

The changing face of web search



Prabhakar Raghavan

Yahoo! Research



What is web search?

- Access to “heterogeneous”, distributed information
 - Heterogeneous in creation
 - Heterogeneous in accuracy
 - Heterogeneous in motives
- Multi-billion dollar business
- Source of new opportunities in marketing
- Strains the boundaries of trademark and intellectual property laws
- A source of unending technical challenges



What is web search?

- Nexus of
 - Sociology
 - Economics
 - Law
- ... with technical implications.



The driver

- Pew Study (US users Aug 2004):

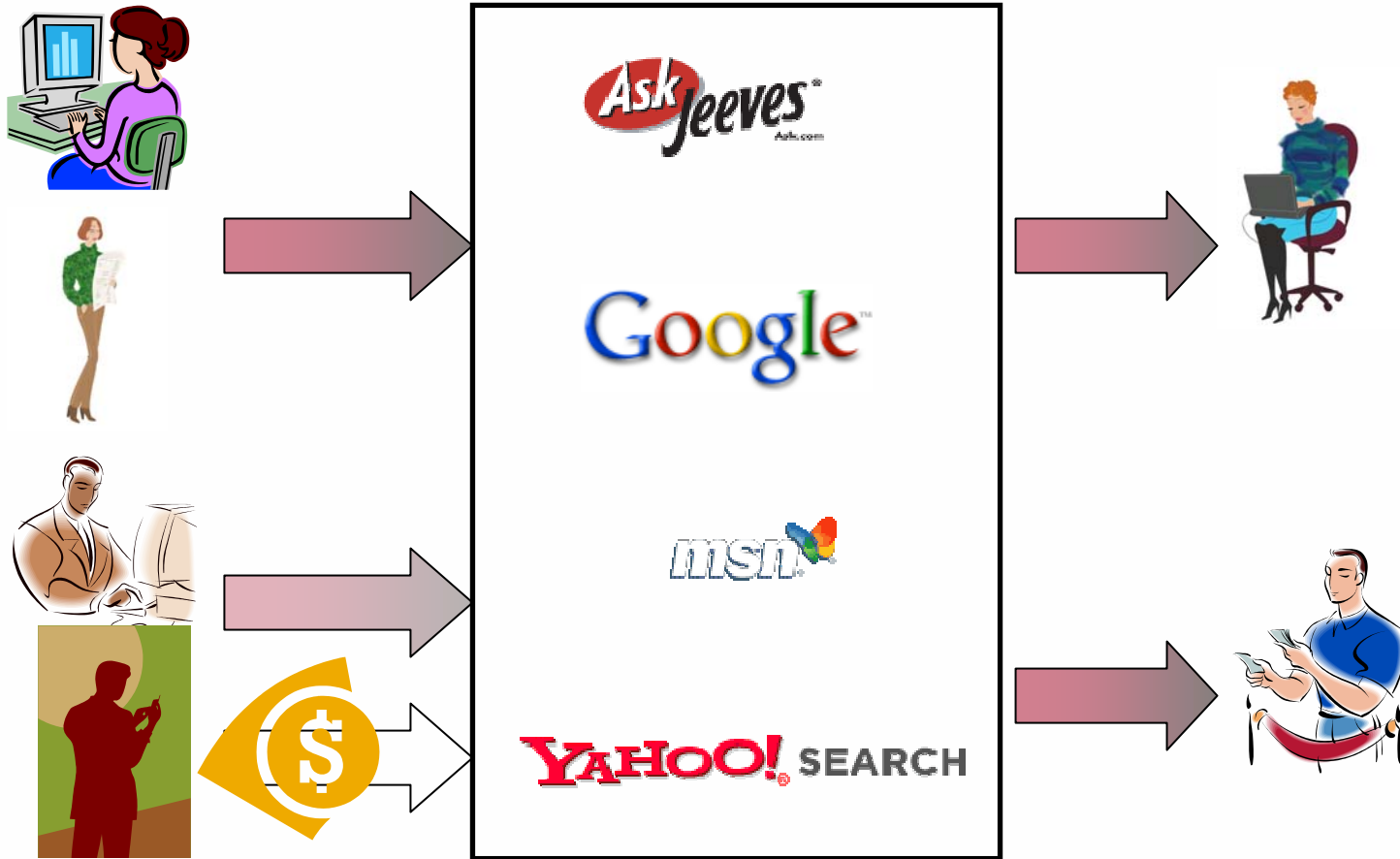
“Getting information is the most highly valued and most popular type of everyday activity done online”.

http://www.pewinternet.org/pdfs/PIP_Internet_and_Daily_Life.pdf

<http://cs276.stanford.edu>



The coarse-level dynamics



Content creators

Content aggregators

Content consumers



Brief (non-technical) history

- Early keyword-based engines
 - Altavista, Excite, Infoseek, Inktomi, Lycos, ca. 1995-1997
- Paid placement ranking: Goto (morphed into Overture → Yahoo!)
 - Your search ranking depended on how much you paid
 - Auction for keywords: casino was expensive!



Brief (non-technical) history

- 1998+: Link-based ranking pioneered by Google
 - Blew away all early engines except Inktomi
 - Great user experience in search of a business model
 - Meanwhile Goto/Overture's annual revenues were nearing \$1 billion



Brief (non-technical) history

- Result: Google added “paid-placement” ads to the side, separate from search results
- 2003: Yahoo follows suit, acquiring Overture (for paid placement) and Inktomi (for search)


 Web [Images](#) [Groups](#) [News](#) [Froogle](#) [Local](#) [more »](#)

nigritude ultramarine

Search

[Advanced Search](#)
[Preferences](#)

Web

Results 1 - 10 of about 185,000 for **nigritude ultramarine**. (0.35 seconds)

[Dash: Nigritude Ultramarine](#)

... as a favor: Link to this post with the phrase **Nigritude Ultramarine**. ... Just placed a link to my **Nigritude Ultramarine** article on my weblog. Cheers! ...

[www.dashes.com/ani/2004/06/04/nigritude_ultra](#) - 101k - Mar 1, 2006 -

[Cached](#) - [Similar pages](#)

[Nigritude Ultramarine FAQ](#)

Nigritude Ultramarine FAQ - frequently asked questions about **nigritude ultramarine** and the realted SEO contest.

[www.nigritudeultramarines.com/](#) - 59k - [Cached](#) - [Similar pages](#)

[SEO contest - Wikipedia, the free encyclopedia](#)

Nigritude ultramarine competition by SearchGuild is widely acclaimed as ...

... comparison of search results for **nigritude ultramarine** during and after the ...

[wikipedia.org/wiki/Nigritude_ultramarine](#) - 37k - [Cached](#) - [Similar pages](#)

[slashdot | How To Get Googled, By Hook Or By Crook](#)

... current 3rd result showcases the "**Nigritude Ultramarine** Fighting Force" who ... When

... discussing **nigritude ultramarine** [slashdot.org] it is important to ...

[slashdot.org/article.pl?sid=04/05/09/1840217](#) - 110k - [Cached](#) - [Similar pages](#)

[Nigritude Ultramarine Search Engine Optimization Contest](#)

... sweeping the web -- or at least search engine optimizers -- a new contest to rank tops for the term **nigritude ultramarine** on Google.

[www.chenginewatch.com/sereport/article.php/3360231](#) - 57k - [Cached](#) - [Similar pages](#)

Sponsored Links

[Business Blogging Seminar](#)

Coming to L.A. March 16

Top bloggers reveal key techniques

[www.blogbusinesssummit.com](#)

Los Angeles, CA

[Full-Time SEO & SEM Jobs](#)

Find companies big & small hiring

full-time SEO & SEM pros right now

[CareerBuilder.com](#)

[SEO Contests](#)

Information on SEO Contests like

the **Nigritude Ultramarine** contest.

[www.seo-contests.com/](#)

[The SEO Book](#)

Nigritude Ultramarine & SEO secrets

Fun, free, raw, & different.

[www.seobook.com](#)

[Ultramarine - Companion](#)

Music - Dance - Electronic

Overstock.com



Ads vs. search results

- Google has said that **ads** (based on vendors bidding for keywords) do not affect **ranking** in search results

Sponsored Links

[CG Appliance Express](#)

Discount Appliances (650) 756-3931
Same Day Certified Installation
www.cgappliance.com
San Francisco-Oakland-San Jose, CA

[Miele Vacuum Cleaners](#)

Miele Vacuums- Complete Selection
Free Shipping!
www.vacuums.com

[Miele Vacuum Cleaners](#)

Miele-Free Air shipping!
All models. Helpful advice.
www.best-vacuum.com

Web

Results 1 - 10 of about 7,310,000 for **miele**. (0.12 seconds)

[Miele, Inc -- Anything else is a compromise](#)

At the heart of your home, Appliances by **Miele**. ... USA. to **miele.com**. Residential Appliances. Vacuum Cleaners. Dishwashers. Cooking Appliances. Steam Oven. Coffee System ...
www.miele.com/ - 20k - [Cached](#) - [Similar pages](#)

[Miele](#)

Welcome to **Miele**, the home of the very best appliances and kitchens in the world.
www.miele.co.uk/ - 3k - [Cached](#) - [Similar pages](#)

[Miele - Deutscher Hersteller von Einbaugeräten, Hausgeräten ...](#) - [[Translate this page](#)]

Das Portal zum Thema Essen & Geniessen online unter www.zu-tisch.de. **Miele** weltweit ...ein Leben lang. ... Wählen Sie die **Miele** Vertretung Ihres Landes.
www.miele.de/ - 10k - [Cached](#) - [Similar pages](#)

[Herzlich willkommen bei Miele Österreich](#) - [[Translate this page](#)]

Herzlich willkommen bei **Miele** Österreich Wenn Sie nicht automatisch weitergeleitet werden, klicken Sie bitte hier! HAUSHALTSGERÄTE ...
www.miele.at/ - 3k - [Cached](#) - [Similar pages](#)

Search =
miele

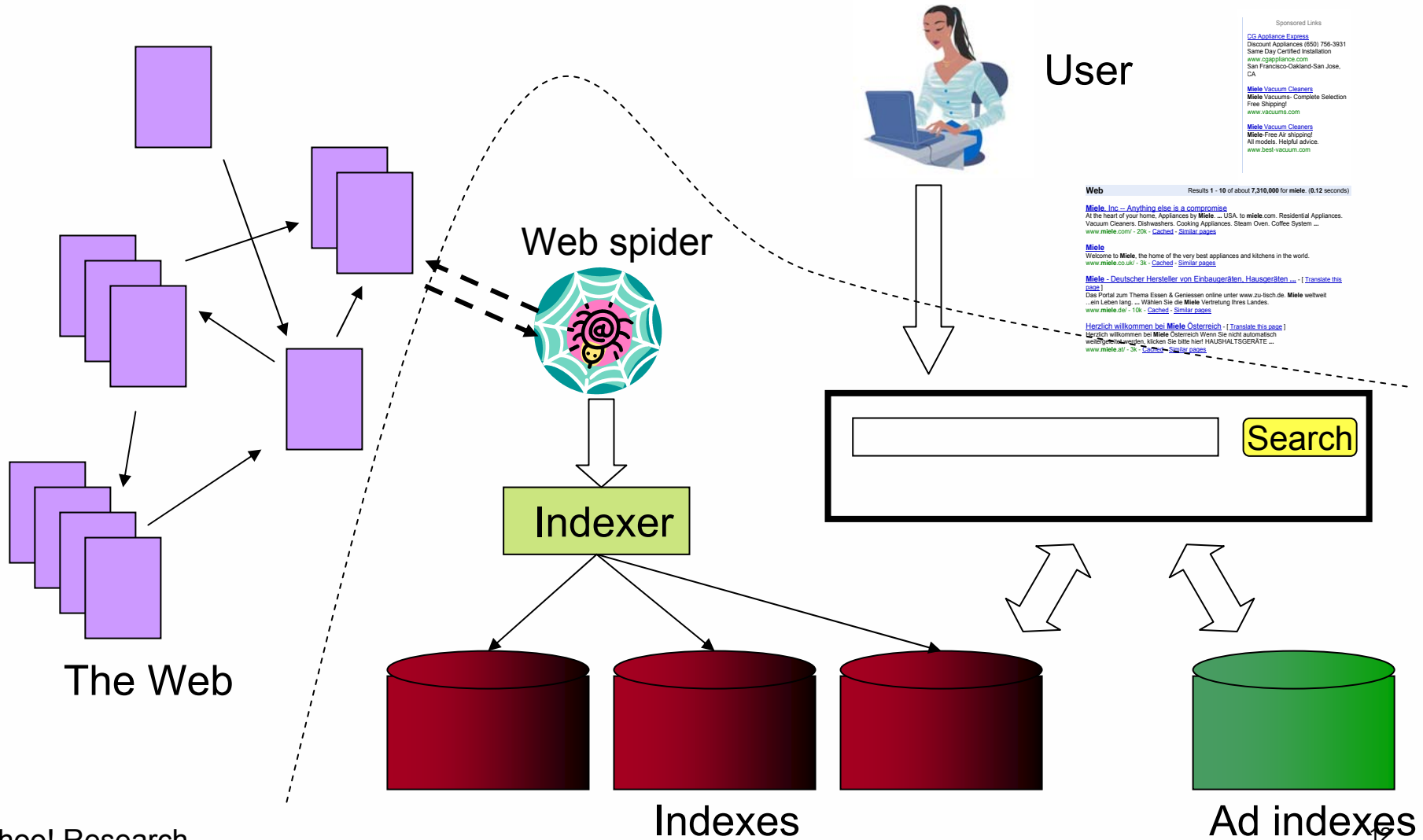


Ads vs. search results

- Other vendors (Yahoo!, MSN) have made similar statements from time to time
 - Any of them can change anytime



Web search basics





“Social” search

Is the Turing test always the
right question?

Photos: [Explore Flickr](#) • [Learn More](#)



Tags / jaguar / clusters

[SEARCH](#)

(Or, try an [advanced search](#).)



[car](#), [cars](#), [auto](#), [etype](#), [automobile](#), [classic](#), [vintage](#), [autoshow](#), [red](#), [show](#)

[→ See more in this cluster...](#)



[zoo](#), [animal](#), [cat](#), [animals](#), [bigcat](#), [seattle](#), [woodlandparkzoo](#), [sleep](#), [edinburgh](#), [caged](#)

[→ See more in this cluster...](#)



[guitar](#), [fender](#)

[→ See more in this cluster...](#)



[aircraft](#), [raf](#)

[→ See more in this cluster...](#)

These are the *most recent* photos tagged with jaguar. [See more...](#)





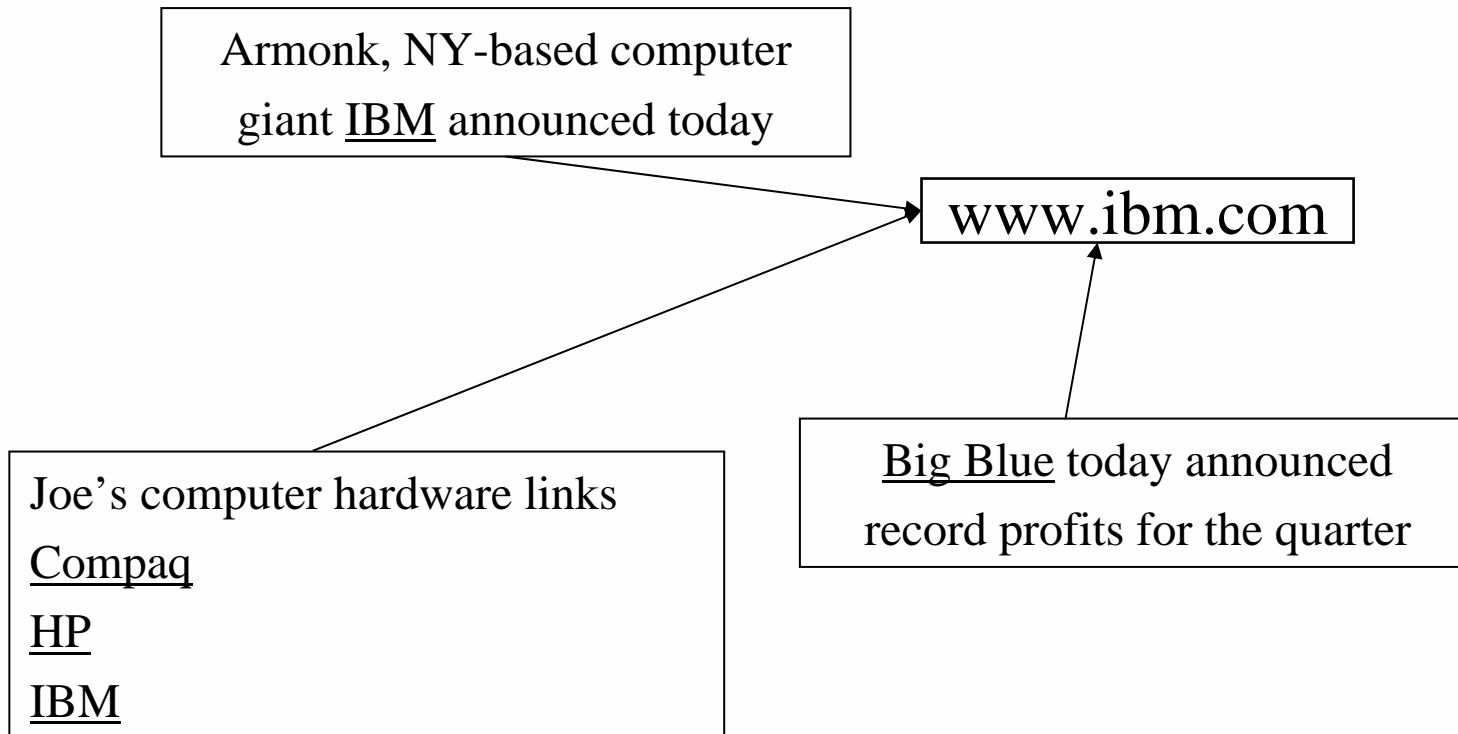
The power of social media

- Flickr – community phenomenon
- Millions of users share and tag each others' photographs (why???)
- The wisdom of the crowd can be used to search
- The principle is not new – anchor text used in “standard” search
- Don't try to pass the Turing test?



Anchor text

- When indexing a document D , include anchor text from links pointing to D .





Challenges in social media

- How do we use these tags for better search?
- What's the ratings and reputation system?
- How do you cope with spam?
- The bigger challenge: where else can you exploit the power of the people?
- What are the incentive mechanisms?



Paid placement

What pays the bills



Paid placement

- Aggregators draw content consumers
 - Search is the “hook”
- Each consumer reveals clues about his current situation
 - The keyword(s) he types (e.g., *miele*)
 - Keyword(s) in his email (gmail)
 - Context information (Yahoo! ...)



Paid placement

- Aggregator gives consumer opportunity to click through to an advertiser
 - Compensated by advertiser for click through
- Whose advertisement is displayed?
 - In the simplest form, auction bids for each keyword
 - Contracts:
 - “At least 20000 presentations of my advertisement to searchers typing the keyword **nfl**, on Super Bowl day”.
 - “At least 100,000 impressions to searchers typing **wilson** in the Yahoo! Tennis category in August”.



Paid placement

- Leads to complex logistical problems: selling contracts, scheduling ads – supply chain optimization
- Interesting issues at the interface of search and paid placement:
 - If you search for *miele*, did you really want the home page of the Miele Corporation at the top?
 - If not, which appliance vendor?



Auctions and pricing

- Overture's original model (still used by Yahoo!):
 - Ads displayed in order of decreasing bid
 - E.g., if advertiser A bids 50 cents, B bids 25, C bids 35 – order ACB
- How do you price slots? Generalized [Vickrey](#)?
- What's good about this?
 - Advertisers like transparency
- What's wrong with this?
 - “Non-performing” ads float to the top
 - Google addressed this with “revenue ordering”:
promote ads based on expected clickability

Web Images Video Audio Directory Local News Shopping More »

YAHOO! SEARCH miele Search the Web My Web (beta)

Search Results Results 1 - 10 of about 8,830,000 for miele - 0.14 sec. (About this page)

to try: miele vacuum, miele appliances, miele dishwasher, More

Miele Vacuums and Accessories
www.thinkvacuums.com Miele vacuum cleaners and accessories.

Pays 8 → **Highest Bidder 10** →

Free Next Day Express On Miele Vacuums
www.vac-cleaner.com Free gift-wrapping. All Miele vacuums models in stock.

Pays 7 → **2nd highest Bidder 8** →

Miele Vacuums and Appliances
www.factorydirectsuperstore.com Full line of vacuums, filters, H built-in appliances.

3rd highest Bidder 7 →

Miele on Yahoo! Shopping
Top Categories: Dishwashers - Cooktops - Washers - Canister Vacuum Cleaners
Yahoo! Shortcut - About

Yahoo! employees: Seeing bad search results for this query? Report them.

Miele USA
Subsidiary providing company, product and dealer information.
Category: Miele > Regional
www.miele.com - Cached - More from this site - Save - Block

Miele - Translate this page
Manufacturing a wide range of home appliances ranging from dishwashers and refrigerators to vacuum cleaners.
Category: Home Appliance Makers > Miele

Miele Vacuum Cleaners - Free Shipping
Extensive reviews, specifications and comparisons of miele vacuums.
www.ristenbatt.com

Over 25 Miele Vacuums at Best-Vacuum.com
Miele authorized vacuum dealer. Free air shipping on Miele vacuums.
www.best-vacuum.com

Shop Miele Appliances
Find Miele home appliances, along with other great brands like...
www.us-appliance.com

Buy Miele Vacuums and Supplies Online
Free insured shipping on all Miele vacuums. Genuine Miele bags,...
totalvac.com

Miele Vacuums and Supplies



Burgeoning research area

- Understanding auctions, incentive mechanisms, pricing, optimization – a huge deal
 - Already multi-billion dollar business, growing fast
 - Interface of microeconomics and CS
 - Early work – Edelman, Ostrovsky, Schwarz
- Advertiser's perspective – spending budget optimally:
Borgs/Chayes/Etesami/Immorlica/Jain/Mahdian
Rusmevichientong/Williamson
- Many open problems, a few papers, some of them quite realistic
 - Marketplace design

WINE



Trademarks and paid placement

- Consider searching Google for **geico**
 - Geico is a large insurance company that offers car insurance

- Sponsored Links

[Car Insurance Quotes](#)

Compare rates and get quotes from top car insurance providers.
[www.dmv.org](#)

[It's Only Me, Dave Pell](#)

I'm taking advantage of a popular case instead of earning my traffic.

[www.davenetics.com](#)

[Fast Car Insurance Quote](#)

21st covers you immediately. Get fast online quote now!

[www.21st.com](#)



Who has the rights to your name?

- Geico sued Google, contending that it owned the trademark “Geico”
 - Thus ads for the keyword **geico** couldn’t be sold to others
 - (Unlikely the writers of the constitution contemplated this issue)
- US Supreme Court ruled: search engines can sell keywords including trademarks



Incentive networks

Joint w/Jon Kleinerg (Cornell)

YAHOO! ANSWERS BETA

Welcome, **praghpragh**
[[Sign Out](#), [My Account](#)]

[Home](#) - [Help](#) - [Forum](#)
[What's going on in Answers?](#)

ask.

Ask a question on any topic and get answers from real people.

(you have **110** characters to work with)

[Post Question](#)

answer.

Share what you know and you might make someone's day.

Featured Question

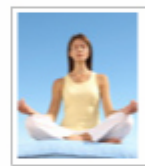


What are the best resources to research African American genealogy?

discover.

Learn something new today by browsing resolved questions.

Featured Topic



See what people are asking about in:

Fitness

[Ask Your Question](#) | [My Q&A](#)

Ready to Participate? - [Get Started!](#)

Our [featured user](#) this week is an aspiring doctor who loves to solve riddles!

Search for questions & answers

 [Search](#)

[Advanced Search](#)

[Save to My Web](#)

Track new questions on Answers

[+ MY YAHOO!](#) [RSS](#)

Share what you know. Answer open questions.



[what is the structure and deployment of the armed forces under the federal government?](#)

Asked by [arne1](#) - [Other - Politics & Government](#) - 3 seconds ago



[What is the outlook for Arena Resources Ticker ARD???](#)

Asked by [Sapito](#) - [Investing](#) - 10 seconds ago



[My dear Sonia Babbar \(Raveena\) where are you. Still waiting for you. always urs. Hirendra?](#)

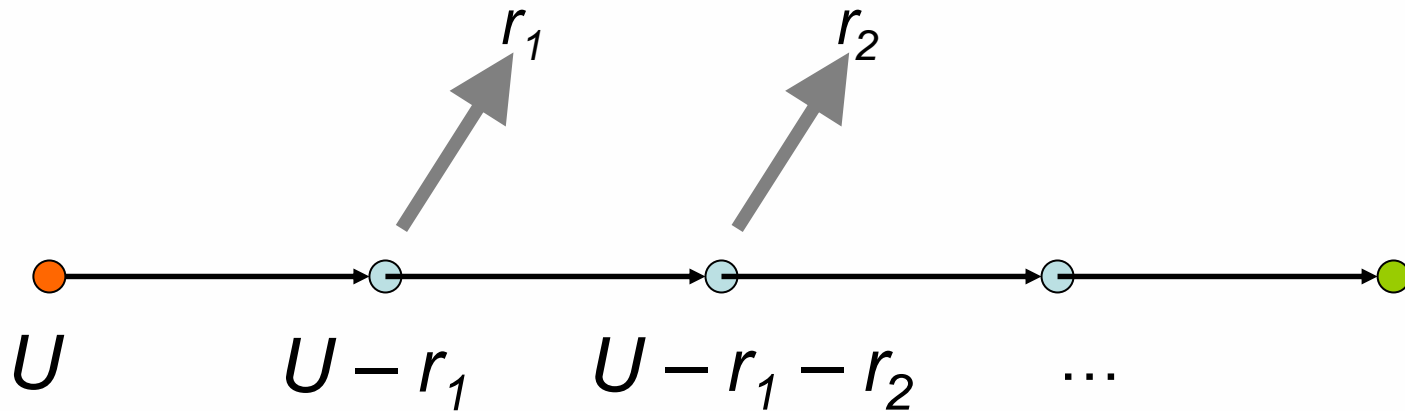


The power of the middleman

- Setting: you have a need
 - For information, for goods ...
- You initiate a request for it and offer a reward for it, to some person X
 - Reward = your utility U
- How much should X “skim off” from your offered reward, before propagating the request?



Propagation



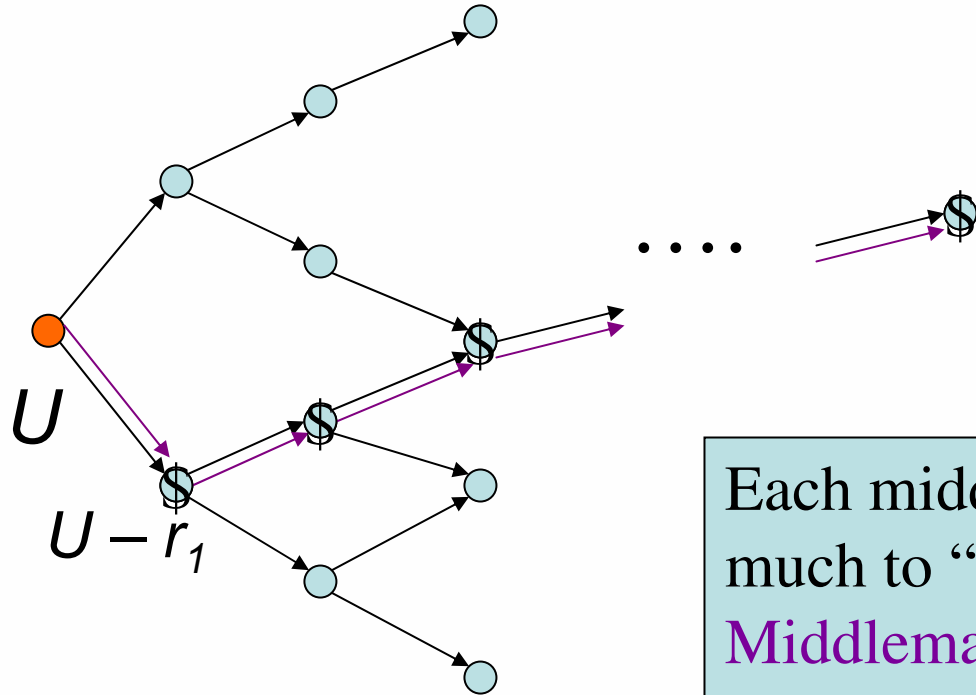
Request propagated repeatedly until it finds an answer.

Target not known in advance.

Middlemen get reward only if answer reached.



More generally



Each middleman decides how much to “skim off”.
Middleman only gets paid if on the path to the answer.

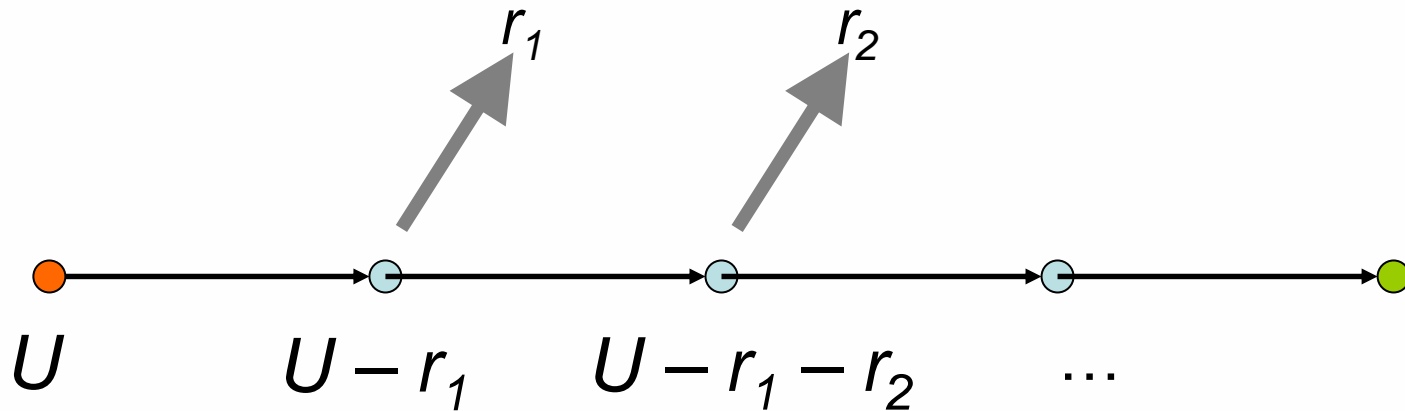


Rewards must be non-trivial

- We will assume that all the $r_i \geq 1$.
- Else, have a form of Zeno's paradox:
 - Source can get away with offering an arbitrarily small reward.
- Equivalently, nodes value their effort in participating.



Back to the line



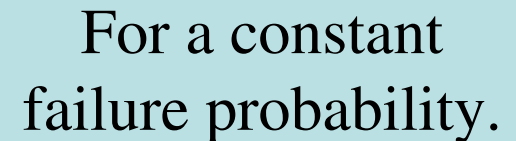
Under *strategic* behavior by each player, how much should a player skim?

$n =$ answer *rarity*: probability a node has the answer = $1/n$, independently of other nodes.



The bad news

- For rarity n , it takes about n hops to get to the answer.
- (By solving a simple recurrence): The initial reward is exponential in n
 - A very inefficient network.
- Was this because of the model, the network, or ... ?



For a constant failure probability.

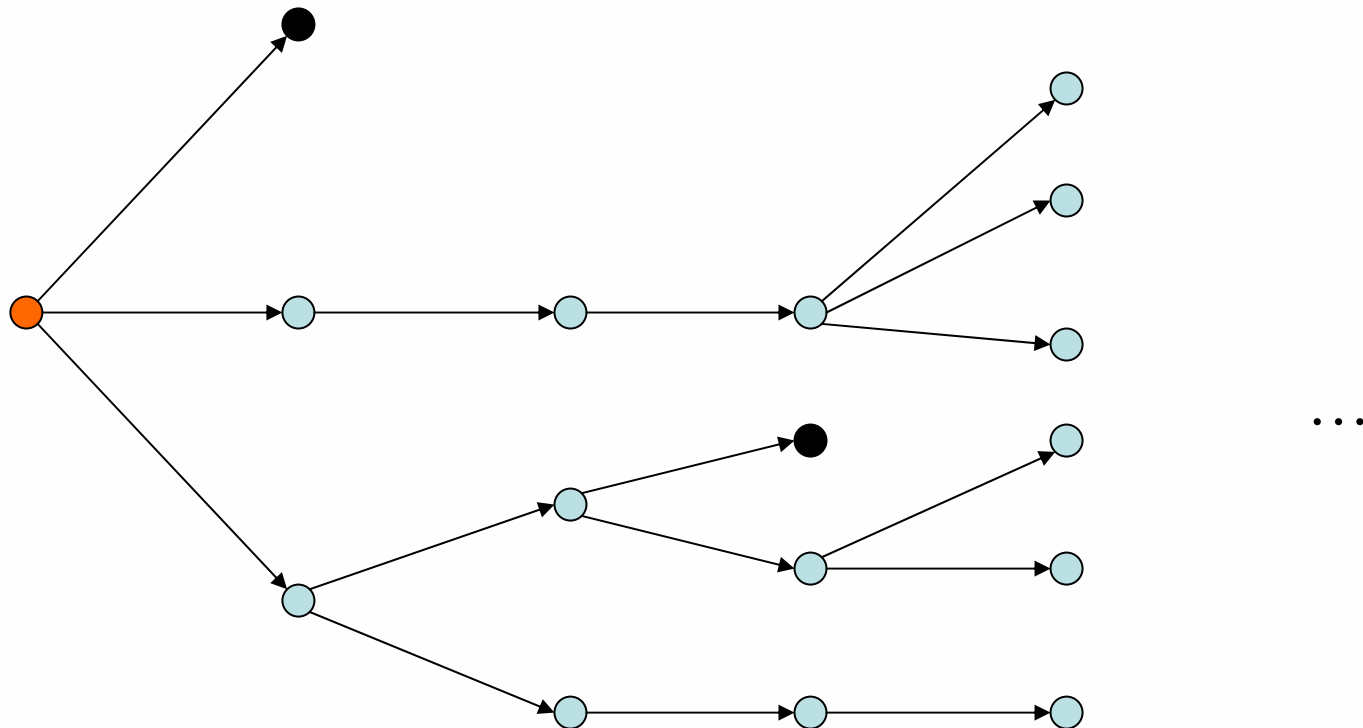


Branching processes

- Branching process: a network where
- Each node has a number of descendants
- Number of descendants is a random variable X
 - drawn from a probability distribution
 - Expectation[X] = b



Branching process example





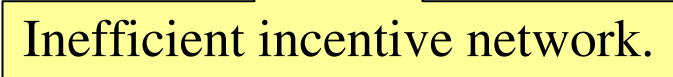

Branching processes

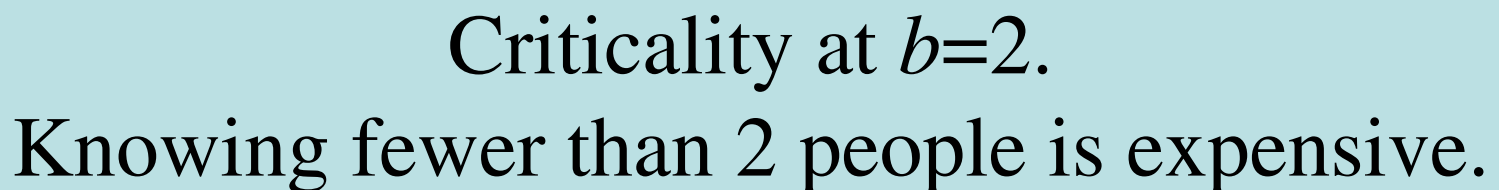
- Classical study of population dynamics and random graph evolution
- Basic fact:
 - If the expected # offspring < 1 , process dies out
 - If the expected # offspring ≥ 1 , process infinite.

$b=1$ called a *critical* process.



Main results

- For $b < 2$, the initial investment must be exponential in the path length from the root to the answer.  Inefficient incentive network.
- For $b > 2$, the initial investment is linear in the path length from the root to the answer.  Efficient incentive network.



Criticality at $b=2$.
Knowing fewer than 2 people is expensive.



Tempting conclusion

- (Sufficient) competition makes incentive networks efficient.
- But ... we haven't fully introduced competition yet.
 - On trees, we have a unique path from origin to each node.

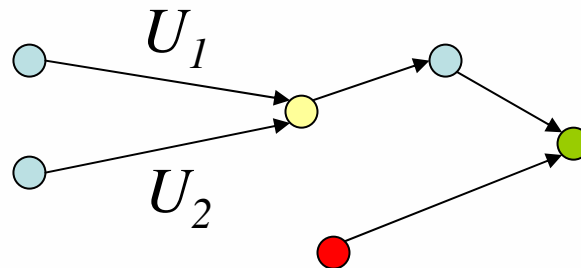
A light blue callout box with a black border and a white arrow pointing upwards from the text to the 'But ...' part of the list above.

Limited form of competition.



Real competition

- Directed acyclic graph with an origin.
- **Rarity n as before.**
- Now a node may receive different reward offers from different parents.
- **Node must consider that descendants can get competing offers.**





Many open questions

- Branching processes:
- What happens at $b=2$?
 - When there are exactly 2 offspring, can show cost linear in path length
- What is the phase transition in the general case?



Directed acyclic graphs

- Full model of competition
- When does competition promote efficiency?
- Given a DAG, how does a node compute its strategy?



The net

- Web search is scientifically young
- It is intellectually diverse
 - The human element
 - The social element
- The technology mirrors the economic, legal and sociological reality

Thank you.



Questions?

pragh@yahoo-inc.com