YAHOO! Research

Combinatorial Prediction Markets

David Pennock

Joint with:

Yiling Chen, Lance Fortnow, Sharad Goel, Mingyu Guo, Joe Kilian, Nicolas Lambert, Eddie Nikolova, Mike Wellman, Jenn Wortman

A Prediction Market

- **Take a random variable, e.g.**

Bin Laden captured by Sept 2009?
(Y/N)

- **Turn it into a financial instrument payoff = realized value of variable**

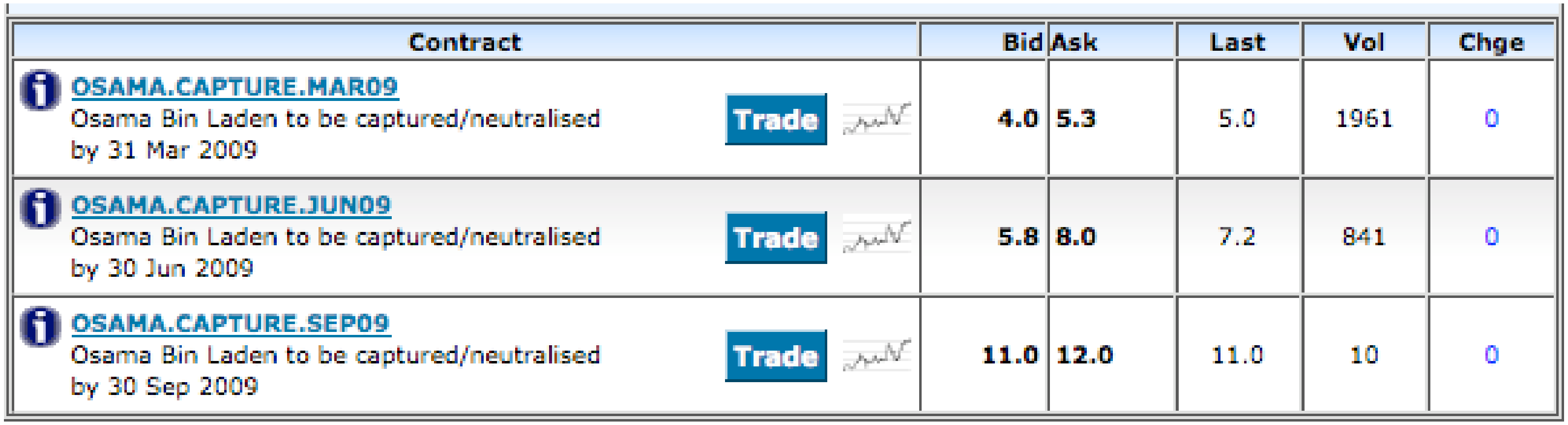
I am entitled to:

$1 if Bin Laden caught

$0 if Bin Laden not caught

intrade™
The Prediction Market

http://intrade.com



Contract		Bid	Ask	Last	Vol	Chge
OSAMA.CAPTURE.MAR09 Osama Bin Laden to be captured/neutralised by 31 Mar 2009	Trade	**4.0**	**5.3**	5.0	1961	0
OSAMA.CAPTURE.JUN09 Osama Bin Laden to be captured/neutralised by 30 Jun 2009	Trade	**5.8**	**8.0**	7.2	841	0
OSAMA.CAPTURE.SEP09 Osama Bin Laden to be captured/neutralised by 30 Sep 2009	Trade	**11.0**	**12.0**	11.0	10	0

Jan 08 - 2:14PM GMT

YAHOO! Research

Prediction Markets

With Money

	? Contract	B Qty	Bid	Ask	A Qty	Last	Vol	Chge
Trade	US.RECESSION.08	1	72.2	73.9	2	74.0	34.9k	+3.0
Trade	BIRDFLU.USA.JUN08	100	6.0	14.0	5	10.0	1323	0
Trade	BIRDFLU.USA.SEP08	10	6.5	16.0	5	11.2	430	0

	? Contract	B Qty	Bid	Ask	A Qty	Last	Vol	Chge
Trade	OSAMA.CAPTURE.MAR08	5	1.9	3.3	1	2.6	4888	0
Trade	OSAMA.CAPTURE.JUN08	4	5.1	5.7	25	5.5	2019	0
Trade	OSAMA.CAPTURE.SEP08	5	8.3	8.8	4	9.1	822	0

	? Contract	B Qty	Bid	Ask	A Qty	Last	Vol	Chge
Trade	2008DEM.NOM.OBAMA	22	71.8	72.0	55	72.0	403.0k	-1.3
Trade	2008DEM.NOM.CLINTON	50	28.5	28.9	4	28.9	549.1k	+1.1

	? Contract	B Qty	Bid	Ask	A Qty	Last	Vol	Chge
Trade	ALABAMA.DEM	0	-	10.0	20	10.0	56	0
Trade	ALABAMA.REP	20	90.0	95.0	5	90.0	22	0
Trade	ALABAMA.FIELD	5	0.1	5.0	20	0.1	0	0
Trade	ALASKA.DEM	20	5.0	10.0	17	7.5	23	0
Trade	ALASKA.REP	20	85.0	95.0	20	92.5	45	0
Trade	ALASKA.FIELD	5	0.1	5.0	20	0.1	0	0

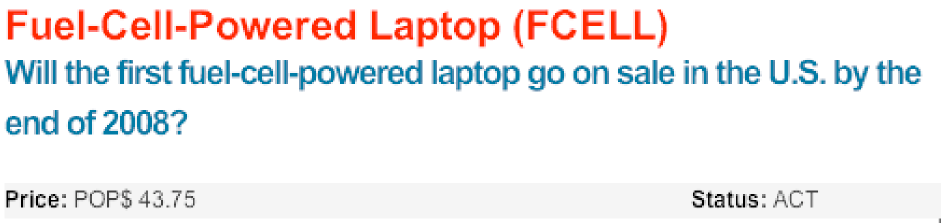
intrade™
The Prediction Market

Without

Androids Beat Humans in Soccer (BOTS)

Will a team of androids beat the human World Cup champs at a game of soccer by 2050?

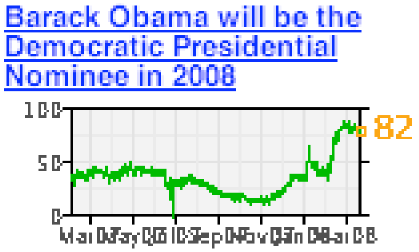
Price: POP$ 47.75 **Status:** ACT



Fuel-Cell-Powered Laptop (FCELL)

Will the first fuel-cell-powered laptop go on sale in the U.S. by the end of 2008?

Price: POP$ 43.75 **Status:** ACT



Barack Obama will be the Democratic Presidential Nominee in 2008

PPX THE POPSCI PREDICTIONS EXCHANGE

newsfutures

Continuous Double Auction

Uber-Hammer of the Financial World

- Buy offers

ACME stock

- Sell offers

$300

$170

$160

$150

$120

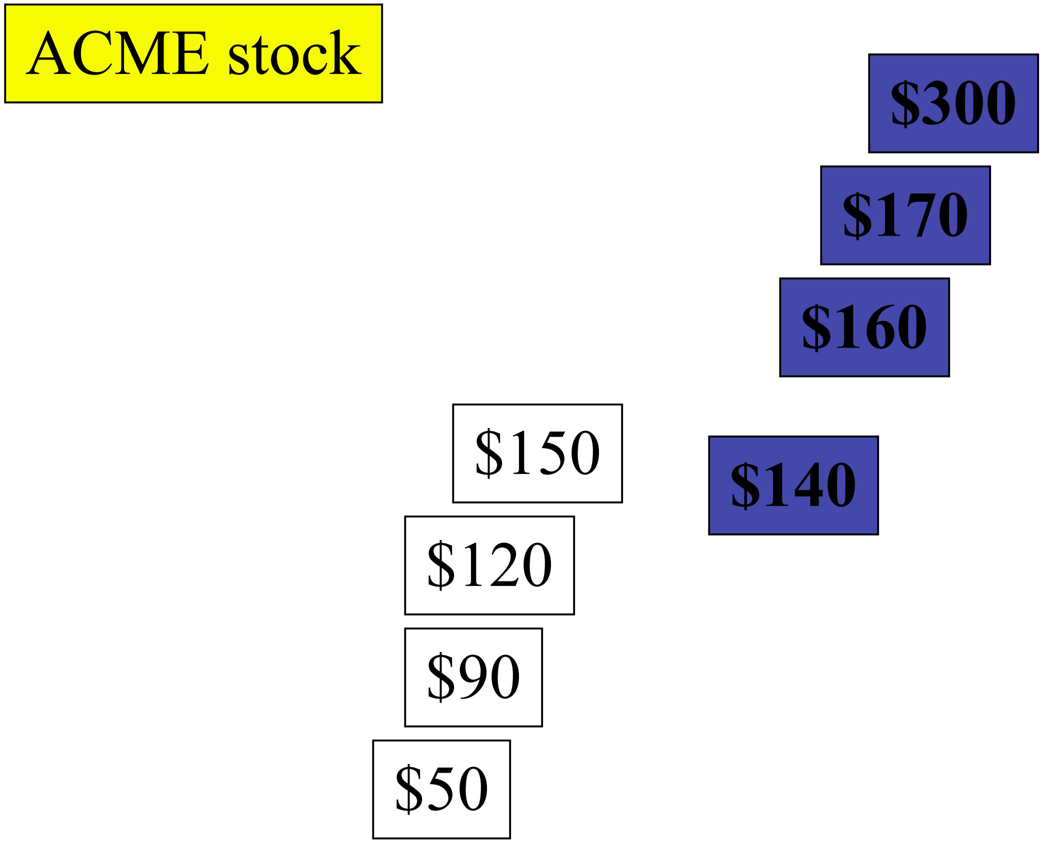
$90

$50

Continuous Double Auction

Uber-Hammer of the Financial World

- Buy offers
- Sell offers



ACME stock

$300

$170

$160

$140

$150

$120

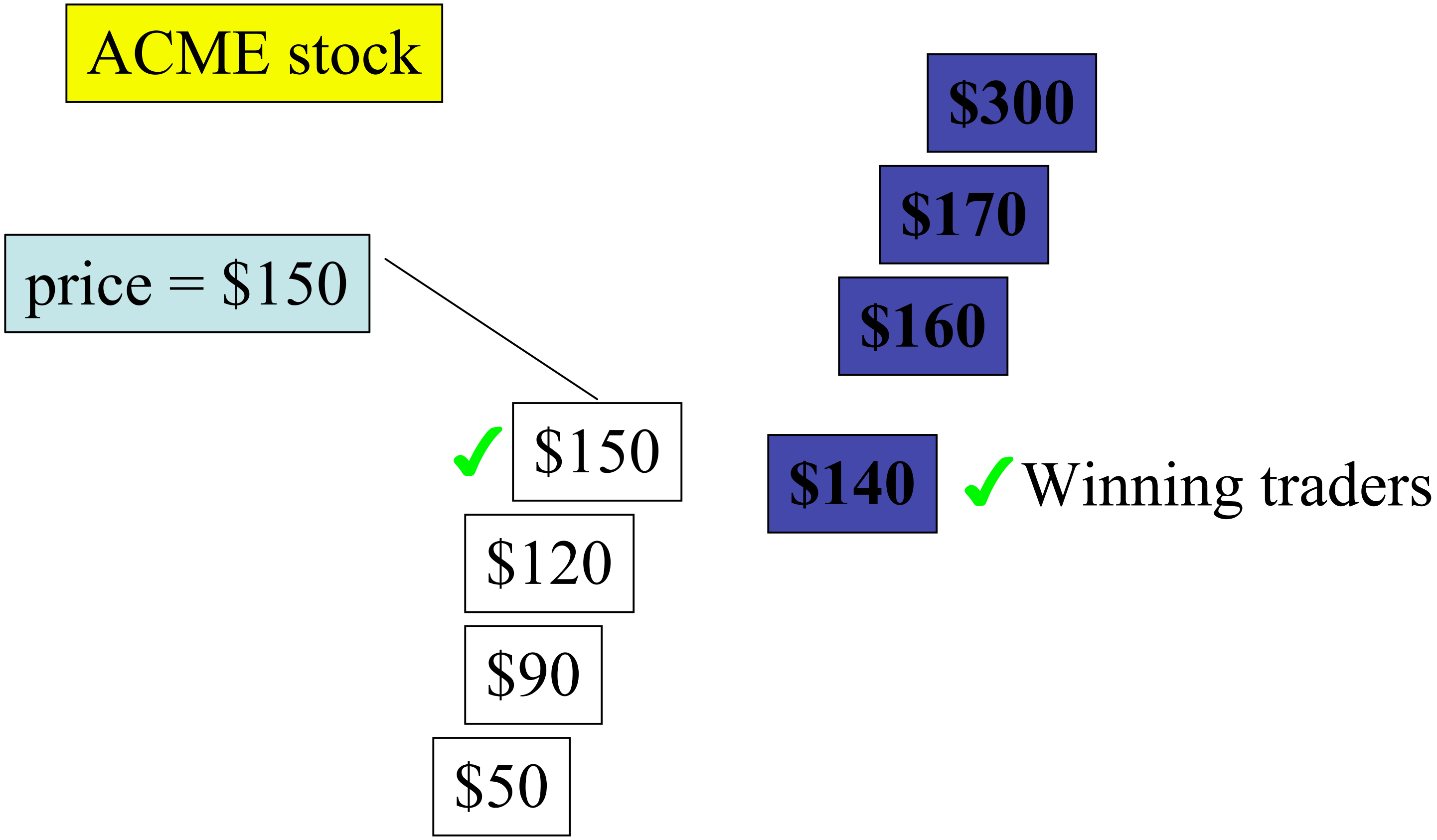
$90

$50

Continuous Double Auction

Uber-Hammer of the Financial World

- Buy offers
- Sell offers



ACME stock

price = $150

✓ $150

$120

$90

$50

$300

$170

$160

$140 ✓ Winning traders

Continuous Double Auction

Uber-Hammer of the Financial World

- Buy offers

ACME stock

$120

$90

$50

- Sell offers

$300

$170

$160

Continuous Double Auction

Uber-Hammer of the Financial World

- Used everywhere
  - – Stocks, options, futures, derivatives
  - – Gambling: BetFair, InTrade
- Related bets? Just use two CDAs
  - – Max[YHOO-10], Max[YHOO-20]
  - – Horse wins, Horse finishes 1st or 2nd
  - – "Power set" instruments: Mutual funds, ETFs, butterfly spreads, "Western Conference wins"
  - – Treats everything like apples and oranges, even hamburgers and cheeseburgers

Continuous Double Auction

Uber-Hammer of the Financial World

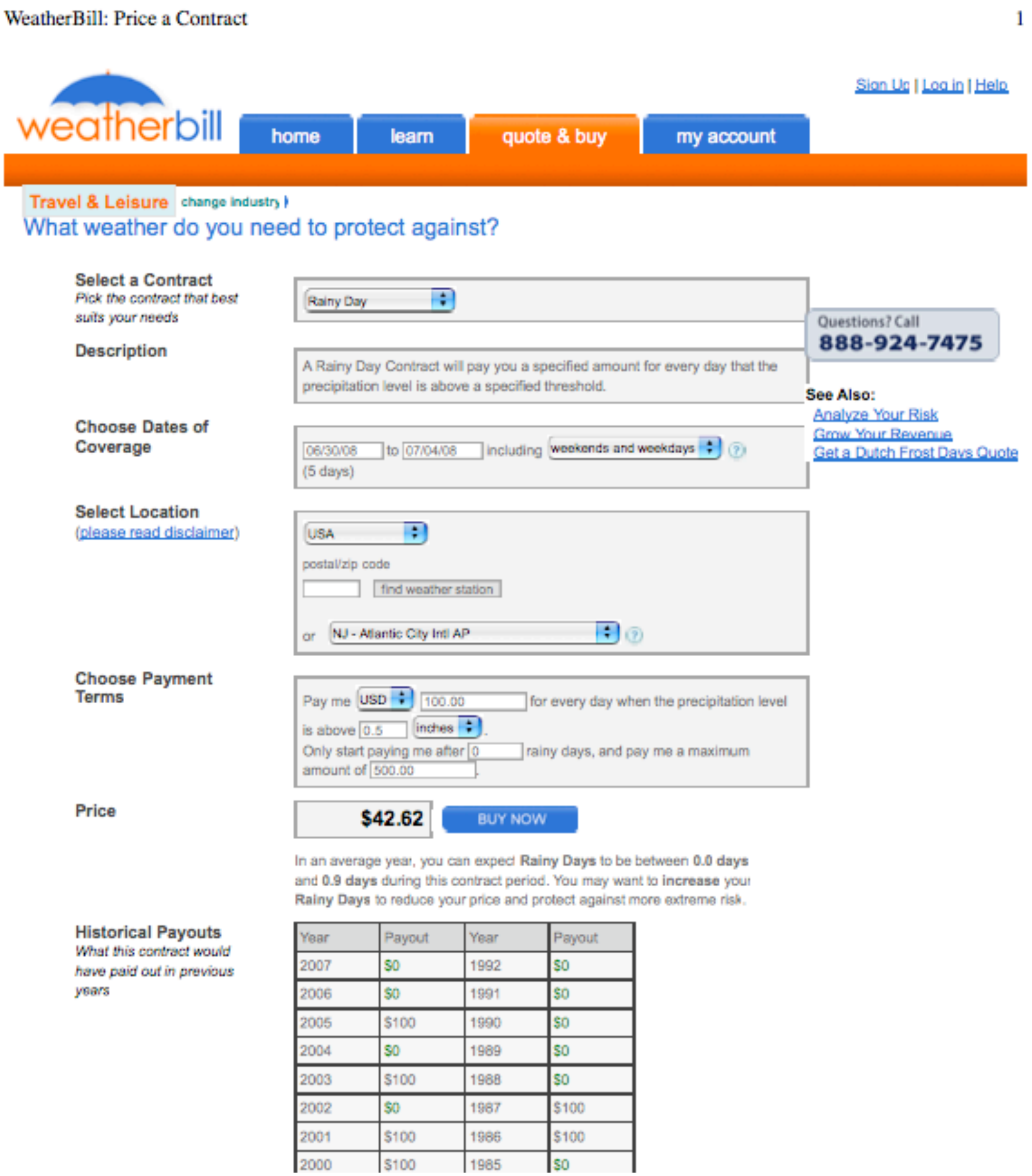
- CDA was invented when auctioneers were people
- Had to be dead simple
- Today, auctioneers are computers...

- ...Yet CDA remains the standard

Like Ordering a Wendy’s Hamburger

- Informal definition: A *combinatorial market* is one where users construct their own bets by mixing and matching options in myriad ways
- Wendy's bags circa March 2008: *"We figured out that there are 256 ways to personalize a Wendy's hamburger. Luckily someone was paying attention in math class."*

Example I: WeatherBill



WeatherBill: Price a Contract 1

Travel & Leisure change industry ▸

What weather do you need to protect against?

Select a Contract
Pick the contract that best suits your needs

Rainy Day

Description

A Rainy Day Contract will pay you a specified amount for every day that the precipitation level is above a specified threshold.

Choose Dates of Coverage

06/30/08 to 07/04/08 including weekends and weekdays
(5 days)

Select Location
(please read disclaimer)

USA
postal/zip code
find weather station
or NJ - Atlantic City Intl AP

Choose Payment Terms

Pay me USD 100.00 for every day when the precipitation level is above 0.5 inches.
Only start paying me after 0 rainy days, and pay me a maximum amount of 500.00.

Price

$42.62 BUY NOW

In an average year, you can expect **Rainy Days** to be between **0.0 days** and **0.9 days** during this contract period. You may want to **increase** your **Rainy Days** to reduce your price and protect against more extreme risk.

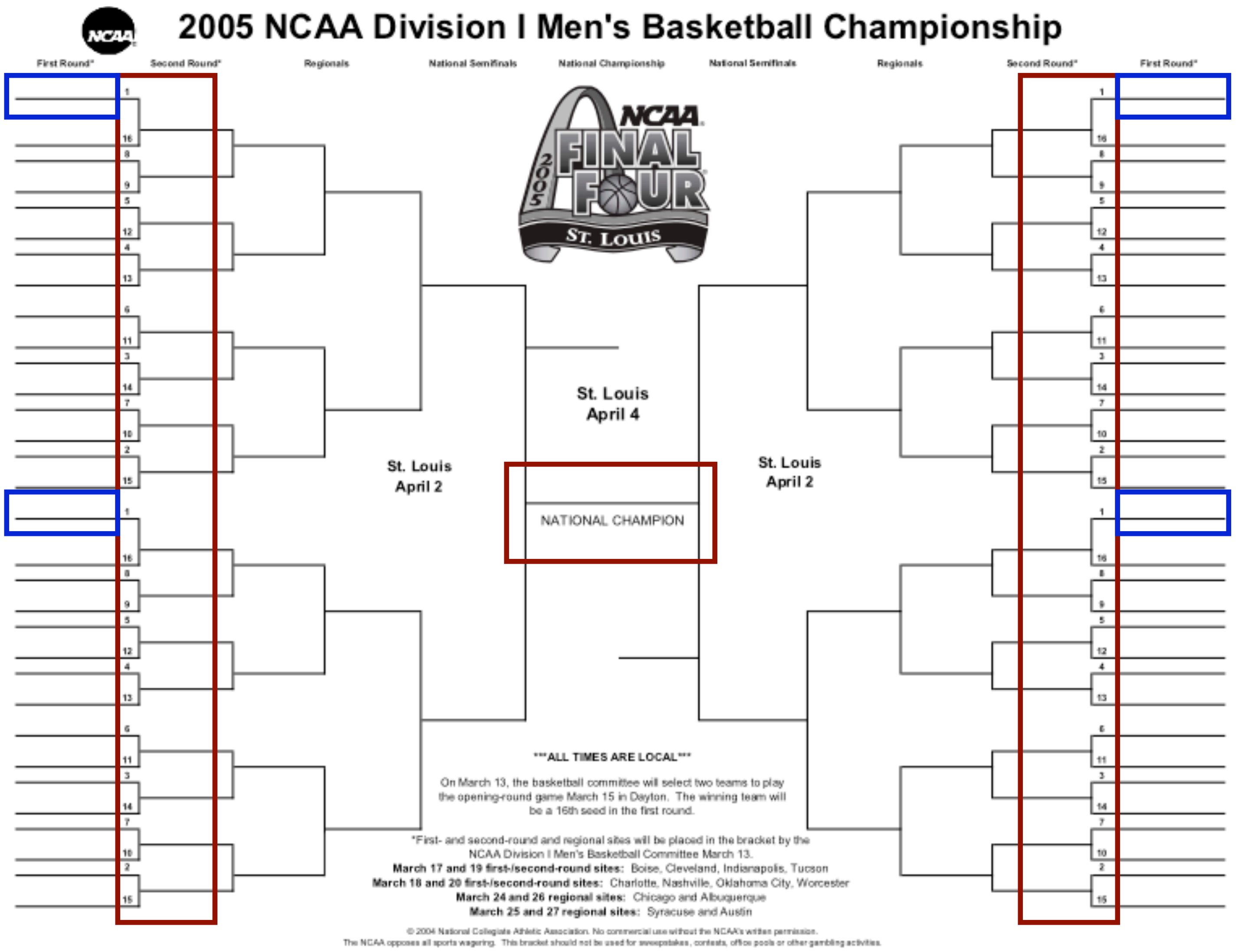
Historical Payouts
What this contract would have paid out in previous years

Year	Payout	Year	Payout
2007	$0	1992	$0
2006	$0	1991	$0
2005	$100	1990	$0
2004	$0	1989	$0
2003	$100	1988	$0
2002	$0	1987	$100
2001	$100	1986	$100
2000	$100	1985	$0

Questions? Call
888-924-7475

See Also:
- Analyze Your Risk
- Grow Your Revenue
- Get a Dutch Frost Days Quote

Example II: March Madness



2005 NCAA Division I Men's Basketball Championship

ALL TIMES ARE LOCAL

On March 13, the basketball committee will select two teams to play the opening-round game March 15 in Dayton. The winning team will be a 16th seed in the first round.

*First- and second-round and regional sites will be placed in the bracket by the NCAA Division I Men's Basketball Committee March 13.

March 17 and 19 first-/second-round sites: Boise, Cleveland, Indianapolis, Tucson
March 18 and 20 first-/second-round sites: Charlotte, Nashville, Oklahoma City, Worcester
March 24 and 26 regional sites: Chicago and Albuquerque
March 25 and 27 regional sites: Syracuse and Austin

Example II: March Madness

- Typical today Non-combinatorial
    - Team wins Rnd 1
    - Team wins Tourney
    - A few other “props”
    - Everything explicit (By def, small #)
    - Every bet indep: Ignores logical & probabilistic relationships

- Combinatorial
    - *Any* property
    - Team wins Rnd k
      Duke > {UNC,NCST}
      ACC wins 5 games
    - $2^{2^{63}}$ possible props (implicitly defined)
    - 1 Bet effects related bets “correctly”; e.g., to enforce logical constraints

Advantages

- More choices -- better hedges
- More information
- Better processing of information: Let traders focus on predicting whatever they want, however they want: Mechanism takes care of logical/probabilistic inference
- Smarter budgeting

Combinatorial Bids vs. Combinatorial Outcomes

- Combinatorial *bids*
  - – Bundling: “Western conference will win”, “Gas prices between 1.75-2.50”
  - – If bids are divisible, almost no disadvantage: use linear programming
- Combinatorial *outcomes*
  - – Outcome space exponential: March Madness, horse racing
  - – Needs combinatorial bids too
  - – Usually intractable but don’t give up hope

[Thanks: Yiling Chen]

Auctioneer vs. Market Maker

- An *auctioneer* only matches buyers & sellers: never takes on any risk. CDA is an example.
- An *automated market maker* is always willing to accept both buy and sell orders at some prices
- Why an institutional market maker? Liquidity!
  - Without market makers, the more expressive the betting mechanism is the less liquid the market is (few exact matches)
  - Illiquidity discourages trading: Chicken and egg
  - Subsidizes information gathering and aggregation: Circumvents no-trade theorems
- Market makers bear risk. But smart pricing algorithms can bound the loss of market makers
  - Market scoring rules [Hanson 2002, 2003, 2006]
  - Family of bounded-loss market makers [Chen & Pennock 2007]
  - Dynamic pari-mutuel market [Pennock 2004]

Combinatorics 1 of 3: Boolean Logic

- Outcomes: All 2^n possible combinations of n Boolean events
- Betting language
  Buy q units of "$1 if Boolean Formula" at price p
  - General: *Any* Boolean formula (2^{2^n} possible)
    - A & not(B)
    - (A&C||F) | (D&E)
    - Oil rises & Hillary wins | Guiliani GOP nom & housing falls
    - Eastern teams win more games than Western in Tourney
  - Restricted languages we study
    - Restricted tournament language
      Team A wins in round i ; Team A beats B, given they meet
    - 2-clauses: A & not(C)

Combinatorics 2 of 3: Permutations

- Outcomes: All possible n! rank orderings of n objects (horse race)
- Betting language
  Buy q units of "$1 if Property" at price p
  - General: *Any* property of ordering
    - A wins
    - A finishes in pos 3,4, or 10th
    - A beats D
    - 2 of {B,D,F} beat A
  - Restricted languages we study
    - Subset betting
      A finishes in pos 3-5 or 9; A,D,or F finish 3rd
    - Pair betting
      A beats F

Combinatorics 3 of 3: Taxonomy

- Outcomes: Cross product of n discretized numbers
- Betting language
  Buy q units of "$1 if Function" at price p
  - General: *Any* mathematical function of the numbers
  - Restricted language we study
    - Taxonomy betting
      Numbers are arranged in a hierarchy
      Parent nodes = sum of children
      Can bet on the range of any node in the hierarchy

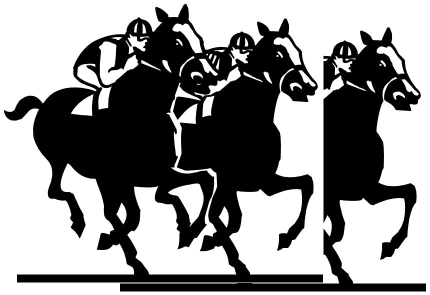
Predicting Permutations

- Predict the ordering of a set of statistics
  - Horse race finishing times
  - Number of votes for several candidates
  - Daily stock price changes
  - NFL Football quarterback passing yards
  - Any ordinal prediction
- Chen, Fortnow, Nikolova, Pennock, EC'07

Market Combinatorics

Permutations

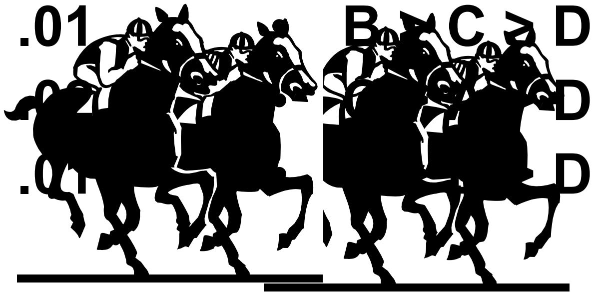
A > B > C	.1
A > C > B	.2
B > A > C	.1
B > C > A	.3
C > A > B	.1
C > B > A	.2



Market Combinatorics

Permutations

D > A > B > C	.01	D > B > C > A	.05
D > A > C > B	.02	D > C > A > B	.1
D > B > A > C	.01	D > C > B > A	.2
A > D > B > C	.01	B > D > C > A	.03
A > D > C > B	.02	C > D > A > B	.1
B > D > A > C	.05	C > D > B > A	.02
A > B > D > C	.01	B > C > D > A	.03
A > C > D > B	.2	C > A > D > B	.01
B > A > D > C	.01	C > B > D > A	.02
A > B > C > D	.01	B > C > D > A	.03
A > C > B > D	.0	D > B	.01
B > A > C > D	.01	D > A	.02



Bidding Languages

- Traders want to bet on *properties* of orderings, not explicitly on orderings: more natural, more feasible
  - A will win ; A will “show”
  - A will finish in [4-7] ; {A,C,E} will finish in top 10
  - A will beat B ; {A,D} will both beat {B,C}
- *Buy 6 units of “$1 if A>B” at price $0.4*
- Supported to a limited extent at racetrack today, but *each in different betting pools*
- Want centralized auctioneer to improve liquidity & information aggregation

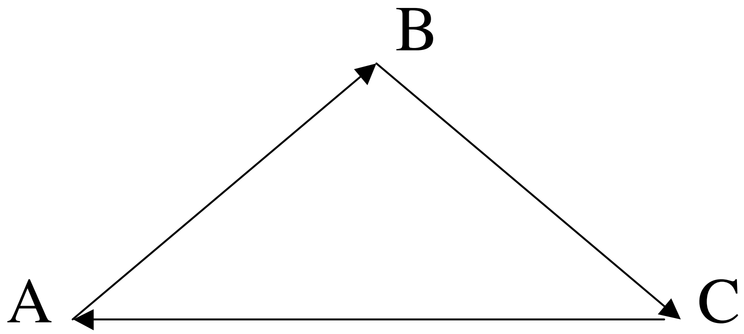
Auctioneer Problem

- **Auctioneer's goal:**
  Accept orders with non-negative worst-case loss (auctioneer never loses money)
- ***The Matching Problem***
- **Formulated as LP**

- **Generalization: *Market Maker Problem*:**
  Accept orders with bounded worst-case loss (auctioneer never loses more than b dollars)

Example

- A three-way match
  - Buy 1 of “$1 if A>B” for 0.7
  - Buy 1 of “$1 if B>C” for 0.7
  - Buy 1 of “$1 if C>A” for 0.7



Pair Betting

- All bets are of the form “A will beat B”
- Cycle with sum of prices > k-1 ==> Match (Find best cycle: Polytime)
- Match =/=> Cycle with sum of prices > k-1

- Theorem: The Matching Problem for Pair Betting is NP-hard (reduce from min feedback arc set)

Subset Betting

- All bets are of the form
  - "A will finish in positions 3-7", or
  - "A will finish in positions 1,3, or 10", or
  - "A, D, or F will finish in position 2"

- Theorem: The Matching Problem for Subset Betting is polytime (LP + maximum matching separation oracle)

Market Combinatorics

Boolean

I am entitled to: $\$1 \text{ if } A1\&A2\&\ldots\&An$

I am entitled to: $\$1 \text{ if } A1\&A2\&\ldots\&\overline{An}$

I am entitled to: $\$1 \text{ if } \overline{A1}\&A2\&\ldots\&An$

I am entitled to: $\$1 \text{ if } \overline{A1}\&A2\&\ldots\&\overline{An}$

I am entitled to: $\$1 \text{ if } A1\&\overline{A2}\&\ldots\&An$

I am entitled to: $\$1 \text{ if } A1\&\overline{A2}\&\ldots\&\overline{An}$

I am entitled to: $\$1 \text{ if } \overline{A1}\&\overline{A2}\&\ldots\&An$

I am entitled to: $\$1 \text{ if } \overline{A1}\&\overline{A2}\&\ldots\&\overline{An}$

- **Betting on complete conjunctions is both unnatural and infeasible**

Market Combinatorics

Boolean

- **A bidding language: write your own security**

I am entitled to: $\$1 \text{ if Boolean\_fn} \mid \text{Boolean\_fn}$

- **For example**

I am entitled to: $\$1 \text{ if } A1 \mid \overline{A2}$

I am entitled to: $\$1 \text{ if } A1\&\overline{A7}$

I am entitled to: $\$1 \text{ if } (A1\&\overline{A7})||A13 \mid (A2||\overline{A5})\&A9$

- **Offer to buy/sell q units of it at price p**
- **Let everyone else do the same**
- **Auctioneer must decide who trades with whom at what price… How? (next)**
- **More concise/expressive; more natural**

The Matching Problem

- **There are many possible matching rules for the auctioneer**
- **A natural one: maximize trade subject to no-risk constraint**
- **Example:**
  - **buy 1 of** $1 if A1 **for $0.40**
  - **sell 1 of** $1 if A1&A2 **for $0.10**
  - **sell 1 of** $1 if A1&$\overline{\text{A2}}$ **for $0.20**
- **No matter what happens, auctioneer cannot lose money**

trader gets $$ in state:

A1A2	A1$\overline{\text{A2}}$	$\overline{\text{A1}}$A2	$\overline{\text{A1A2}}$
0.60	0.60	-0.40	-0.40
-0.90	0.10	0.10	0.10
0.20	-0.80	0.20	0.20
-0.10	-0.10	-0.10	-0.10

Fortnow; Kilian; Pennock; Wellman

Complexity Results

- **Divisible orders: will accept any q* ≤ q**
- **Indivisible: will accept all or nothing**

LP

reduction from X3C

# events	divisible	indivisible
O(log n)	polynomial	NP-complete
O(n)	co-NP-complete	Σ_2^p complete

reduction from SAT

reduction from T∃∀BF

- **Natural algorithms**
  - **divisible: linear programming**
  - **indivisible: integer programming;
logical reduction?**

[Thanks: Yiling Chen]

Automated Market Makers

- **n disjoint and exhaustive outcomes**
- **Market maker maintain vector Q of outstanding shares**
- **Market maker maintains a cost function C(Q) recording total amount spent by traders**
- **To buy ΔQ shares trader pays C(Q+ ΔQ) – C(Q) to the market maker; Negative "payment" = receive money**
- **Instantaneous price functions are** $p_i(Q) = \frac{\partial C(Q)}{\partial q_i}$
- **At the beginning of the market, the market maker sets the initial Q^0, hence subsidizes the market with $C(Q^0)$.**
- **At the end of the market, $C(Q^f)$ is the total money collected in the market. It is the maximum amount that the MM will pay out.**

New Results: Pricing LMSR market maker

- Subset betting on permutations is #P-hard (call market polytime!)
- Pair betting on permutations is #P-hard
- 2-clause Boolean betting #P-hard
- Restricted tourney betting is polytime (uses Bayesian network representation)
- Approximation techniques for general case
- Published in EC’08 and STOC’08

Overview: Complexity Results

	Permutations			Boolean			Taxonomy	
	General	Pair	Subset	General	2-clause	Restrict Tourney	General	Tree
Auctioneer	NP-hard EC’07	NP-hard EC’07	Poly EC’07	NP-hard DSS’05	co-NP-complete DSS’05	?	?	?
Market Maker (LMSR)	#P-hard EC’08	#P-hard EC’08	#P-hard EC’08	#P-hard EC’08 **Approx** **STOC’08**	#P-hard EC’08	Poly STOC’08	#P-hard AAMAS ‘09	Poly AAMAS ‘09

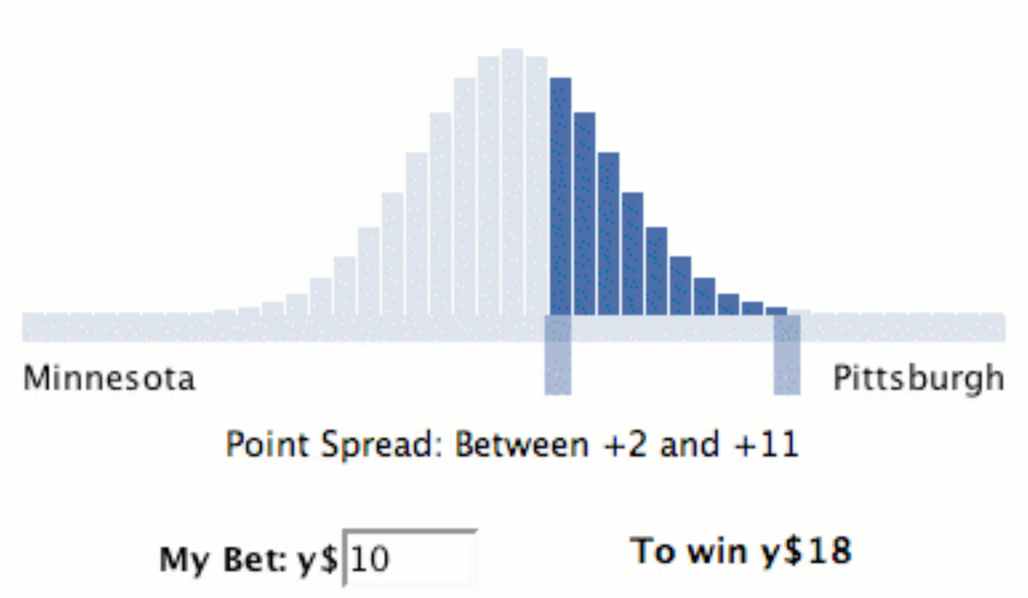
More Info

What is (and what good is) a combinatorial prediction market?

`http://blog.oddhead.com/2008/12/22/what-is-and-what-good-is-a-combinatorial-prediction-market/`

- **March Madness bet constructor**
- Bet on any team to win any game
  - Duke wins in Final 4
- Bet “exotics”:
  - Duke advances further than UNC
  - ACC teams win at least 5
  - A 1-seed will lose in 1st round

New Prediction Game: Yoopick



Point Spread: Between +2 and +11

My Bet: y$ 10

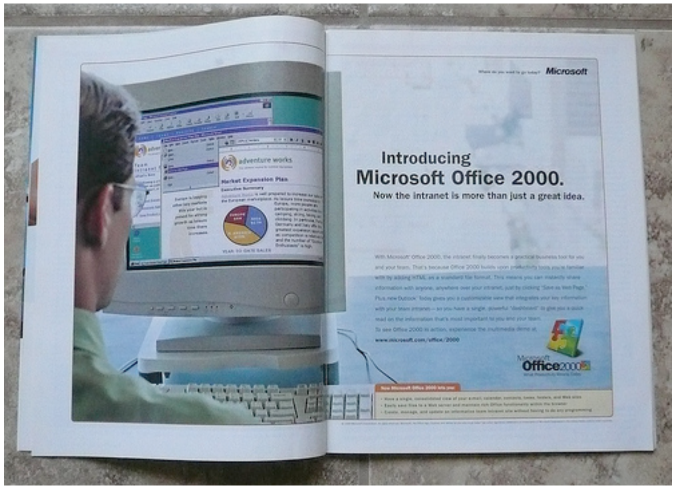
To win y$18

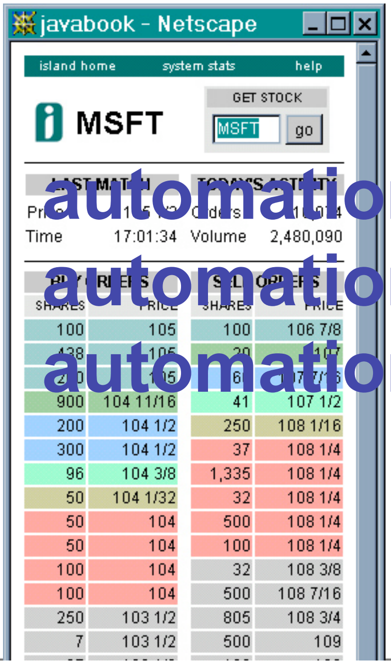
CFTC Role

- MayDay 2008: CFTC asks for help
- Q: What to do with prediction markets?
- Right now, the biggest prediction markets are overseas, academic (1), or just for fun
- CFTC may clarify, drive innovation
- Or not

Advertising Then and Now

- Then: Think real estate
  Phone calls
  Manual negotiation
  “Half doesn’t work”





LAST MATCH		TODAY'S ACTIVITY	
Price	[illegible]	Orders	[illegible]
Time	17:01:34	Volume	2,480,090

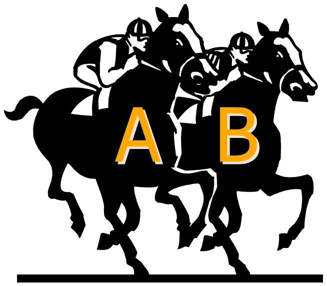
BUY ORDERS		SELL ORDERS	
SHARES	PRICE	SHARES	PRICE
100	105	100	106 7/8
438	105	20	107
[illegible]	105	[illegible]	107 7/16
900	104 11/16	41	107 1/2
200	104 1/2	250	108 1/16
300	104 1/2	37	108 1/4
96	104 3/8	1,335	108 1/4
50	104 1/32	32	108 1/4
50	104	500	108 1/4
50	104	100	108 1/4
100	104	32	108 3/8
100	104	500	108 7/16
250	103 1/2	805	108 3/4
7	103 1/2	500	109

automation
automation
automation

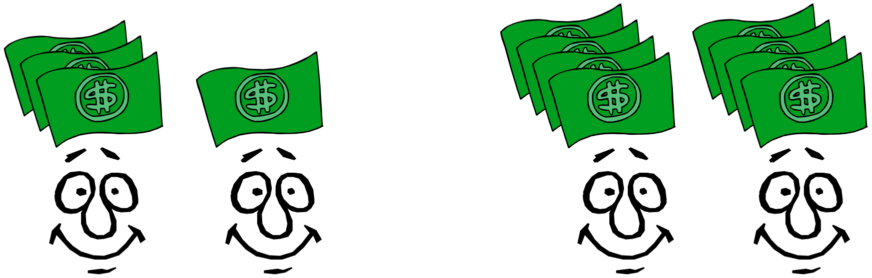
- Now: Think Wall Street
  Computer learns what ad is best
  Computer mediates ad sales: Auction
  Computer measures which ads work
  Advertisers buy *contextual events*:
    User i views/clicks/converts
    on page j at time t

Dynamic Parimutuel Market: An Automated Market Maker

What is a pari-mutuel market?

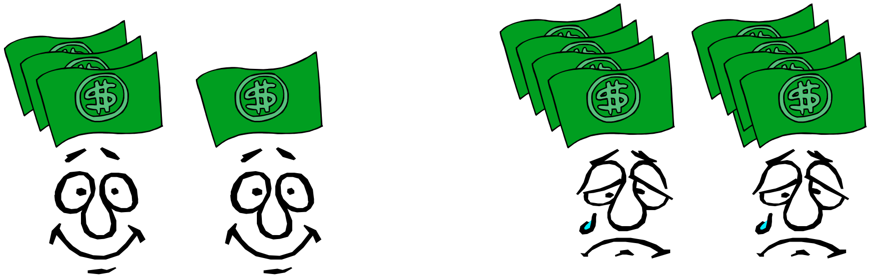


- E.g. horse racetrack style wagering
- Two outcomes: A B
- Wagers:



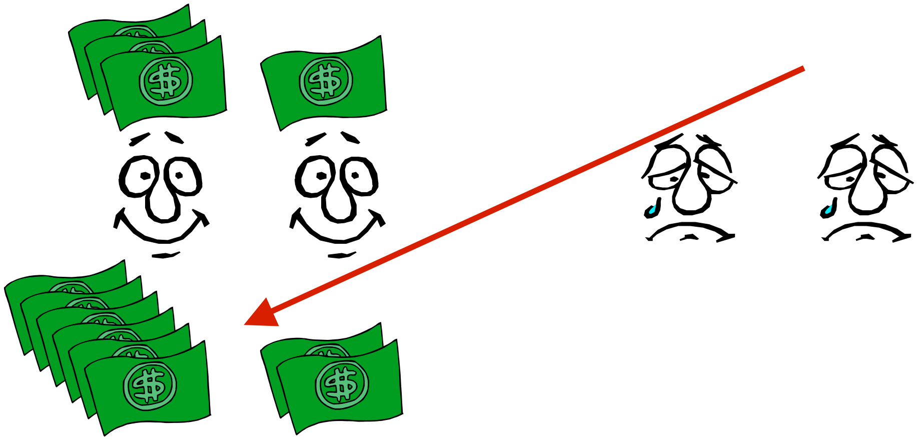
What is a pari-mutuel market?

- E.g. horse racetrack style wagering
- Two outcomes: √A B
- Wagers:



What is a pari-mutuel market?

- E.g. horse racetrack style wagering
- Two outcomes: √ A B
- Wagers:

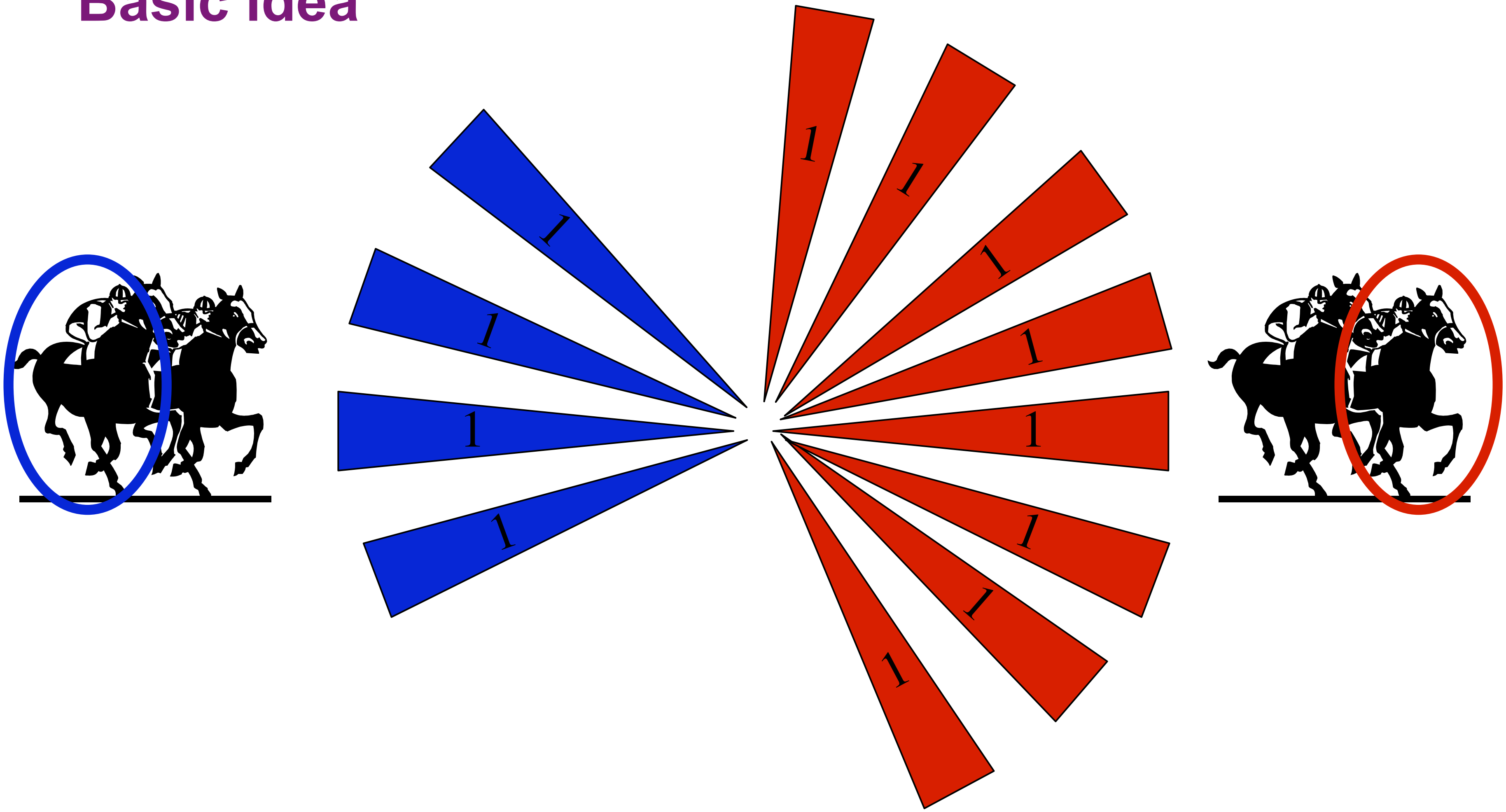


What is a pari-mutuel market?

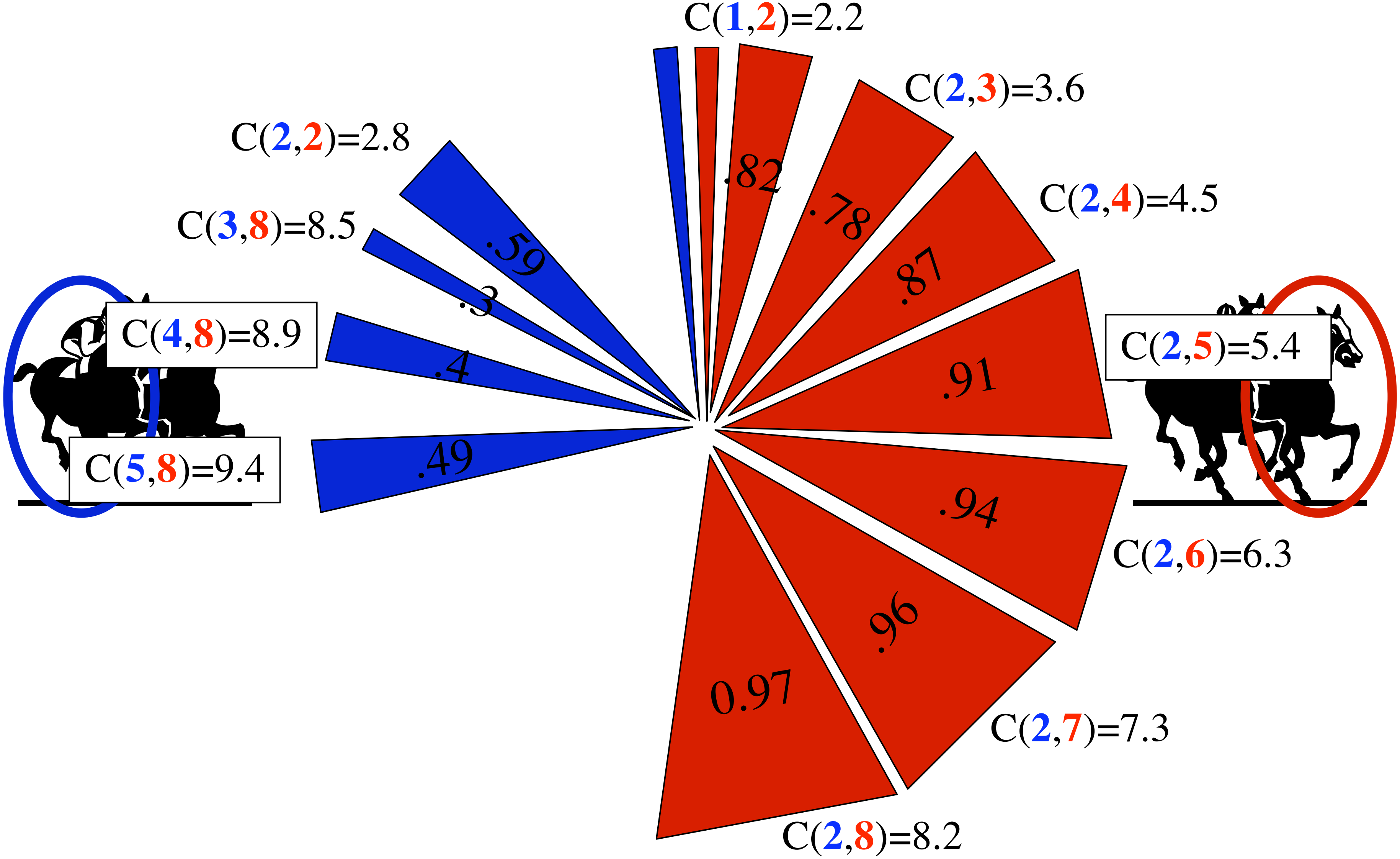
- **Before outcome is revealed, "odds" are reported, or the amount you *would* win per dollar *if* the betting ended now**
  - **Horse A: $1.2 for $1; Horse B: $25 for $1; … etc.**
- **Strong incentive to wait**
  - **payoff determined by *final* odds; every $ is same**
  - **Should wait for best info on outcome, odds**
  - **⇒ No continuous information aggregation**
  - **⇒ No notion of "buy low, sell high" ; no cash-out**

Pari-Mutuel Market

Basic idea



Dynamic Parimutuel Market



Share-ratio price function

- One can view DPM as a market maker
- Cost Function:

$$C(Q) = \sqrt{\sum_{i=1}^{n} q_i^{\ 2}}$$

- Price Function:

$$p_i(Q) = \frac{q_i}{\sqrt{\sum_{j=1}^{n} q_j^{\ 2}}}$$

- Properties
  - No arbitrage
  - $\text{price}_i/\text{price}_j = q_i/q_j$
  - $\text{price}_i < \$1$
  - payoff if right = $C(Q_{\text{final}})/q_o > \$1$

Mech Design for Prediction

	Financial Markets	Prediction Markets
Primary	Social welfare (trade) Hedging risk	Information aggregation
Secondary	Information aggregation	Social welfare (trade) Hedging risk

Mech Design for Prediction

- Standard Properties
  - ~~Efficiency~~
  - Inidiv. rationality
  - ~~Budget balance~~
  - ~~Revenue~~
  - Truthful (IC)
  - Comp. complexity
- Equilibrium
  - General, Nash, ...

- PM Properties
  - #1: Info aggregation
  - Expressiveness
  - Liquidity
  - Bounded budget
  - Truthful (IC)
  - Indiv. rationality
  - Comp. complexity
- Equilibrium
  - Rational expectations

Competes with: experts, scoring rules, opinion pools, ML/stats, polls, Delphi