The Emergent Science of the Internet and the Worldwide Web

Christos H. Papadimitriou

UC Berkeley

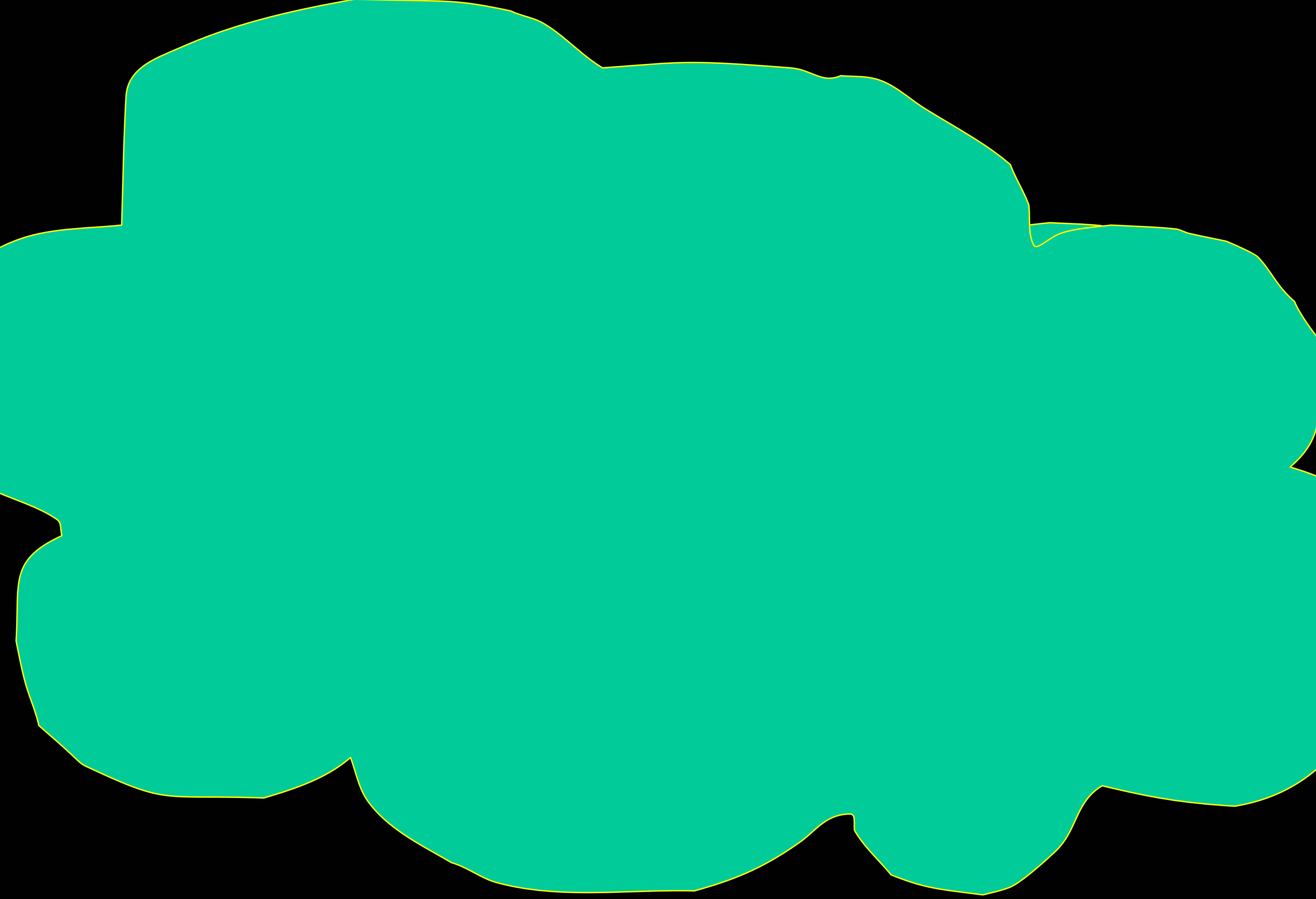
Google christos

- Goal of TCS (1950-2000):

Develop a mathematical understanding of the capabilities and limitations of the von Neumann computer and its software –the dominant and most novel computational artifacts of that time

(Mathematical tools: combinatorics, logic)

- What should Theory's goals be today?



The Internet

- Huge, growing, open, end-to-end
- Built and operated by 15.000 companies in various *(and varying)* degrees of competition
- The first computational artefact that must be studied by observations, measurements, and the development of falsifiable theories (like the universe, the brain, the cell, the market)

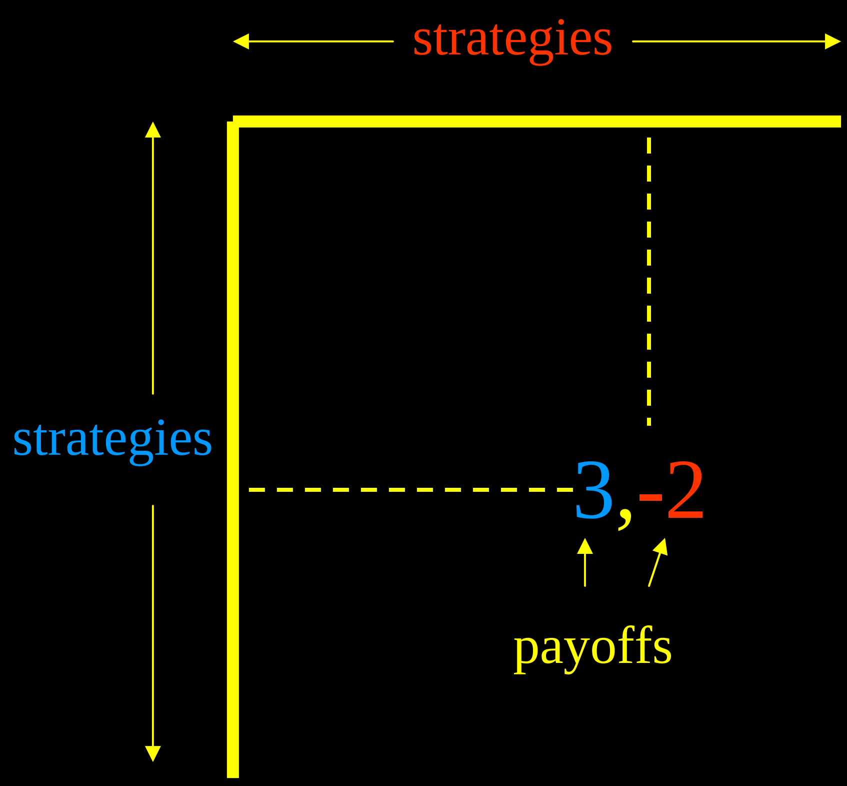
The Internet (cont.)

- ***The platform for the worldwide web,***
  an information repository that is to an unprecedented degree universal, unstructured, heterogeneous, available, and critical
- Theoretical understanding urgently needed
- *Tools: math economics and game theory, probability, graph theory, spectral theory*

Sources on Game Theory and Microeconomics

- Osborne and Rubinstein *A Course in Game Theory*, MIT, 1994
- Mas-Colell, Whinston, and Greene *Microeconomic Theory*, Oxford 1995
- Kreps *A Course on Microeconomic Theory*
- Varian *Microeconomics*
- http://www.cs.berkeley.edu/~christos/games/cs294.html and …/focs01.ppt

Game Theory



(NB: also, many players)

e.g.

matching pennies

1,-1	-1,1
-1,1	1,-1

prisoner's dilemma

3,3	0,4
4,0	1,1

chicken

0,0	0,1
1,0	-1,-1

concepts of rationality

- undominated strategy

  (problem: too weak)
- (weakly) dominating srategy (*alias* “duh?”)

  (problem: too strong, rarely exists)
- Nash equilibrium (or double best response)

  (problem: may not exist)
- randomized Nash equilibrium

Theorem [Nash 1952]: Always exists.

if a digraph with all in-degrees ≤ 1 has a source,

then it must have a sink

$\Rightarrow$ Sperner's Lemma

$\Rightarrow$ Brouwer's fixpoint Theorem

($\Rightarrow$ Kakutani's Theorem $\Rightarrow$ market equilibrium)

$\Rightarrow$Nash's Theorem

?

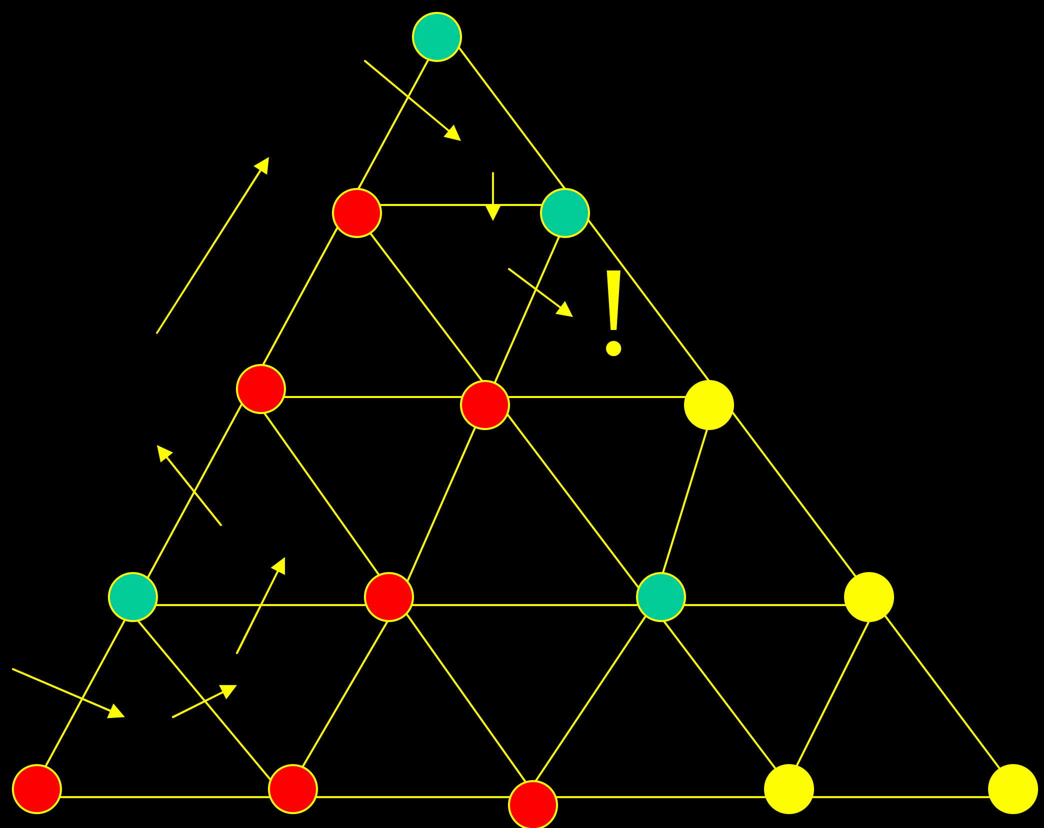
$\Rightarrow$ min-max theorem for zero-sum games

$\Rightarrow$ linear programming duality

$\in$ P

Sperner's Lemma: Any "legal" coloring of the triangulated simplex has a trichromatic triangle

Proof:



Sperner $\Rightarrow$ Brouwer

Brouwer's Theorem: Any continuous function from the simplex to itself has a fixpoint.

Sketch: Triangulate the simplex

Color vertices according to "which direction they are mapped"

Sperner's Lemma means that there is a triangle that has "no clear direction"

Sequence of finer and finer triangulations, convergent subsequence of the centers of Sperner triangles, QED

Brouwer $\Rightarrow$ Nash

For any pair of mixed strategies x,y (distributions over the strategies) define

$\varphi(x,y) = (x', y')$, where x' maximizes $\text{payoff}_1(x',y) - |x - x'|^2$, and similarly for y'.

Any Brouwer fixpoint is now a Nash equilibrium

Nash $\Rightarrow$ von Neumann

If game is zero-sum, then double best response is a max-min pair:

Therefore, $\min_y \max_x xAy^T = \max_x \min_y xAy^T$

The critique of mixed Nash equilibrium

- Is it really rational to randomize?

  (*cf:* bluffing in poker, tax audits)
- If (x,y) is a Nash equilibrium, then any y' *with the same support* is as good as y

  (corollary: problem is combinatorial!)
- Convergence/learning results mixed
- ***There may be too many Nash equilibria***

is it in P?

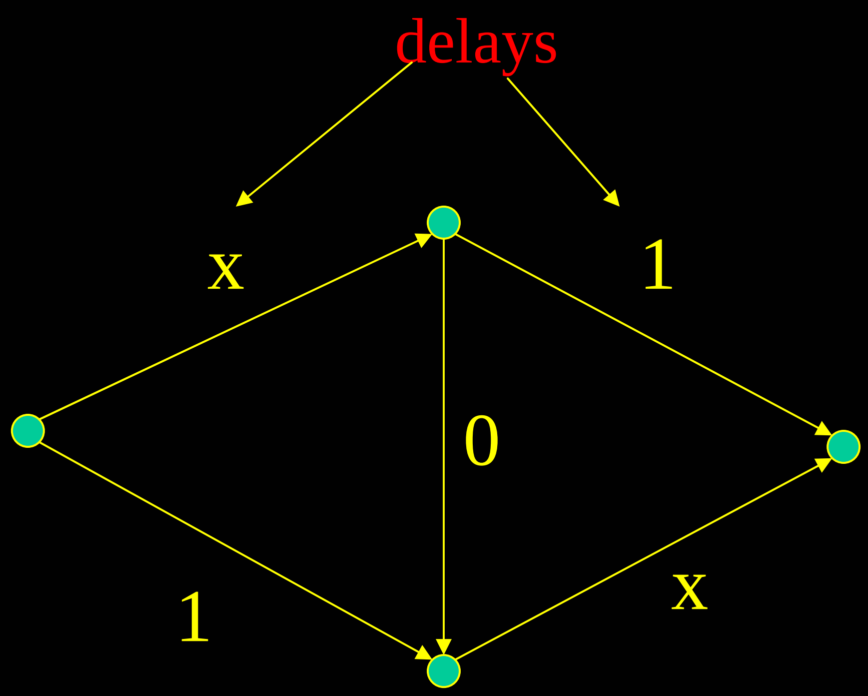
The price of anarchy

$$\frac{\text{cost of worst Nash equilibrium}}{\text{“socially optimum” cost}}$$

[Koutsoupias and P, 1998]

Also: [Spirakis and Mavronikolas 01, Roughgarden 01, Koutsoupias and Spirakis 01]

Selfishness can hurt you!



Social optimum: 1.5

Anarchical solution: 2

Worst case?

Price of anarchy $= 2$ (4/3 for linear delays)

[Roughgarden and Tardos, 2000, Roughgarden 2002]

The price of the Internet architecture?

Simple net creation game

(with Fabrikant, Maneva, Shenker PODC 03)

- Players: Nodes $V = \{1, 2, \ldots, n\}$
- Strategies of node i: all possible subsets of $\{[i,j]: j \neq i\}$
- Result is *undirected* graph $G = (s_1,\ldots,s_n)$
- Cost to node i:

$$c_i[G] = \underbrace{\alpha\cdot}_{\text{cost of edges}} | s_i | + \Sigma_i \, \underbrace{\text{dist}_G}_{\text{delay costs}}(i,j) \cdot (\text{traffic}_{ij})$$

$$(w_i \cdot w_j)$$

Nash equilibria?

- (*NB: Best response is NP-hard…*)
- Let us fix $w_i = 1$
- If $\alpha < 1$, then the only Nash equilibrium is the clique
- If $1 < \alpha < 2$ then social optimum is clique, Nash equilibrium is the star (price of anarchy = 4/3)

Nash equilibria (cont.)

- $\alpha > 2$? The price of anarchy is at least 3
- Upper bound: $\sqrt{\alpha}$
- **Conjecture:** For large enough α, all Nash equlibria are trees.
- If so, the price of anarchy is at most 5.
- **General w_i** : *Are the degrees of the Nash equilibria proportional to the w_i's?*

mechanism design
(or *inverse* game theory)

- agents have utilities – but these utilities are known *only to them*
- game designer prefers certain outcomes *depending on players' utilities*
- designed game (mechanism) has designer's goals as dominating strategies (or other rational outcomes)

mechanism design (math)

- n players, set K of outcomes, for each player i a possible set U_i of utilities of the form u: $K \rightarrow R^+$
- designer preferences P: $U_1 \times \ldots \times U_n \rightarrow 2^K$
- mechanism: strategy spaces S_i, plus a mapping G: $S_1 \times \ldots \times S_n \rightarrow K$

Theorem (The Revelation Principle): If there is a mechanism, then there is one in which all agents truthfully reveal their secret utilities (direct mechanism).

Proof: common-sense simulation

Theorem (Gibbard-Satterthwaite): If the sets of possible utilities are too rich, then only dictatorial *P*'s have mechanisms.

Proof: Arrow's Impossibility Theorem

- but… if we allow mechanisms that use Nash equilibria instead of dominance, then almost anything is implementable
- but… these mechanisms are extremely complex and artificial

  (complexity-theoretic critique would be welcome here…)

- but… if outcomes in K include payments ($K = K_0 \times R^n$) and utilities are *quasilinear* (utility of "core outcome" plus payment) *and designer prefers to optimize the sum of core utilities*, then the Vickrey-Clarke-Groves mechanism works

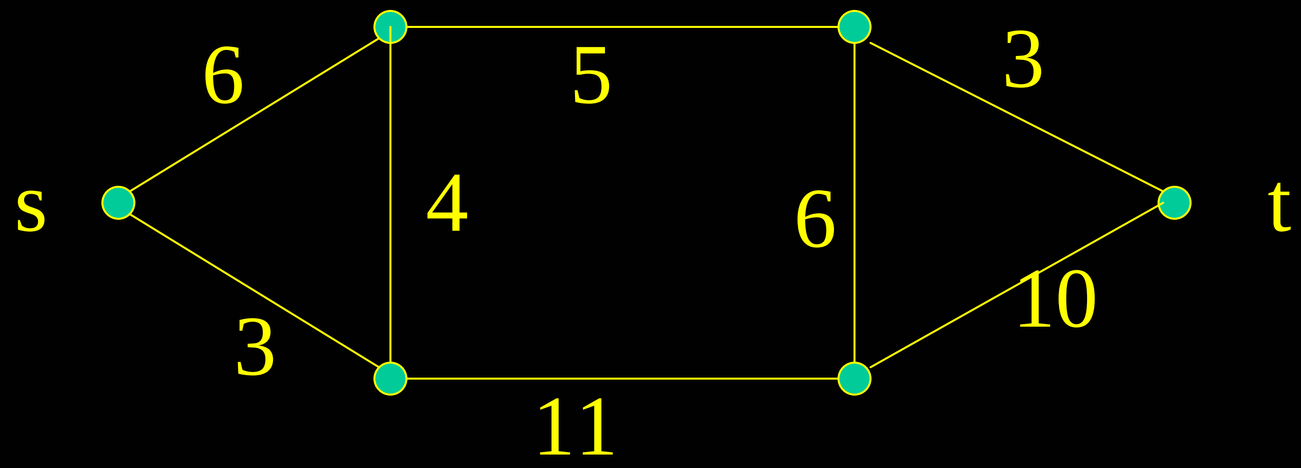
e.g., Vickrey auction

- sealed-highest-bid auction encourages gaming and speculation
- Vickrey auction: Highest bidder wins, pays second-highest bid

Theorem: Vickrey auction is a truthful mechanism.

Theorem: It maximizes social benefit *and* auctioneer expected revenue.

e.g., shortest path auction

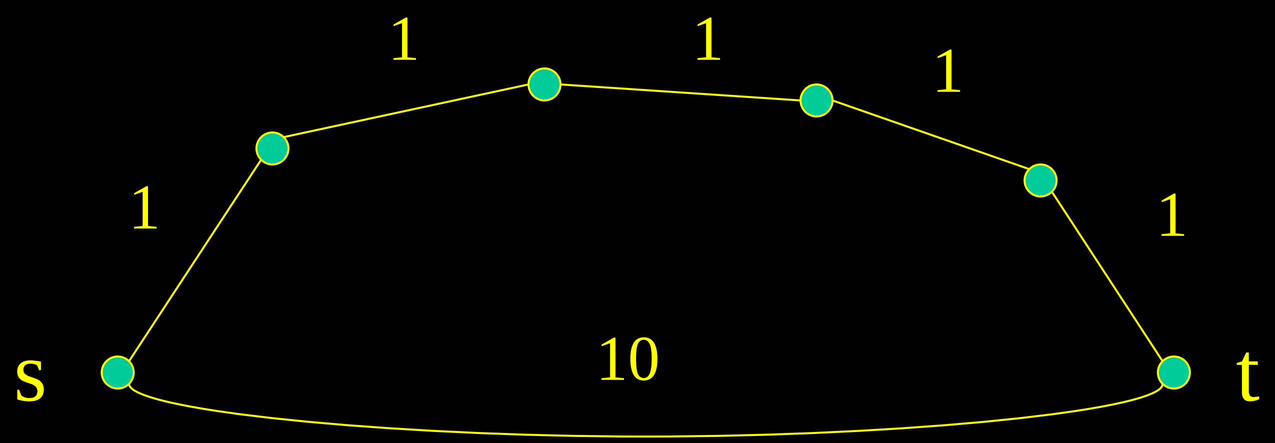


pay *e* its declared cost c(*e*),

plus a bonus equal to $\text{dist}(s,t)|_{c(e)=\infty} - \text{dist}(s,t)$

Theorem [Suri & Hershberger 01]:
Payments can be computed by one shortest path computation.

Problem:



Theorem [Elkind, Sahai, Steiglitz, 03]: This is inherent for truthful mechanisms.

But…

- …in the Internet (the graph of autonomous systems) VCG overcharge would be only about 30% on the average [FPSS 2002]
- Could this be the manifestation of rational behavior at network creation?

Also…

- In Internet routing, VCG[e] *depends on the origin and destination.*
- Can be computed with little overhead on top of BGP (the standard protocol for interdomain routing).
- **Theorem** [with Mihail and Saberi, 2003]**:** In a random graph with average degree d, the expected VCG overcharge is constant (***conjectured:*** $\sim 1/d$)

e.g., 2-processor scheduling

[Nisan and Ronen 1998]

- two players/processors, *n* tasks, each with a different execution time on each processor
- each execution time is known only to the appropriate processor
- designer wants to minimize makespan

(= maximum completion time)

- each processor wants to minimize its own completion time

Idea: Allocate each task to the most efficient processor (i.e., minimize total work). Pay each processor for each task allocated to it an amount equal to the time required for it *at the other processor*

Fact: Truthful and 2-approximate

Theorem (Nisan-Ronen) : No mechanism can achieve ratio better than 2

Sketch: By revelation, such a mechanism would be truthful.

wlog, Processor 1 chooses between proposals of the form (partition, payment), where the payment depends only on the partition and Processor 2's declarations

Theorem (Nisan-Ronen, continued):

Suppose all task lengths are 1, and Processor 1 chooses a partition and a payment

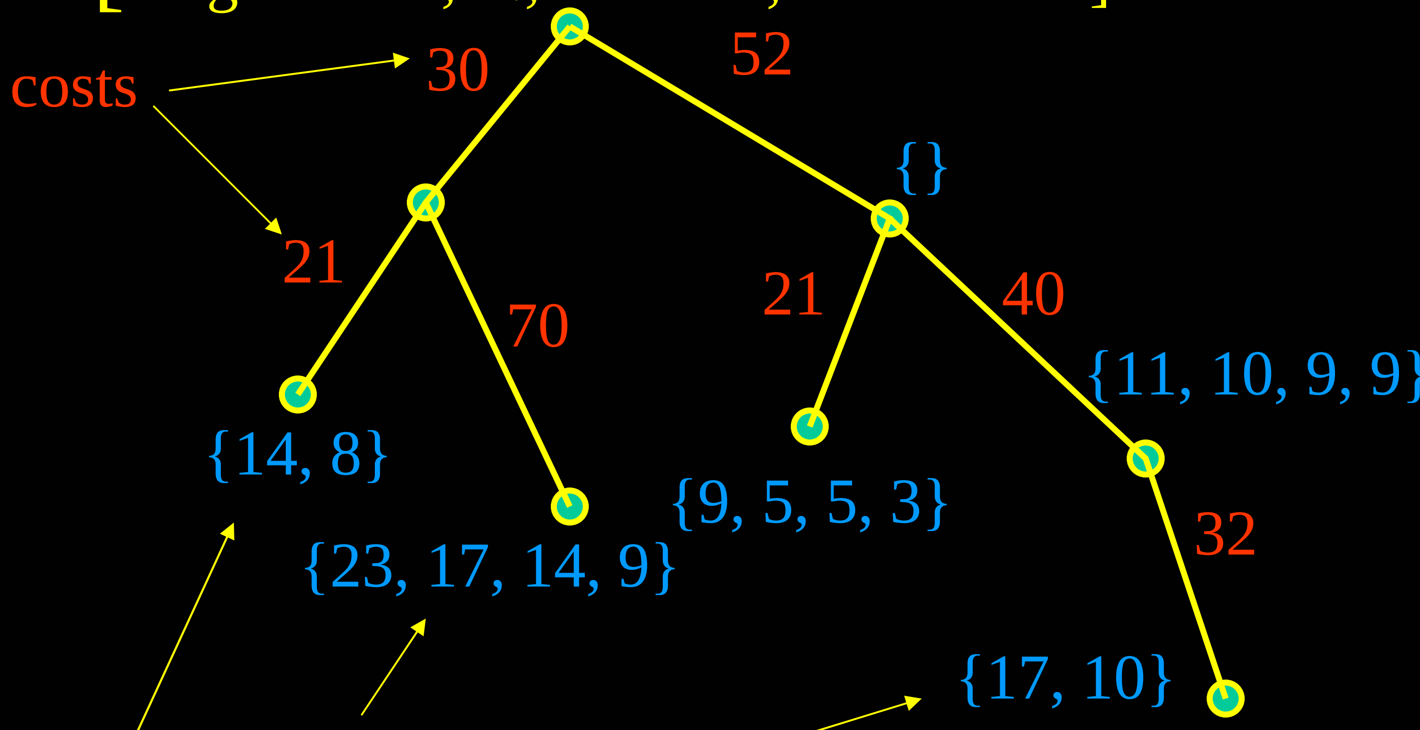
If we change the 1-lengths in the partition to ε and all others to 1 + ε, it is not hard to see that the proposals will remain the same, and Processor 1 will choose the same one

But this is ~2-suboptimal, QED

Also: k processors, randomized 7/4 algorithm.

e.g., pricing multicasts

[Feigenbaum, P., Shenker, STOC2000]



utilities of *agents* in the node
(u_i = the intrinsic value of the information
to agent *i*, *known only to agent i*)

We wish to design a protocol that will result in the computation of:

- x_i (= 0 or 1, will i get it?)
- v_i (how much will i pay? (0 if $x = 0$))

protocol must obey a set of desiderata:

- $0 \le v_i \le u_i,$

- $\lim_{u_i \to \infty} x_i = 1$

- *strategy proofness*: $(w_i \stackrel{\text{def}}{=} u_i \cdot x_i - v_i)$
  $w_i(u_1 \ldots u_i \ldots u_n) \ge w_i(u_1 \ldots u_i' \ldots u_n)$

- welfare maximization

$$\sum u_i x_i - c[T] = \max$$

marginal cost mechanism

- budget balance

$$\sum v_i = c(T[x])$$

Shapley mechanism

But…

In the context of the Internet, there is another desideratum:

Tractability: the protocol should require few (constant? logarithmic?) messages per link.

This new requirement changes drastically the space of available solutions.

- $0 \le v_i \le u_i$
- $\lim_{u_i \to \infty} x_i = 1$
- *strategy proofness*: $(w_i \stackrel{\text{def}}{=} u_i \cdot x_i - v_i)$

  $w_i(u_1 \ldots u_i \ldots u_n) \ge w_i(u_1 \ldots u_i' \ldots u_n)$

- welfare maximization

  $\sum w_i = \max$

  marginal cost mechanism

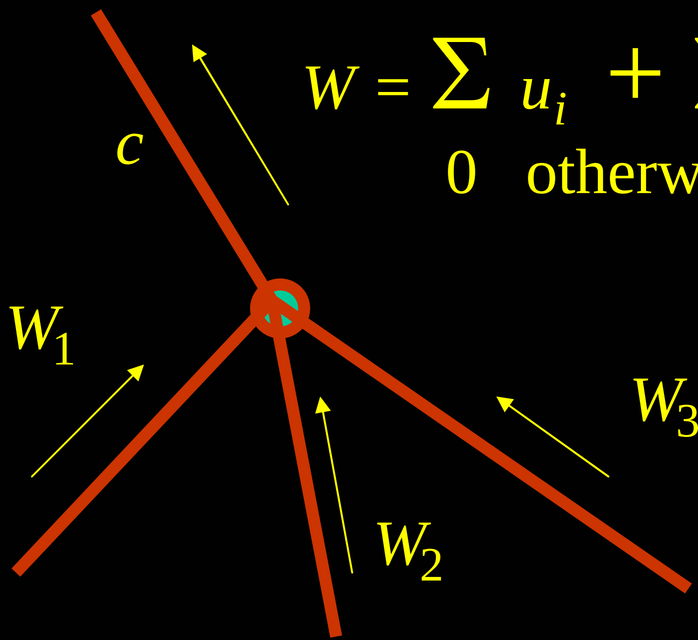
- ~~budget balance~~

  $\cancel{\sum v_i = c(T[x])}$

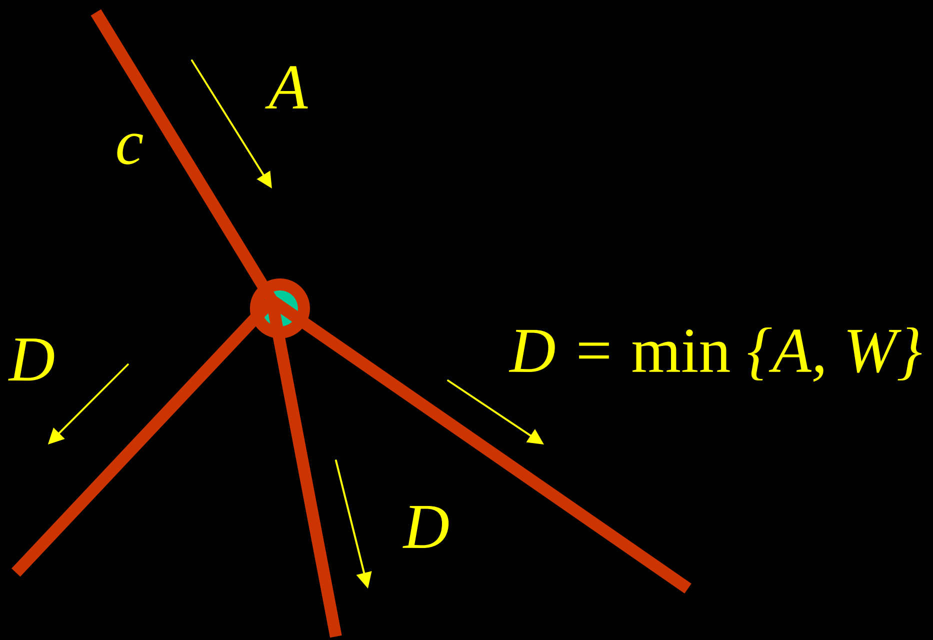
  ~~Shapley mechanism~~

Bottom-up phase

$$W = \begin{cases} \sum u_i + \sum W_j - c, & \text{if } > 0 \\ 0 & \text{otherwise} \end{cases}$$



Top-down phase

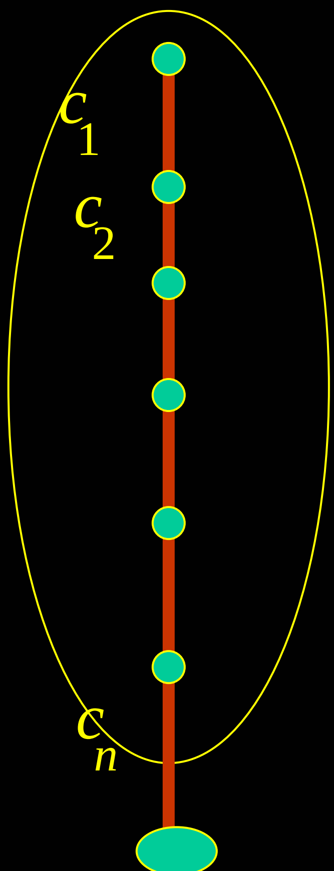


$$v_i = \max\{0, u_i - D\}$$

Theorem: The marginal cost mechanism is tractable.

Theorem: "The Shapley value mechanism is intractable."

Model: Nodes are linear decision trees, and they exchange messages that are linear combinations of the *u*'s and *c*'s



It reduces to checking whether $Au > Bc$ by two sites, one of which knows u and the other c, where A, B are nonsingular

agents drop out one-by-one

$\{u_1 < u_2 < \ldots < u_n\}$

Algorithmic Mechanism Design

- central problem
- few results outside "social welfare maximization" framework (n.b.[Archer and Tardos 01])
- VCG mechanism often breaks the bank
- approximation rarely a remedy (n.b.[Nisan and Ronen 99, Jain and Vazirani 01])
- wide open, radical departure needed

algorithmic aspects of auctions

- Optimal auction design [Ronen 01]
- Combinatorial auctions [Nisan 00]
- Auctions for digital goods [Goldberg, Hartline, 01]
- On-line auctions [Kearns, Wong 02]
- Communication complexity of combinatorial auctions [Nisan-Segal 01]

So…. Game Theory and Math Economics:

- Deep and elegant
- Different
- Exquisite interaction with CS
- Relevant to the Internet
- Wide open algorithmic aspects
- Mathematical tools of choice for the "new TCS"