Microsoft® Research

Faculty Summit

10

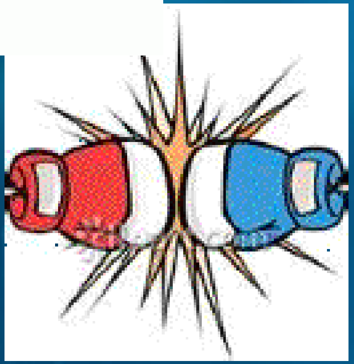
YEAR ANNIVERSARY

Some Vignettes from Learning Theory

Robert Kleinberg
Cornell University

Microsoft Faculty Summit, 2009

Prelude: Tennis or Boxing?



- You’re designing a sporting event with *n* players of unknown quality
- Spectators want to see matches between the highest-quality players
  - No preference for variety or for seeing upsets
- Tennis solution: single-elimination tournament
- Boxing solution: players challenge the current champion until he/she is defeated
- Which is optimal? Or is a third alternative better?

Online Learning

Microsoft Research

Algorithms that make decisions with uncertain consequences, guided by past experience

```
rdk:~$ dig microsoft.com

; <<>> DiG 9.3.6-P1 <<>> microsoft.com
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 65039
;; flags: qr rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 5, ADDITIONAL: 5

;; QUESTION SECTION:
;microsoft.com.                 IN      A

;; ANSWER SECTION:
microsoft.com.          3583    IN      A       207.46.232.182
microsoft.com.          3583    IN      A       207.46.197.32

;; AUTHORITY SECTION:
microsoft.com.          3153    IN      NS      ns2.msft.net.
microsoft.com.          3153    IN      NS      ns5.msft.net.
microsoft.com.          3153    IN      NS      ns1.msft.net.
microsoft.com.          3153    IN      NS      ns3.msft.net.
microsoft.com.          3153    IN      NS      ns4.msft.net.

;; ADDITIONAL SECTION:
ns1.msft.net.           140615  IN      A       65.55.37.62
ns2.msft.net.           140615  IN      A       64.4.59.173
ns3.msft.net.           140615  IN      A       213.199.161.77
ns4.msft.net.           140615  IN      A       207.46.66.126
ns5.msft.net.           140615  IN      A       65.55.226.140

;; Query time: 5 msec
;; SERVER: 171.64.7.99#53(171.64.7.99)
;; WHEN: Fri Jul 10 12:49:02 2009
;; MSG SIZE  rcvd: 241
```

Multi-armed Bandits

- Decision maker picks one of *k* actions (slot machines) in each step, observes random payoff
- Try to minimize “regret”
  - Opportunity cost of not knowing the best action a priori



0.3	0.7	0.4
0.5	0.1	0.6
0.2	0.2	0.7
0.3	0.8	0.5
0.6	0.1	0.4
2.2	vs.	2.6

Multi-armed Bandits

Microsoft Research

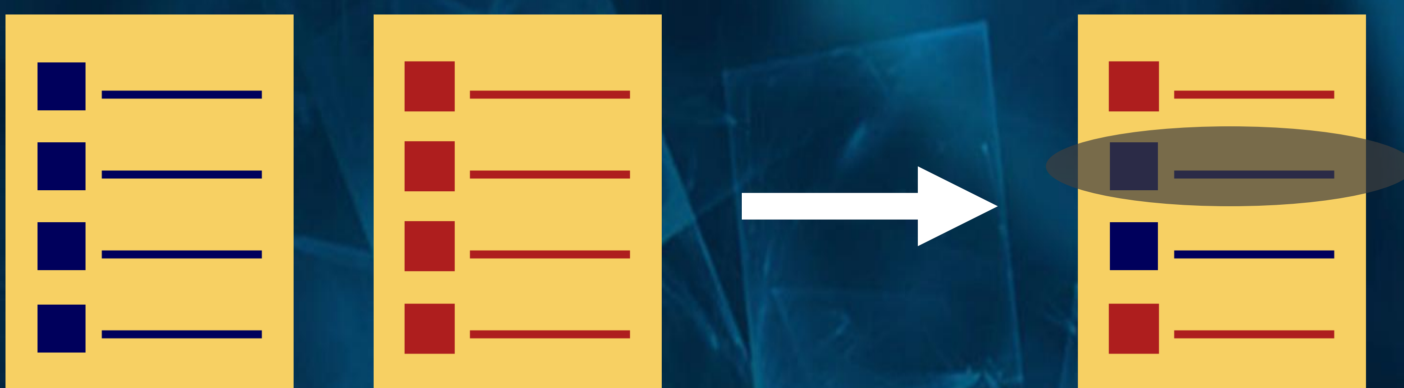
Studied for more than 50 years, but the theory is experiencing a renaissance influenced by the Web

0.3	0.7	0.4
0.5	0.1	0.6
0.2	0.2	0.7
0.3	0.8	0.5
0.6	0.1	0.4
2.2	vs.	2.6

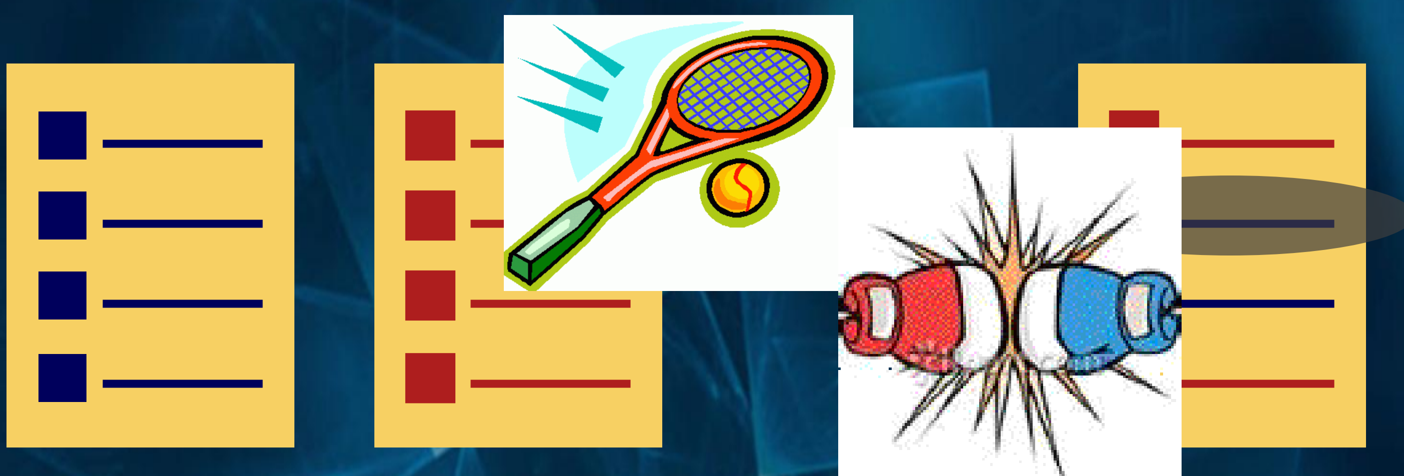
Example: Learning to Rank

- You have many different ranking functions for constructing a list of search results
- Interactively learn which is best for a user or population of users
- Elicit quality judgments using “interleaving experiments.” (Radlinski, Korup, Joachims, CIKM’08)



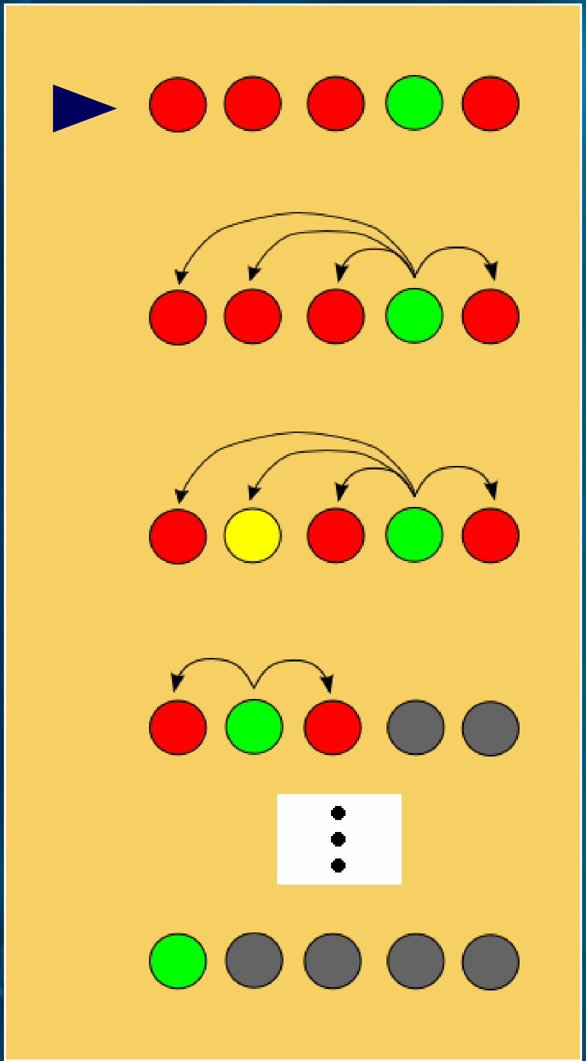
Example: Learning to Rank

- Much more reliable than other ways of detecting retrieval quality from "implicit feedback"
  - E.g. abandonment rate, query reformulation rate, position of the clicked links
- This is like multi-armed bandits, but with a twist: you can compare two slot machines, but you can't just pick one and observe its payoff



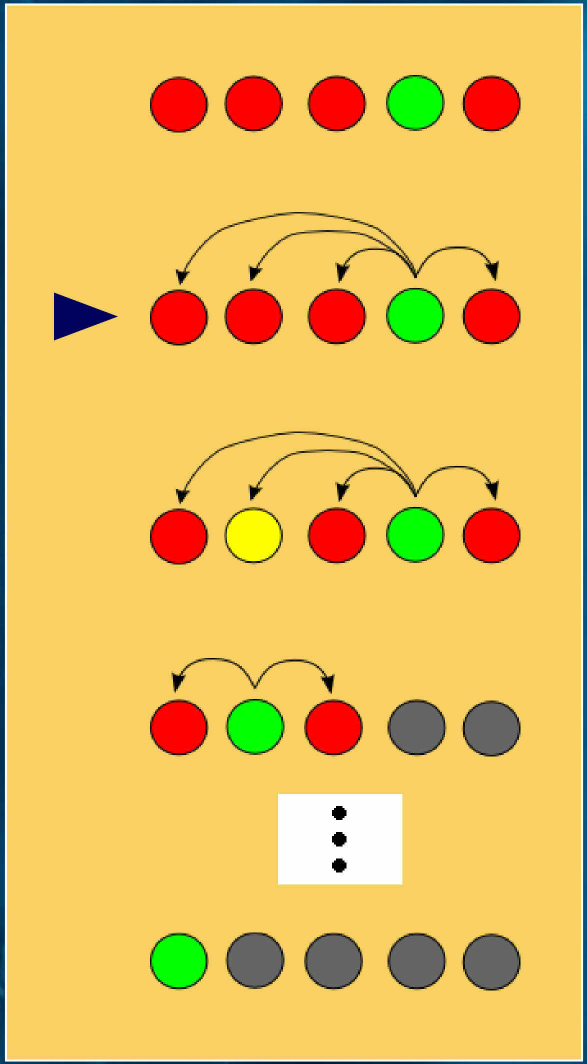
Interleaved Filter

- Choose arbitrary “incumbent”



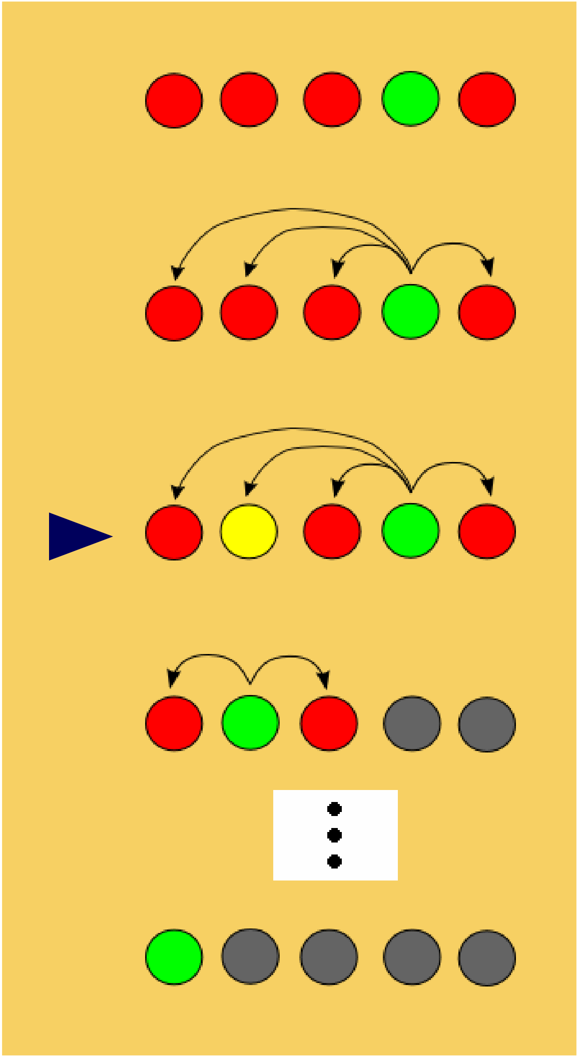
Interleaved Filter

- Choose arbitrary "incumbent"
- Play matches against all other players in round-robin fashion…
  (noting mean, confidence interval)



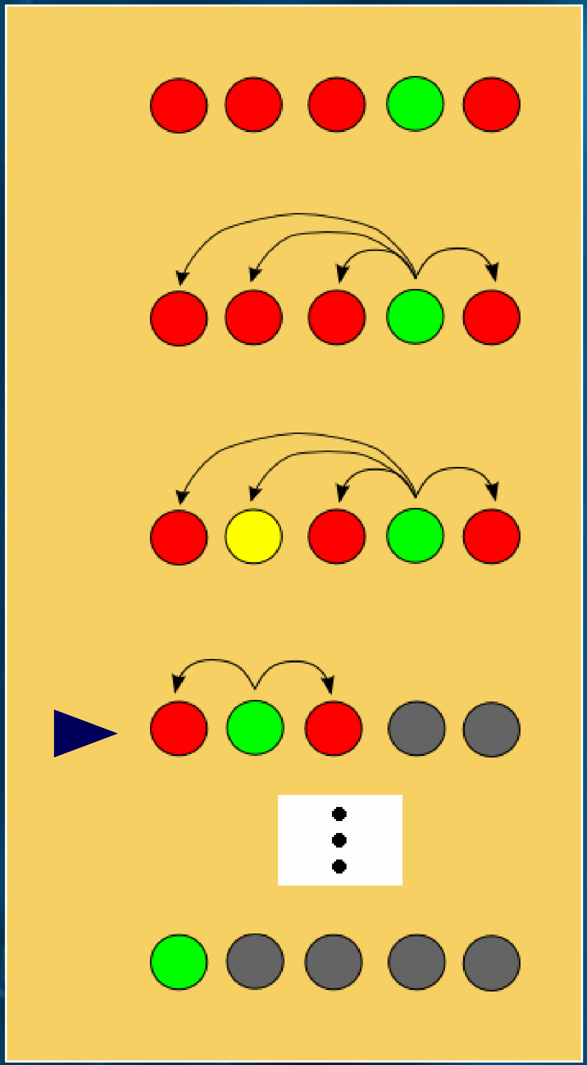
Interleaved Filter

- Choose arbitrary “incumbent”
- Play matches against all other players in round-robin fashion…
  (noting mean, confidence interval)
- … until a challenger is better with high confidence



Interleaved Filter

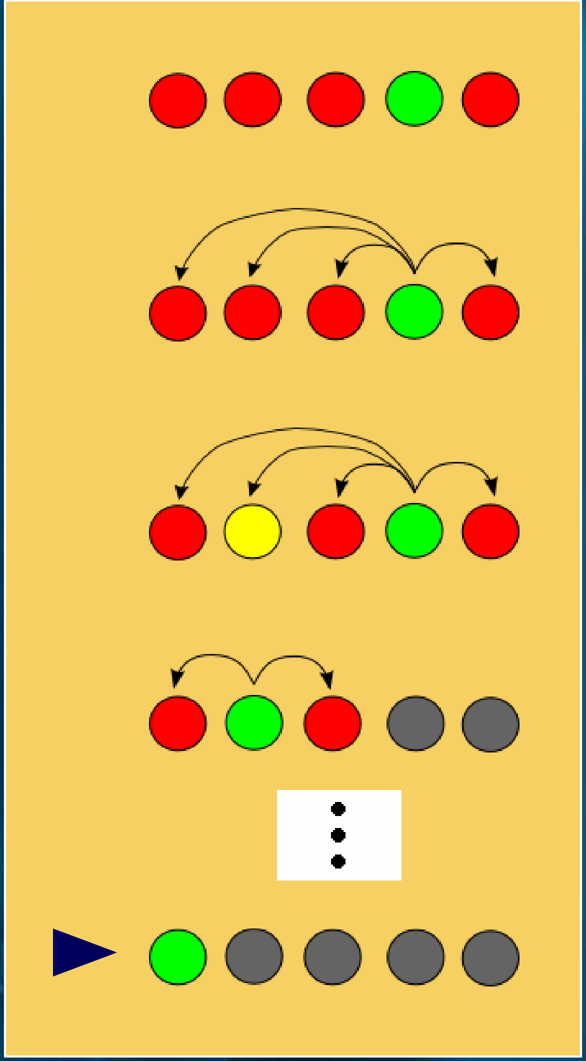
- Choose arbitrary "incumbent"
- Play matches against all other players in round-robin fashion…
  (noting mean, confidence interval)
- … until a challenger is better with high confidence
- Eliminate old incumbent and all empirically worse players
- Repeat process with new incumbent…



Interleaved Filter

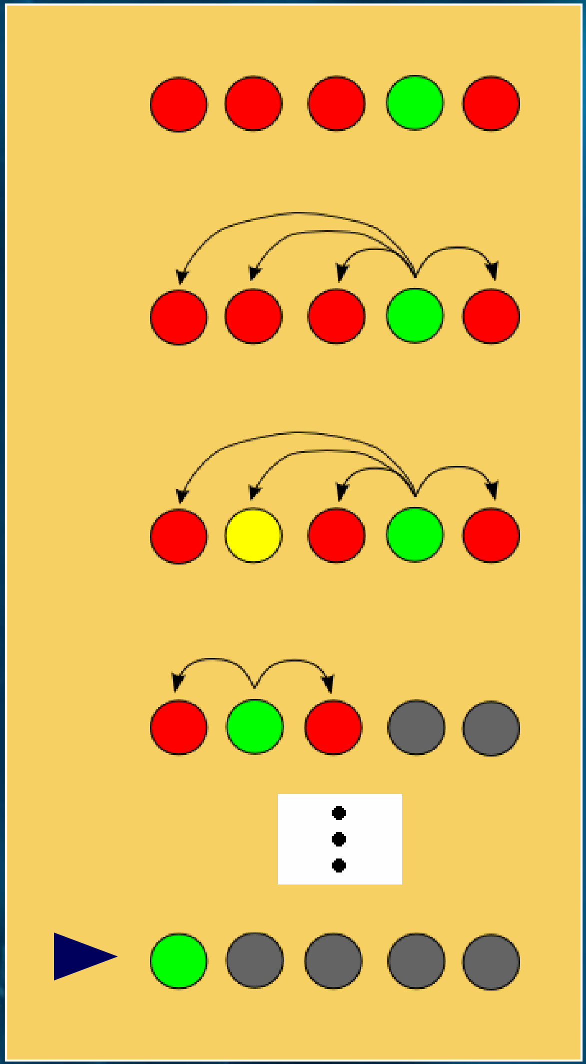
Microsoft Research

- Choose arbitrary "incumbent"
- Play matches against all other players in round-robin fashion…
  (noting mean, confidence interval)
- … until a challenger is better with high confidence
- Eliminate old incumbent and all empirically worse players
- Repeat process with new incumbent…
- … until only one player is left



Interleaved Filter

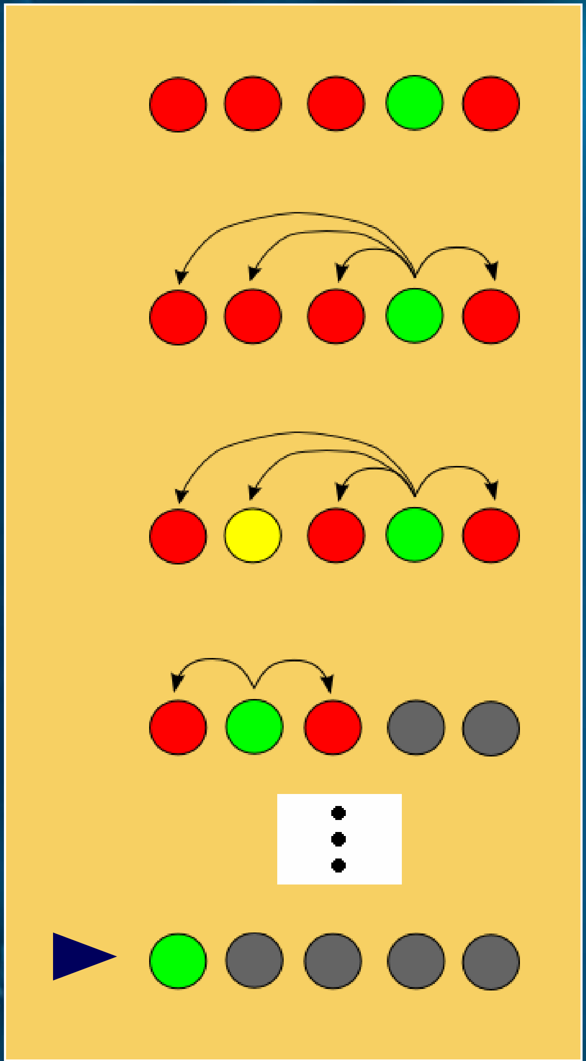
- This algorithm is information theoretically optimal
- Boxing is better than tennis!



Interleaved Filter

- This algorithm is information theoretically optimal
- Boxing is better than tennis!
- **Thank you, Microsoft!**

  Yisong Yue, the lead student on the project, is supported by a Microsoft Graduate Research Fellowship



Vignette #2: Learning with Similarity Information

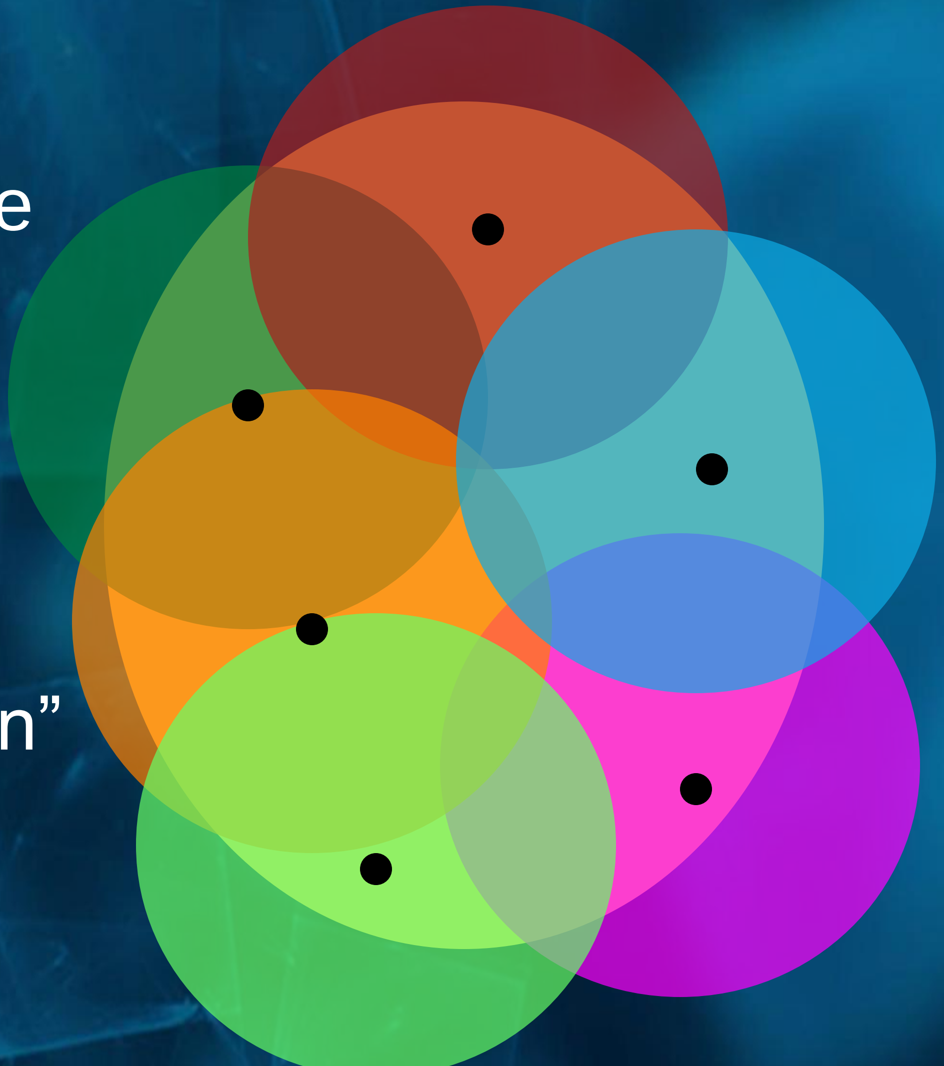
- Recall the multi-armed bandit problem
- Can we use this for web advertising?
- Slot machines are banner ads, which one should I display on my site?

Vignette #2: Learning with Similarity Information

- Recall the multi-armed bandit problem
- Can we use this for web advertising?
- Slot machines are banner ads, which one should I display on my site?
- Scalability issue: there are 10^5 bandits, not 3!
- On the other hand, some ads are similar to others, and this should help

Solution:
The Zooming Algorithm

- The set of alternatives (ads) are a metric space
- We designed a bandit algorithm for metric spaces, that starts out exploring a “coarse” action set and “zooms in” on regions that are performing well



Solution: The Zooming Algorithm

Microsoft Research

- The set of alternatives (ads) are a metric space
- We designed a bandit algorithm for metric spaces, that starts out exploring a "coarse" action set and "zooms in" on regions that are performing well

Thank you, Microsoft!!

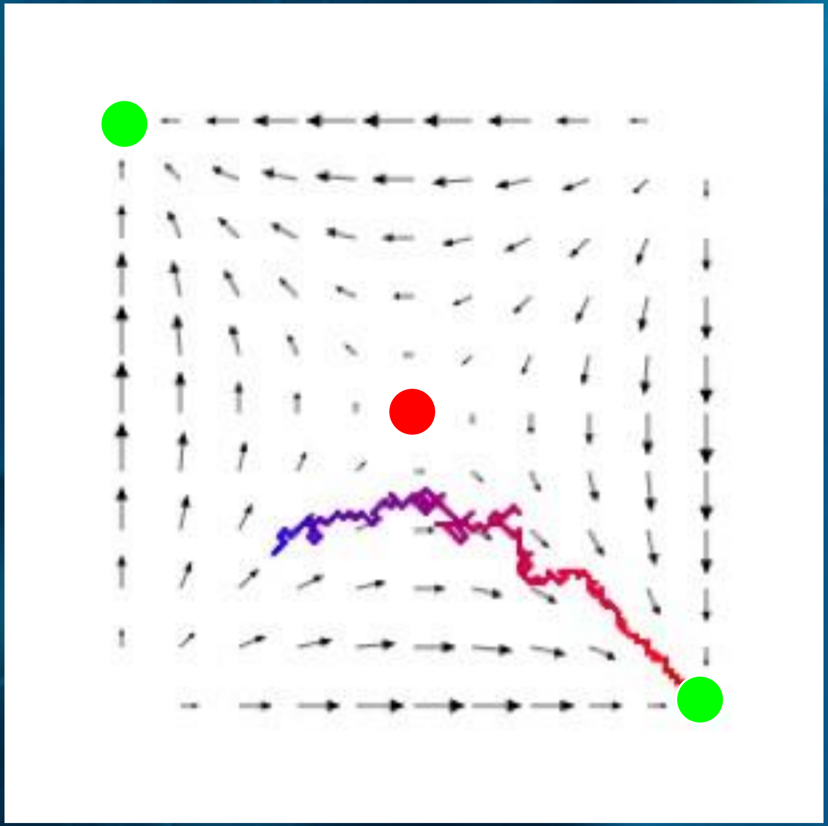
Alex Slivkins

- One of many collaborations with MSR over six years … a major influence on my development as a computer scientist

What Next?

Microsoft Research

- Often, the systems we want to analyze are composed of many interacting learners
- How does this influence the system behavior?
- Answering these questions requires combining:
  - Game theory
  - Learning theory
  - Analysis of algorithms



Thank you, Microsoft!!!

Microsoft Research

- Joining our team next year…
  - Katrina Ligett (Ph.D. CMU, 2009)
  - Shahar Dobzinski (Ph.D. Hebrew U., 2009)
- …the top graduates this year in online learning theory and algorithmic game theory
- An unprecedented postdoc recruiting success for myself and Cornell
- Brought to you by the Microsoft Research New Faculty Fellowship!