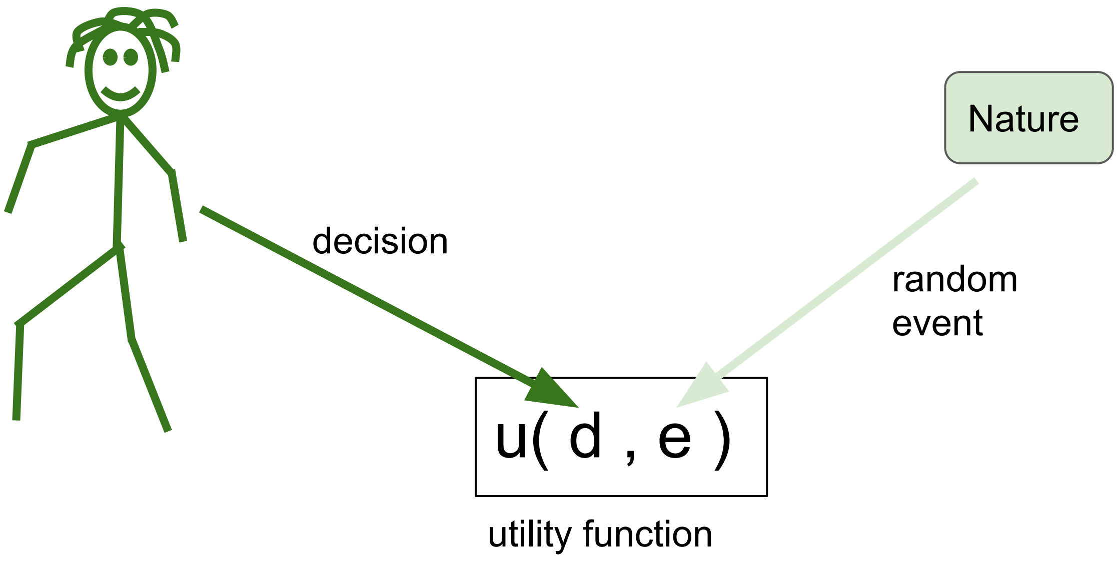


Starting point for definitions

What is the “value” of information?

Starting point for definitions

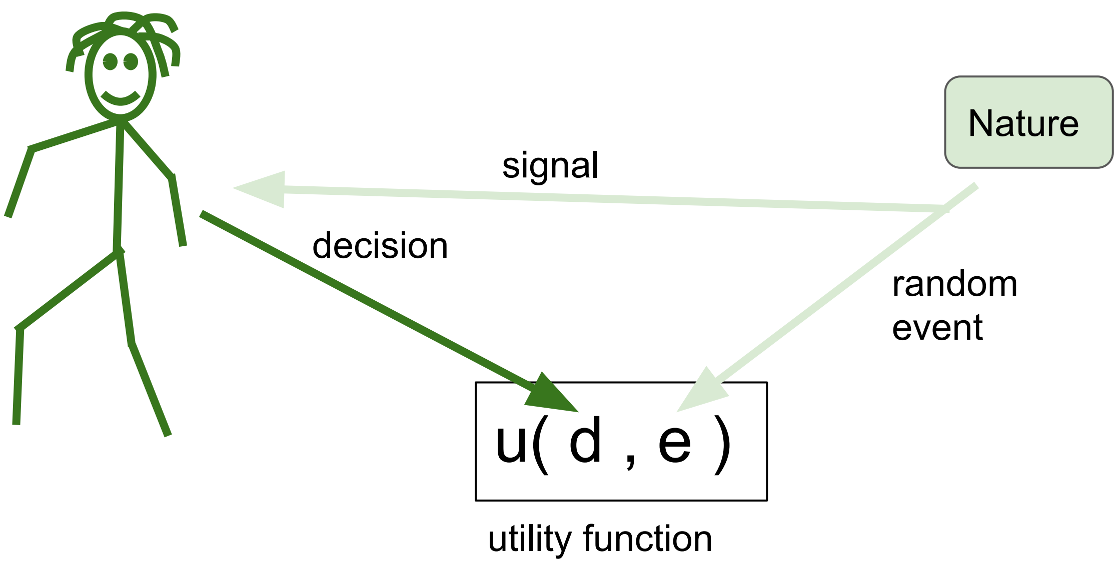
What is the “value” of information? (Context: decision prob)



Starting point for definitions

What is the “value” of information? (Context: decision prob)

The utility for observing that information, then acting.



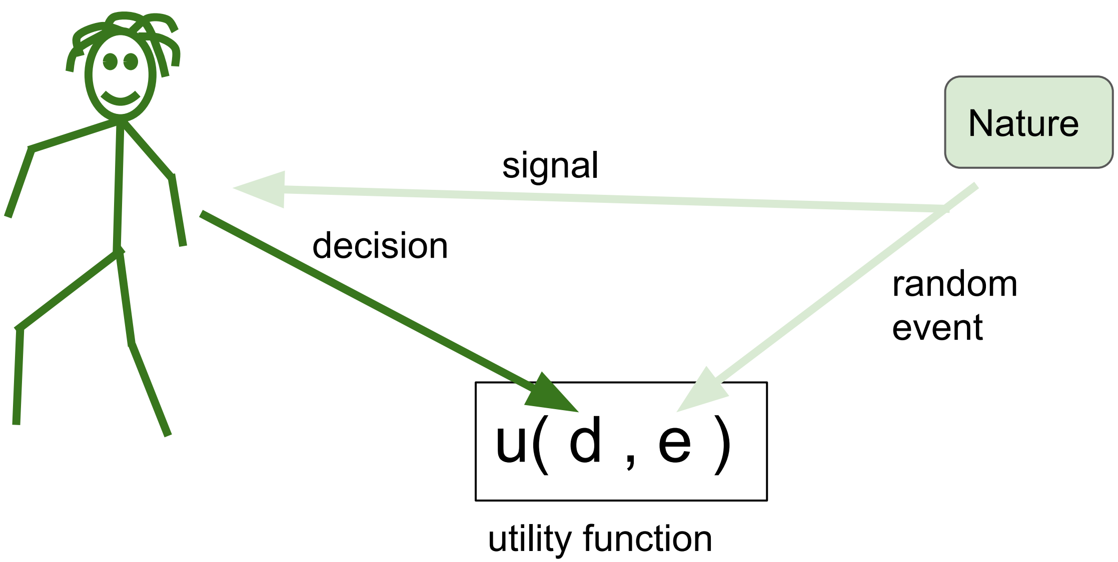
1. Nature draws signal and event
2. Agent observes signal
3. Agent chooses decision

Starting point for definitions

What is the “value” of information? (Context: decision prob)

The utility for observing that information, then acting.

Let V(A) = E_a[util of optimal decision knowing A=a | A=a].



1. Nature draws signal and event
2. Agent observes signal
3. Agent chooses decision

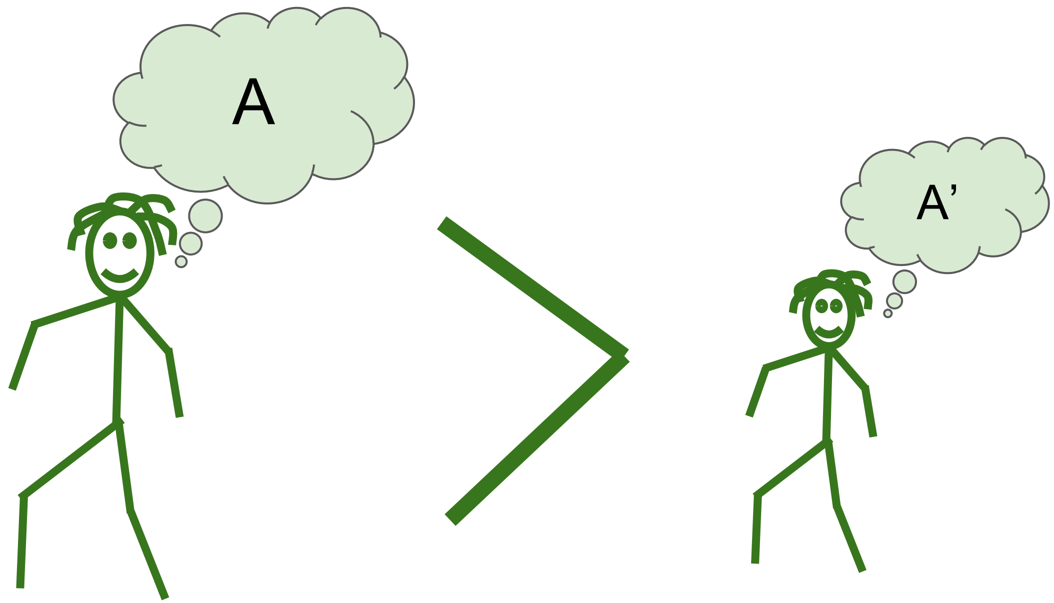
Borgers et al 2013

Capturing “marginal” information

Given A, suppose A’ is independent conditional on A.

Then A’ contains “strictly less” information (is a “garbling”).

→ we use the relation A > A’ (which forms a lattice)



cf. Blackwell criterion

The definitions

A and B are substitutes for a given decision problem if:
for all A' < A,

$$V(A',B) - V(A') \geq V(A,B) - V(A)$$

(and symmetrically for B' < B.)
"marginal value of B is **smaller** the more we know of A"

They are complements if:
For all A' < A,

$$V(B) - V(\varnothing) \geq V(A',B) - V(A')$$

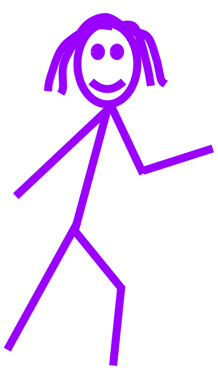
(and symmetrically for B' < B.)
"marginal value of B is **larger** the more we know of A"

Recap / big picture

- V(A) = “expected utility to observe A, then act optimally” in a particular decision problem
- V(B,A) - V(A) = “marginal utility of obtaining B if we will already observe A”
- A and B are ***substitutes*** if, the more one knows of A, the **smaller** the marginal utility of obtaining B
- A and B are ***complements*** if, the more one knows of A, the **larger** the marginal utility of obtaining B

Outline:

1. Develop definitions of informational substitutes
2. A useful tool and some equivalent definitions
3. (How) is information aggregated in prediction markets?
4. How to acquire information under constraints?

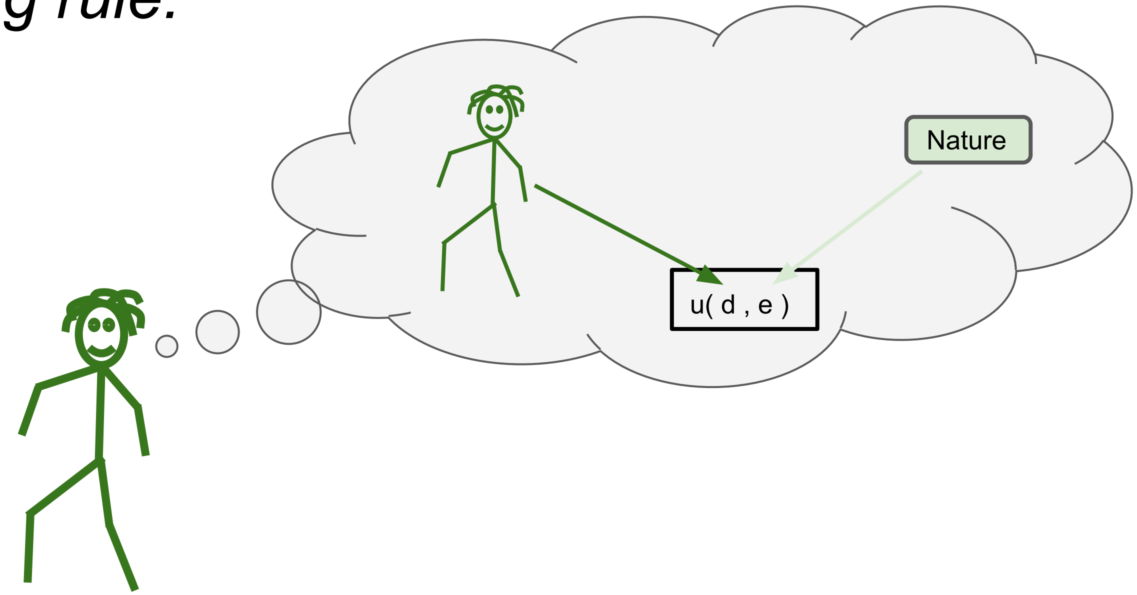


Key tool: Reduce decisionmaking to prediction

Lemma (“revelation principle”):

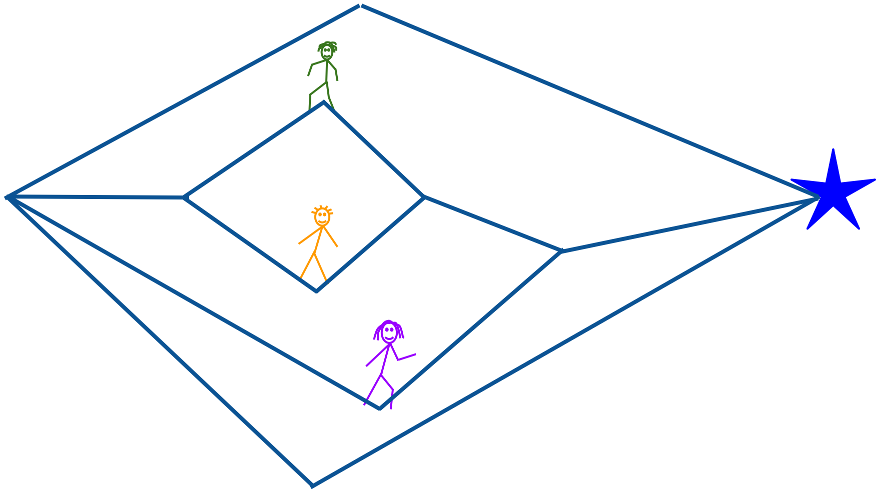
For any decision problem, there is a payoff-equivalent prediction problem.

In it, the agent is asked to predict E and is paid by a proper scoring rule.



Characterization 1: submodularity

1. Signals are substitutes iff V is a submodular function on the signal lattice.
(complements ⇔ supermodular)



Characterization 2: entropy

2. Each decision problem corresponds to a generalized entropy function such that:

A and B are substitutes iff, the more “bits” of information are known about A, the fewer “bits” are revealed by B.

(complements ⇔ more bits of A, more bits of B)

Characterization 3: distance

2. Each decision problem corresponds to a generalized divergence (“distance”) function.

Consider the distance our belief moves when learning B (i.e. by Bayesian updating on B).

Characterization 3: distance

2. Each decision problem corresponds to a generalized divergence ("distance") function.

Consider the distance our belief moves when learning B (i.e. by Bayesian updating on B).

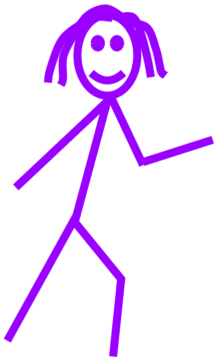
A and B are substitutes iff, the more is known about A, the smaller the distance our beliefs move when updating on B.

(complements ⇔ more info about A, *larger* distance)

Note: log scoring rule = Shannon entropy, KL-divergence

Outline:

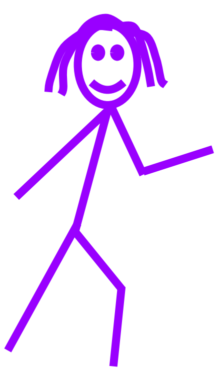
1. Develop definitions of informational substitutes
2. A useful tool and some equivalent definitions
3. (How) is information aggregated in prediction markets?
4. How to acquire information under constraints?

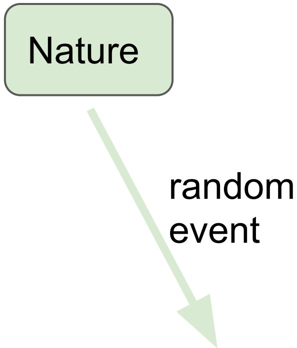


Prediction Markets

Prediction market: toy model of financial markets.

There are “securities” tied to future events (e.g. elections).
When the event occurs, shares of the security pay off.



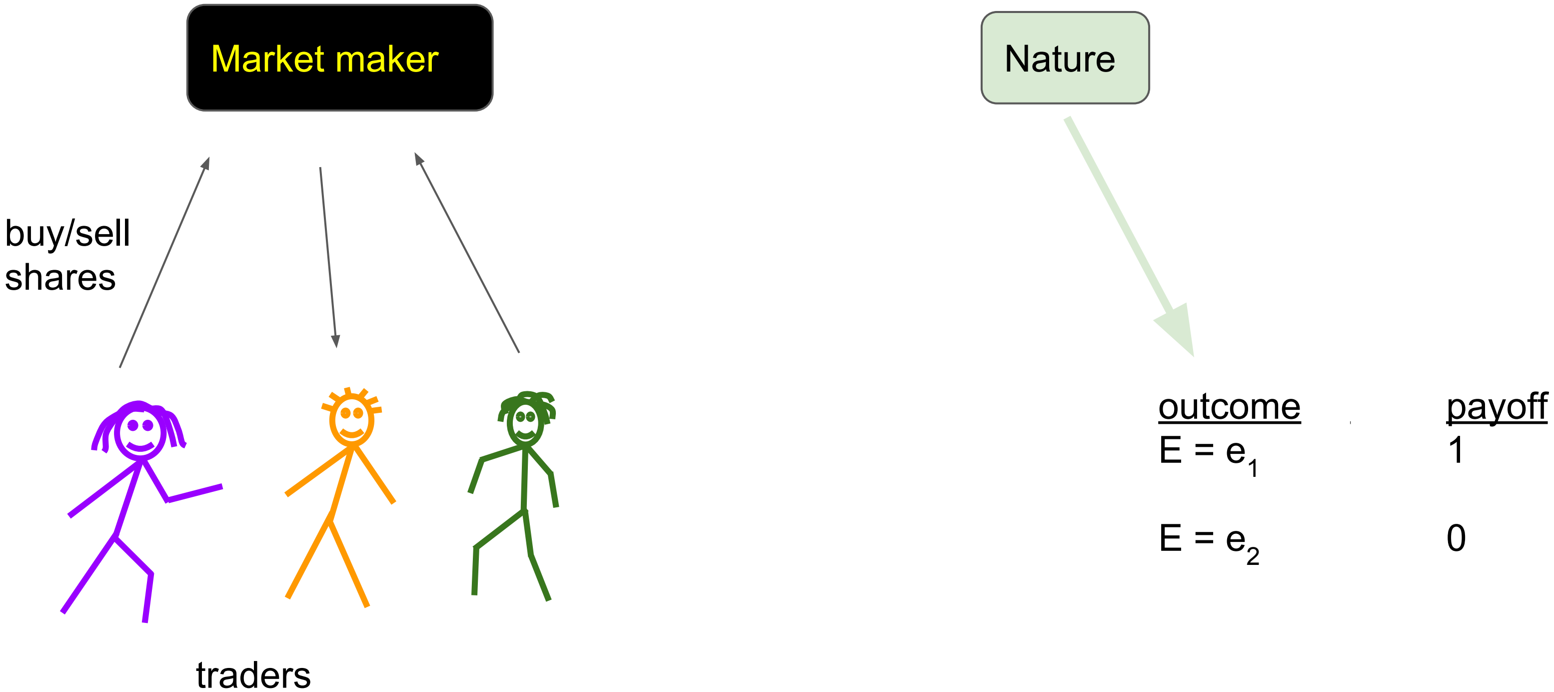


outcome	payoff
$E = e_1$	1
$E = e_2$	0

Prediction Markets

Prediction market: toy model of financial markets.

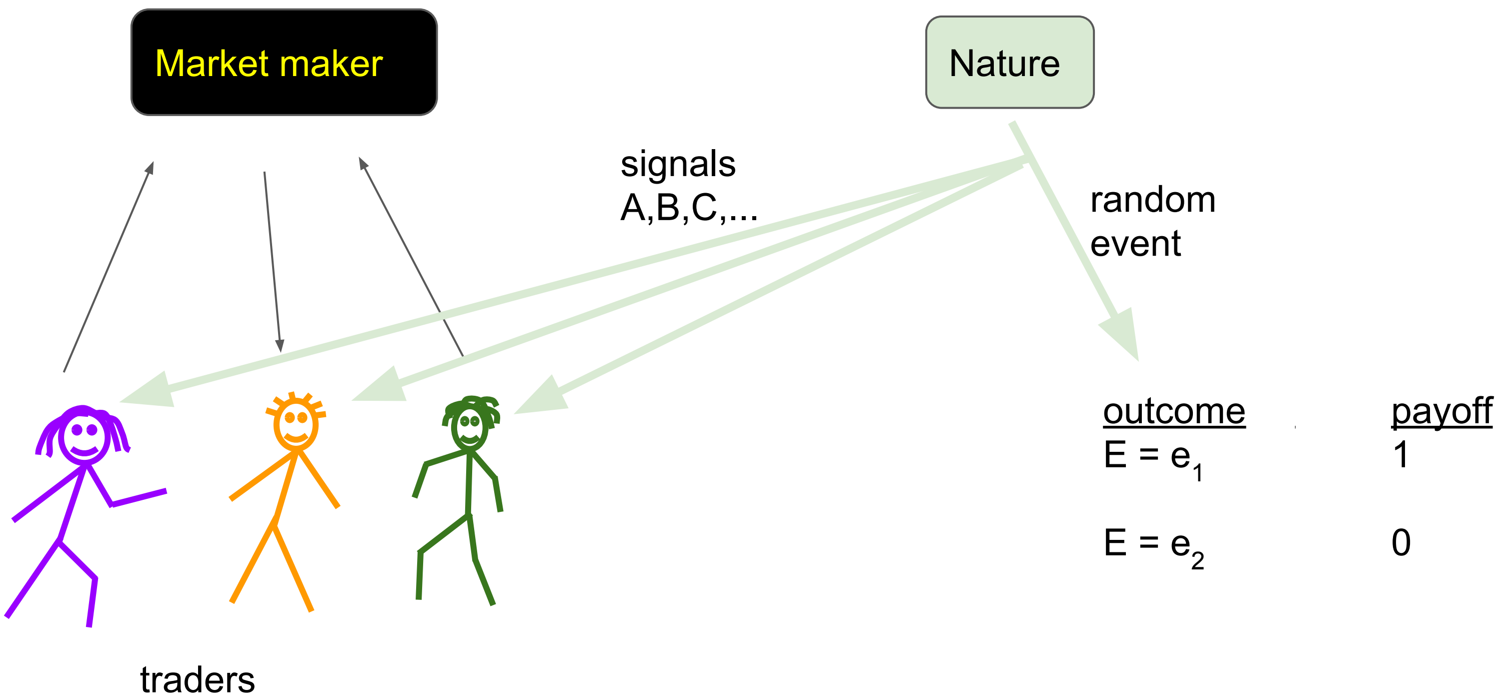
There are “securities” tied to future events (e.g. elections). When the event occurs, shares of the security pay off.



Prediction Markets

Prediction market: toy model of financial markets.

There are “securities” tied to future events (e.g. elections).
When the event occurs, shares of the security pay off.



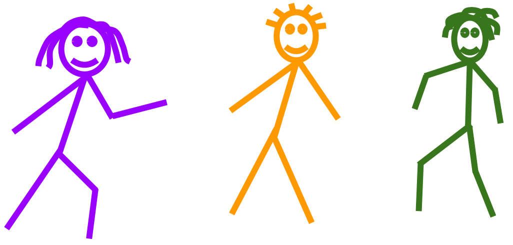
Efficient Market Hypothesis

Is information about events aggregated in markets?

Fama (1970), Kyle (1985), ….

Ostrovsky (2013): Information is always aggregated in markets.

OK, but how?

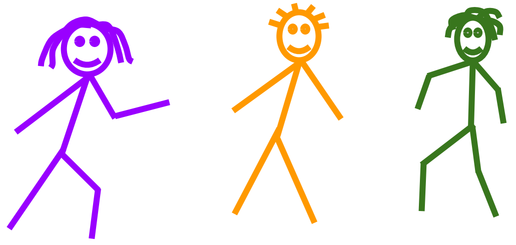


Known results in prediction markets

For the log scoring rule:

- conditionally indep signals ⇒ immediately aggregated.
  (Chen, Dimitrov, Sami, Reeves, Pennock, Hanson, Fortnow, and Gonen 2010)

- unconditionally indep signals ⇒ not aggregated until the last possible moment.
  (Gao, Zhang, Chen 2013)

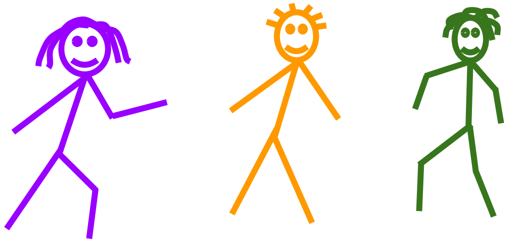


Our results

For any scoring rule and any information structure:

1. Information is **immediately aggregated** if and only if traders' signals are **substitutes**.

2. Information is **not aggregated** until the last possible moment if and only if traders' signals are **complements**.

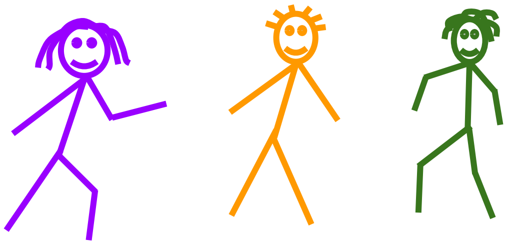


Ideas

Main ideas are very intuitive.

→ Key point: In equilibrium, nobody is deceived!

(they are only under-informed. You cannot bluff in equilibrium.)



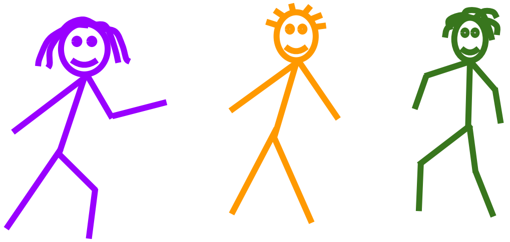
Ideas

Main ideas are very intuitive.

→ Key point: In equilibrium, nobody is deceived!
(they are only under-informed. You cannot bluff in equilibrium.)

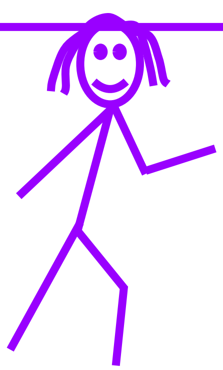
→ Hence, the problem is all about *how much information to reveal* and *when to reveal it*.

→ Markets reward you (essentially) in proportion to the amount of information you reveal at a given time.
(Recall entropy characterization....)



Outline:

1. Develop definitions of informational substitutes
2. A useful tool and some equivalent definitions
3. (How) is information aggregated in prediction markets?
4. How to acquire information under constraints?



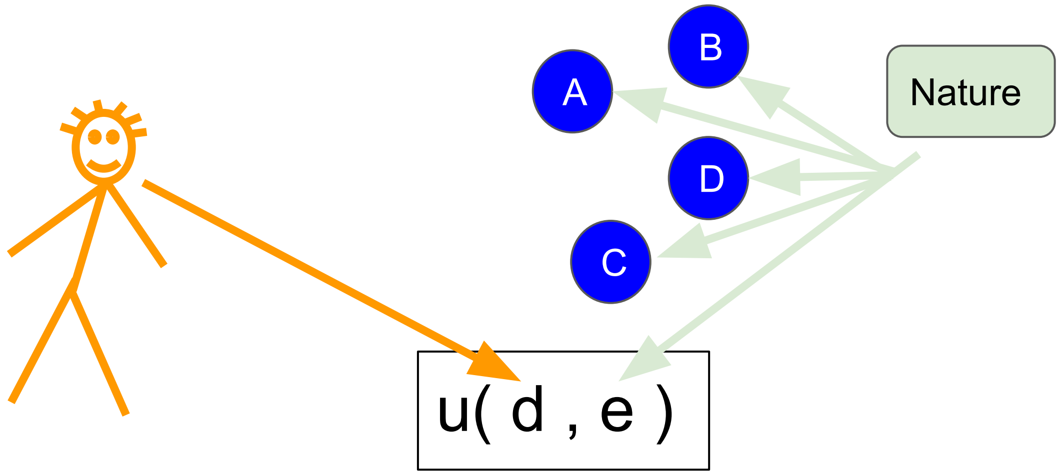
(Approximately) optimal information acquisition

Input:

- a decision problem u(d,e)
- description of signals A, B, … with prices π_A, π_B, …
- Budget B

Output:

- A set of signals to purchase to maximize expected utility

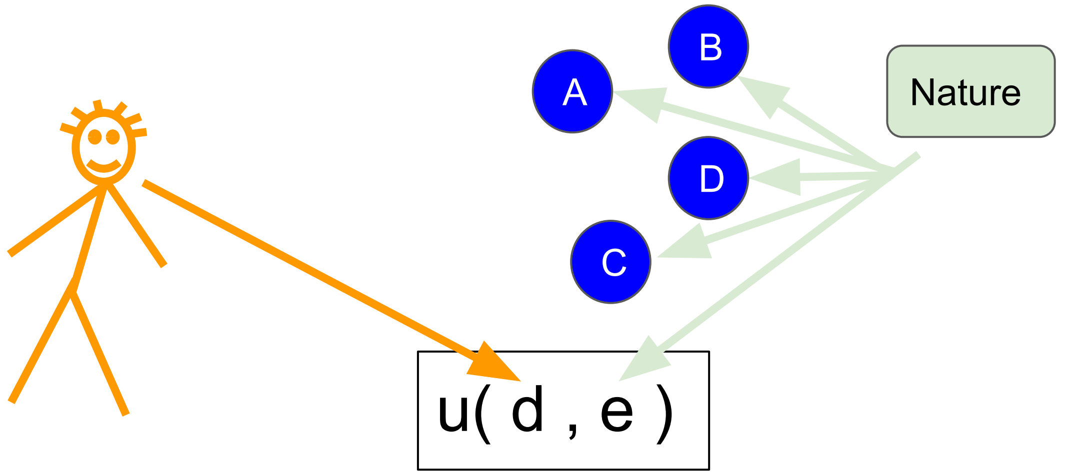


(Approximately) optimal information acquisition

Results:

If signals are substitutes, there exists a 1-1/e approximation algorithm (via reduction to submodular maximization).

In the general case, the problem is as hard as general set function maximization (via a reverse reduction).



(Approximately) optimal information acquisition

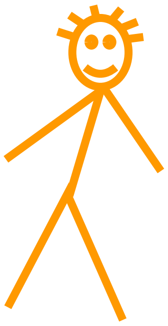
Ideas:

(1) if signals are substitutes, we can implement a submodular value oracle.

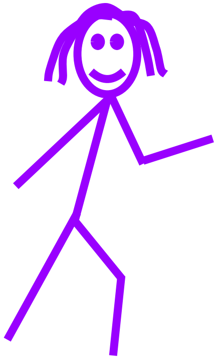
(2) given a general set function, we can construct a matching information structure and decision problem.

PS. this works for all kinds of constraints, e.g. matroid constraints etc.

PPS. issues of representation size come up.



Recap and Conclusion

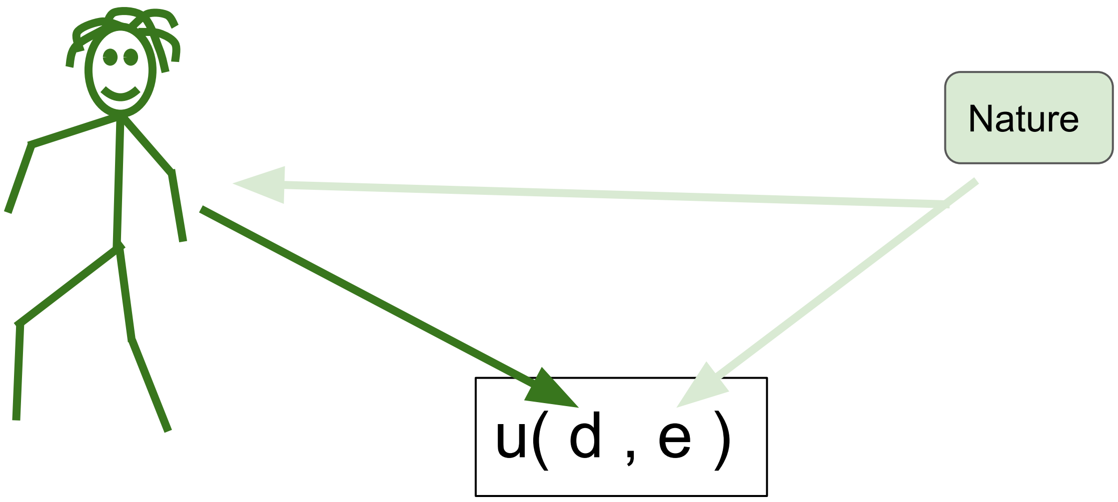


Conclusion

We:

1. developed definitions of informational substitutes and informational complements.

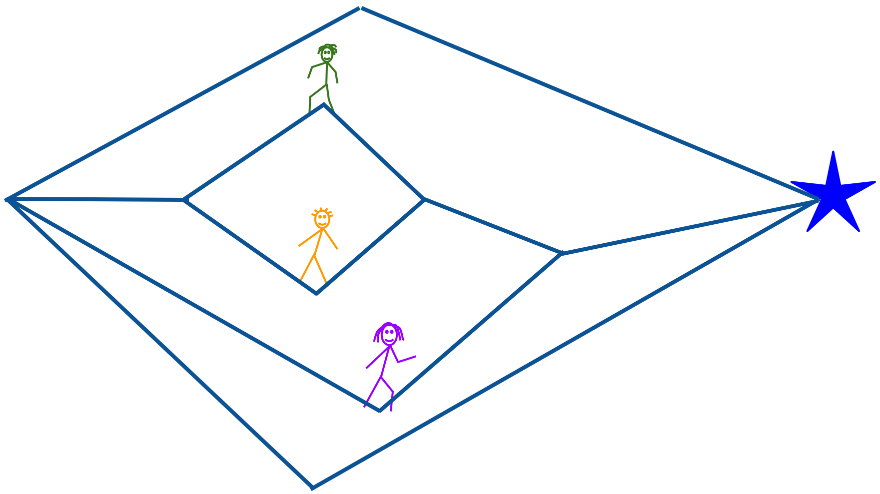
   "substitutes = diminishing marginal value of information"



Conclusion

We:

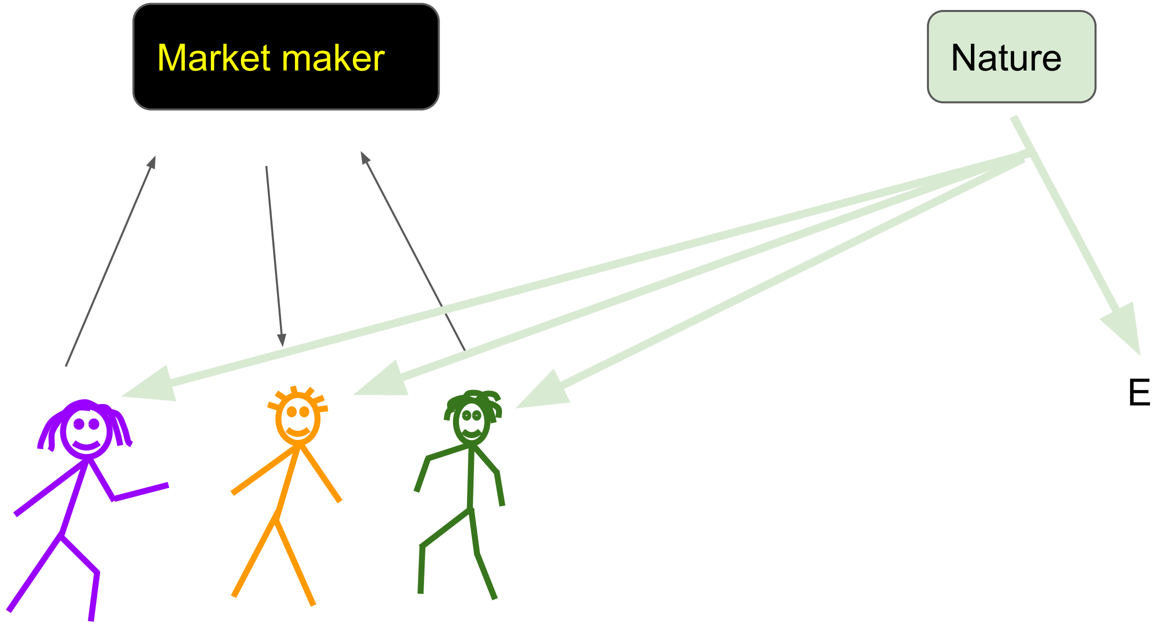
2. saw some equivalent definitions (submodularity, entropy, distance).



Conclusion

We:

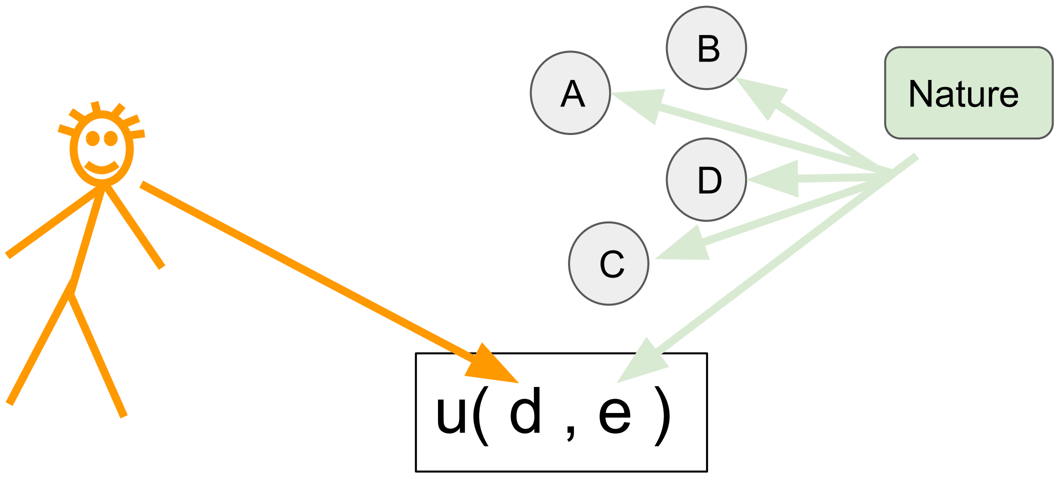
3. saw that substitutes (complements) characterize best-case (worst-case) information aggregation in prediction markets



Conclusion

We:

4. saw that substitutes imply efficient algorithms for information acquisition problems
(which are hard in general)



Thanks!