Brokering Actions on the Web

Patrick Pantel

Joint work with:
Tom Lin (UW), Michael Gamon (MSR),
Anitha Kannan (MSR), Ariel Fuxman (MSR)

April 2012
Structured Recommendations

Related Movies:
- Dreams
- Ran
- Dersu Uzala
- Throne of Blood
- Yojimbo
- Stray Dog

Entity Detection

Structured Data

Search Results:
- Seven Samurai

Also try: seven samurai movie, seven samurai remake, more...

Seven Samurai (1954)
movies.yahoo.com

Yahoo's B+: A Japanese farming village, constantly besieged and pillaged by an army of bandits, recruits seven independent... more

Running Time: 2 hr 40 min
Directed by: Akira Kurosawa
Starring: Takashi Shimura, Toshiro Mifune, Yoshio Inaba, Seiji Miyaguchi, ... more

Play trailer
Task completion

Price prediction

Aggregate ratings

Task completion
Faceted search
Did you know...

A HUGE number of Internet users simply search for entities. They come to Bing to find where to manipulate them (function overload).

**BIG opportunity to act as a broker.**
Current Experience

JETBeam RRT-0 Review
JETBeam Flashlights
JETBeam RRT-0 R5
JETBeam Raptor
JETBeam M1X
JETBeam RRT-0 R5
240 Lumens

JETBeam RRT-0
Going Gear JETBeam RRT-0 R5 with Infinite Ramping LED Flashlight [jetbeamrtr0r5] - The popular RRT-0 has been updated to be an even more versatile flashlight,...
goinggear.com/index.php?main_page=product_info&cPath=14_22&products_id=438 · Mark as spam

Jetbeam RRT-0
Now In Stock! Model: RRT-0 (S2) Specially designed for Military, Law Enforcement, Self-defense, Hunting, Search & Rescue and General Outdoors The Rapid...
www.jetbeamusa.com/products/RRT%2d0.html · Mark as spam

JETBeam RRT-0 Overview With Lots of Pictures
I got my JETBeam RRT-0 Raptor today. All I can say is to JETBeam! This thing is a ... PICTURES QUICKLY.... BEFORE MY HEAD EXPLODES!!! Ok... it didn't ...
323 replies since November 2009
www.candlepowerforums.com/vb/showthread.php?248375-JETBeam-RRT-0-Overview-With-Lots-of... · Mark as spam

JETBeam Flashlights
Specification: Model: JETBeam RRT-0(R2) Flashlight; Specially designed for Military, Law Enforcement, Self-defense, Hunting, Search & Rescue and Outdoorsman
jetbeamblog.com · Mark as spam
Better experience

Actions easily accessible

Recognize entity in query

Active Objects
Better experience

Recognize entity in query

Actions easily accessible

---

**Jetbeam RRT-0**
Now In Stock! Model: RRT-0 (S2) Specially designed for Military, Law Enforcement, Self-defense, Hunting, Search & Rescue and General Outdoors The Rapid ...

Get more ... www.candlepowerforums.com/vb/showthread.php?248375-JETBeam-RRT-0-Overview-With-Lots-of-... Mark as spam

**JETBeam Flashlights**
Specification: Model: JETBeam RRT-0(R2) Flashlight; Specially designed for Military, Law Enforcement, Self-defense, Hunting, Search & Rescue and Outdoorsman jetbeamblog.com · Mark as spam
Better experience

Recognize entity in query

Actions easily accessible

Web Images Videos Shopping News Maps More | MSN Hotmail

**ENTITY**

**ALL RESULTS** 1-10 of 6,600 results · Advanced

**JETBeam RRT-0** R5 with Infinite Ramping LED Flashlight... Going Gear JETBeam RRT-0 R5 with Infinite Ramping LED Flashlight [jetbeammrtr05] - The popular RRT-0 has been updated to be an even more versatile flashlight. ... goinggear.com/index.php?main_page=product_info&cPath=14_22&products_id=438 · Mark as spam

**Jetbeam RRT-0**
Now In Stock! Model: RRT-0 (S2) Specially designed for Military, Law Enforcement, Self-defense, Hunting, Search & Rescue and General Outdoors The Rapid... www.jetbeamusa.com/products/RRT%2d0.html · Mark as spam

**JETBeam RRT-0** Overview With Lots of Pictures
I got my JETBeam RRT-0 Raptor today. All I can say is to JETBeam! This thing is a... PICTURES QUICKLY.... BEFORE MY HEAD EXPLODES!!! Ok... it didn’t...
Films

Recognize entity in query

Actions easily accessible

Web Images Videos Shopping News Maps More | MSN Hotmail

bing

Pandorum

Web Images Videos Shopping Movies More

ENTITY

ALL RESULTS

1-10 of 116,000 results · Advanced

PANDORUM - Now Available on DVD & Blu-ray
PANDORUM Movie Official Site - Now Available on DVD & Blu-ray. Get Trailers and Clips, Synopsis, Games, Downloads, Gallery and more.
www.pandorummovie.com · Mark as spam

Pandorum - Wikipedia, the free encyclopedia
Plot · Cast · Production · Release
Pandorum is a 2009 German-British science fiction thriller film written by Travis Milloy, directed by Christian Alvart and produced by Paul W.S. Anderson.
en.wikipedia.org/wiki/Pandorum · Mark as spam

Pandorum (2009) - IMDb
Action/Horror/Sci-Fi · R · 108 min
With Dennis Quaid, Ben Foster, Cam Gigandet, Antje Traue. A pair of crew members aboard a spaceship wake up with no knowledge of...
www.imdb.com/title/tt11148729 · Mark as spam

Actions
Watch Trailer
Watch Film Online
Read User Reviews
Read Critics Reviews
Buy DVD/Blu-ray
Jinni Taste Match

Active Objects
Politicians

Actions easily accessible

Recognize entity in query
1) INTRODUCTION

2) ACTIONS IN WEB SEARCH

- plan vacation
- get in shape
- hilton orlando reviews
- sea world location
- how to lose weight

3) ACTIVE OBJECTS

4) MINING ENTITY TYPES

- Informational
- Navigational
- Transactional
  - [Broder, 2002]
  - [Rose and Levinson, 2004]

- Advice
- Locate
- ... Download
- Obtain
- Interact
- ...

- get address (landmark)
- add to Netflix queue (film)

- read reviews (hotel)
- buy (camera)
**Actions vs Intents**

**User Intents and Goals**
- plan vacation
- get in shape

**Query**
- hilton orlando reviews
- sea world location
- how to lose weight

**Query Intent**
- Informational
  - Advice
  - Locate
- Navigational
  - ... (not specific example)
- Transactional
  - Download
  - Obtain
  - Interact

**Finer-grained Intents**
- get address(landmark)
- add to Netflix queue(film)

**Actions on Entities**
- read reviews(hotel)
- buy(camera)

[Broder, 2002]
[Rose and Levinson, 2004]
Do web queries contain entities?

Entity Distribution in Web Search Queries

- **43% entity**
  - (e.g., “GoldenEye”, “Horne Auto”)

- **14% entity category**
  - (e.g., “golf cart battery”, “global sim card”)

- **15% no entity**
  - (e.g., “xxx”, “good reading quotes”)

- **28% website**
  - (e.g., “yahoo mail”, “girlybox.com”)

- **4% category**

- **10% category + refiner**

- **14% entity + refiner**

**Schema.org types for entity-bearing queries**

- **creativework** 40%
- **organization** 37%
- **person** 8%
- **product** 9%
- **event** 3%
- **other** 3%

*From a query traffic-weighted sample*
## Ontology of Actions

<table>
<thead>
<tr>
<th>Navigational (need satisfied by reading content, or could be satisfied by written transcript of content)</th>
<th>Transactional (navigating to a web-mediated action)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Login Action (on a Website entity)</td>
<td>Apply for Job (on a LocalBusiness / Organization entity)</td>
</tr>
<tr>
<td>Search Action (on a Website entity)</td>
<td>Buy (Shopping on a Product entity)</td>
</tr>
<tr>
<td>Informational (need satisfied by reading content, or could be satisfied by written transcript of content)</td>
<td>Buy Tickets (on an Event / Product / Person entity)</td>
</tr>
<tr>
<td>Find Location(s) (on an Organization entity)</td>
<td>Content Creation (on a Website entity)</td>
</tr>
<tr>
<td>Find Lyrics (on a CreativeWork / MusicalTrack entity)</td>
<td>Discuss Online (on any entity)</td>
</tr>
<tr>
<td>Find Recipe For (on a food)</td>
<td>Download (on a CreativeWork or Software entity)</td>
</tr>
<tr>
<td>Find Where to Buy (on a Product entity)</td>
<td>Listen to Music (on a CreativeWork or Website entity)</td>
</tr>
<tr>
<td>Get Contact Information (on an Organization entity)</td>
<td>Manage Account (on a Local Business / Website / Org entity)</td>
</tr>
<tr>
<td>Get Directions To (on an Organization / Location entity)</td>
<td>Pay Bill (on a Website / Organization entity)</td>
</tr>
<tr>
<td>Get Domain Information (on a Website entity)</td>
<td>Play Game (on a Game entity)</td>
</tr>
<tr>
<td>Get Event Details (on an Event entity)</td>
<td>Rent (on a CreativeWork / Product entity)</td>
</tr>
<tr>
<td>Get Event Results (on an Event entity)</td>
<td>Reservation (on a Hotel entity)</td>
</tr>
<tr>
<td>Product Detail (on a Product entity)</td>
<td>Schedule Appointment (on a LocalBusiness entity)</td>
</tr>
<tr>
<td>Learn (on any entity)</td>
<td>Sell (Shopping on a Product entity)</td>
</tr>
<tr>
<td>Learn / Educational (on a Person / Product / Organization entity)</td>
<td>Use Service On (e.g., translate, on a Website)</td>
</tr>
<tr>
<td>Learn / Trivia (on any entity)</td>
<td>Watch Video About (on any entity)</td>
</tr>
<tr>
<td>Operating Hours (on an Organization entity)</td>
<td>Web Chat</td>
</tr>
<tr>
<td>Read Articles (on a News / Magazine entity)</td>
<td>Other</td>
</tr>
<tr>
<td>Read Guide (on a Product entity)</td>
<td>Shopping (category of actions including reviews and buying)</td>
</tr>
<tr>
<td>Read Help (on a Product entity)</td>
<td>Various/Unknown</td>
</tr>
<tr>
<td>Read News About (on any entity)</td>
<td></td>
</tr>
<tr>
<td>Read News About (on any entity)</td>
<td></td>
</tr>
<tr>
<td>Read Reviews (Shopping on a CreativeWork / Product / Service entity)</td>
<td></td>
</tr>
<tr>
<td>Read Spoilers (on a CreativeWork)</td>
<td></td>
</tr>
<tr>
<td>Research (focused information gathering, on any entity)</td>
<td></td>
</tr>
<tr>
<td>Search Database of (e.g., obituaries, on an Organization / Website)</td>
<td></td>
</tr>
<tr>
<td>See Menu (on a Restaurant)</td>
<td></td>
</tr>
<tr>
<td>See Pictures (on a Person / Product / Organization entity)</td>
<td></td>
</tr>
<tr>
<td>Side Effects / Safety (on a Product entity)</td>
<td></td>
</tr>
<tr>
<td>Stock Price (on an Organization entity)</td>
<td></td>
</tr>
</tbody>
</table>

Actions are tied to entity types
47 actions in current list

Note: No existing Actions equivalent for Schema.org
How many Actions should there be?

Discovery Rate of New Actions

Rapidly decreasing discovery rate
1) INTRODUCTION

2) ACTIONS IN WEB SEARCH

- plan vacation
- get in shape

- hilton orlando reviews
- sea world location
- how to lose weight

- Informational
- Navigational
- Transactional

[Broder, 2002]

- Advice
- Locate
- Download
- Obtain
- Interact

[Rose and Levinson, 2004]

- get address(landmark)
- add to Netflix queue(film)

- read reviews(hotel)
- buy(camera)

3) ACTIVE OBJECTS

4) MINING ENTITY TYPES
Learning Actions from Web Usage Logs

- Three months of us-en web logs
- Annotate with Freebase entities
- Keep queries with an entity in set of 21 types
- Filter out navigational queries
- Filter out clicked hosts that weren’t clicked at least 100 times

<table>
<thead>
<tr>
<th>(query, host) pairs over 3 months</th>
<th>2,164,579</th>
</tr>
</thead>
<tbody>
<tr>
<td>entities</td>
<td>235,385</td>
</tr>
<tr>
<td>contexts</td>
<td>129,088</td>
</tr>
<tr>
<td>hosts</td>
<td>58,123</td>
</tr>
</tbody>
</table>
**Goal**: Define a theory (story) for how actionable queries are generated.

The story for \( p(\text{actionable query}) \),

or more formally

The story for \( p(\phi, \theta, a, n | \alpha, \beta) \)

For each action \( a \)
\[
\phi_a \sim \text{Dirichlet}(\beta) \quad \text{(action→contexts die)}
\]

For each query \( q \)
\[
\theta \sim \text{Dirichlet}(\alpha)
\]

For each context position in \( q \) (pre or post)
\[
\text{action } a \sim \text{Multinomial}(\theta) \\
\text{ngram } n \sim \text{Multinomial}(\phi_a)
\]
The story for $p(\phi, \theta, a, n \mid \alpha, \beta)$

For each action $a$

$\phi_a \sim \text{Dirichlet}(\beta)$

For each query $q$

$\theta \sim \text{Dirichlet}(\alpha)$

For each context position in $q$ (pre or post)

- action $a \sim \text{Multinomial}(\theta)$
- ngram $n \sim \text{Multinomial}(\phi_a)$

action $a \sim \text{Multinomial}(\theta)$

ngram $n_1 \sim \text{Multinomial}(\phi_a)$

ngram $n_2 \sim \text{Multinomial}(\phi_a)$

"ebert Star Wars review"

read reviews action
Clicked hosts matter...

*buy action*
- amazon.com
- ebay.com
- walmart.com

*read reviews action*
- rottentomatoes.com
- metacritic.com
- efilmcritic.com
The story for $P(\phi, \theta, \omega, a, n, c \mid \alpha, \beta, \iota)$

For each action $a$
$\phi_a \sim \text{Dirichlet}(\beta)$ (to contexts)
$\omega_a \sim \text{Dirichlet}(i)$ (to clicks)

For each query $q$
$\theta \sim \text{Dirichlet}(\alpha)$
action $a \sim \text{Multinomial}(\theta)$
ngram $n_1 \sim \text{Multinomial}(\phi_a)$
ngram $n_2 \sim \text{Multinomial}(\phi_a)$
click $c \sim \text{Multinomial}(\omega_a)$
The type matters...
The story for $P(\phi, \theta, \tau, \omega, t, a, n, c \mid \alpha, \beta, \gamma, \iota)$

For each action $a$
- $\phi_a \sim \text{Dirichlet}(\beta)$
- $w_a \sim \text{Dirichlet}(i)$

For each type $t$
- $\tau_t \sim \text{Dirichlet}(\gamma)$

For each query $q$
- $\theta \sim \text{Dirichlet}(\alpha)$
- type $t \sim \text{Multinomial}(\theta)$
- action $a \sim \text{Multinomial}(\theta)$

For each action $a$
- $a \sim \text{Multinomial}(\tau_t)$
- ngram $n_1 \sim \text{Multinomial}(\phi_a)$
- ngram $n_2 \sim \text{Multinomial}(\phi_a)$
- click $c \sim \text{Multinomial}(\omega_a)$
We also have *entity* data...
The story for \( P(\phi, \theta, \tau, \psi, \omega, t, a, e, n, c \mid \alpha, \beta, \gamma, \eta, \iota) \)

For each action \( a \)
- \( \phi_a \sim \text{Dirichlet}(\beta) \)
- \( \omega_a \sim \text{Dirichlet}(\iota) \)

For each type \( t \)
- \( \tau_t \sim \text{Dirichlet}(\gamma) \)
- \( \psi_t \sim \text{Dirichlet}(\eta) \)

For each query \( q \)
- \( \theta \sim \text{Dirichlet}(\alpha) \)
- type \( t \sim \text{Multinomial}(\theta) \)
- action \( a \sim \text{Multinomial}(\tau_t) \)
- entity \( e \sim \text{Multinomial}(\psi_t) \)
- ngram \( n_1 \sim \text{Multinomial}(\phi_a) \)
- ngram \( n_2 \sim \text{Multinomial}(\phi_a) \)
- click \( c \sim \text{Multinomial}(\omega_a) \)
The story for $P(\phi, \theta, \tau, \psi, \omega, \sigma, t, a, e, s, n, c \mid \alpha, \beta, \gamma, \eta, \tau, \varepsilon)$

For each action/type pair, \{a, t\}
- $\phi_a \sim \text{Dirichlet}(\beta)$
- $\omega_a \sim \text{Dirichlet}(\iota)$
- $\sigma_a \sim \text{Beta}(\varepsilon)$

For each type t
- $\tau_t \sim \text{Dirichlet}(\gamma)$
- $\psi_t \sim \text{Dirichlet}(\eta)$

For each query q
- $\theta \sim \text{Dirichlet}(\alpha)$
- type $t \sim \text{Multinomial}(\theta)$
- action $a \sim \text{Multinomial}(\tau_t)$
- entity $e \sim \text{Multinomial}(\psi_t)$
- switch $s_1 \sim \text{Bernoulli}(\sigma_a)$
- switch $s_2 \sim \text{Bernoulli}(\sigma_a)$
- if ($s_1$) ngram $n_1 \sim \text{Multinomial}(\phi_a)$
- if ($s_2$) ngram $n_2 \sim \text{Multinomial}(\phi_a)$
- click $c \sim \text{Multinomial}(\omega_a)$
Apply Generative Model

A new query comes in: (e.g., “New York City hotels”)

Entity Recognition(query) → entity (“New York City”)
entity → types (“city”, “employer”, “travel destination”)
(query, entity) → context (“Ø”, “hotels”)

Historical Data(query) → distribution over hosts
• EM Posterior Probabilities to give us likelihood of each action cluster.
• action cluster → action phrase (“book hotel in”)
• (action cluster, action phrase, historical data) → best hosts (“travel.bing.com”)

query: jetbeam rrt-0
Action words from Web Trigrams

- Patterns (similar to Hearst patterns) on a Web Trigram corpus to get actions.

  - Filter Adverbs
    (e.g., “honestly”, “quickly”)

  - Filter noise
    (the 25% with lowest frequency/unigram count
    e.g., “a”, “boy”)

  - 13,417 action words
    make        download
    find        torrent
    say         eBay
    Pay         login
    Buy         podcast
    help        ...
    ...

  - Method scales to longer actions, e.g., 4-grams for 2-word actions ("read review")

  - Finds modern/web actions that older annotated corpora might miss.
Not all actions can be recommended over the Web (e.g., “shock” or “kill”). How do we find the ones that can?

- **13,417 action words**
  - make
  - download
  - find
  - torrent
  - say
  - eBay
  - Pay
  - login
  - Buy
  - podcast
  - help
  - ...

- **Web Action ngram pattern**
  - “(x) at (y)” where y has the form of a web site URL
  - “buy at Amazon.com”
  - “download at cnet.com”

- **1,279 web actions**
  - buy
  - review
  - shop
  - unsubscribe
  - book
  - download
  - ...

---

**Active Objects**
Human Annotation of Action Phrases

From each model, we first automatically generate:

19, Read biography of
19, See pictures of
19, Read blog of
19, Contact
19, Read interview with
19, Watch video of

12, Download
12, Find reviews of
12, Update
12, Get help for

24, Apply for jobs at
24, View career options at
24, View map of
24, Read news about
1, Find locations of
1, Find address of
1, Get stock quote of

Use as training data
Evaluation Setup

Model 2

Model 3

Model 4

Model 5

Model 6

Data

2,164,579 (query, host) pairs over 3 months

235,385 entities

129,088 contexts

21 types

58,123 hosts

+ Filter out Navigational Queries

50 clusters, 2-step learning over 100 total EM iterations, 2 folds per model
## Evaluation Framework

**Query 1:** "40/42 walmart address"
- **Entity:** walmart
- **Types:** location/location

<table>
<thead>
<tr>
<th>Action</th>
<th>perf</th>
<th>exc</th>
<th>good</th>
<th>fair</th>
<th>bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create account at (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Download (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find address of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find coupons for (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find map of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow sports teams of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get coupons for (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get lyrics of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get tickets for (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get weather in (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Install (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Login to (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read reviews of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read reviews of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>See map of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>See pictures of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>See schedule of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torrent (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>View map of (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watch movies with (Walmart)</td>
<td>☑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Annotation Guidelines**

- **perfect action**: Exactly the explicit intent of the user as stated in the query.
- **excellent action**: The presumed likely intent of the user as stated in the query.
- **good action**: Likely to be interesting to the user, although not the stated intent.
- **fair action**: Possibly of interest to some users who issue the query.
- **bad action**: Unlikely to be of interest to any user who issues this query.

*... or click [Nav] if all users who issue this query just want one specific website, and would not be receptive to any set of actions.*
Performance on HEAD vs. TAIL vs. Type-Balanced queries

nDCG vs. Query Sets
(with 95% confidence bounds)

More complex models are discovering more diverse actions
Tail is dominated by “People” type
Examples

Query: Webster University  
Entity: Webster University  
Context: ("", "")  
Types: /business/employer, education/university, /location/location

<table>
<thead>
<tr>
<th>Model 2 (context)</th>
<th>Model 3 (+click)</th>
<th>Model 4 (+type)</th>
<th>Model 5 (+entity)</th>
<th>Model 6 (+switch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Torrent</td>
<td>1. Torrent</td>
<td>1. Read reviews of</td>
<td>1. Read reviews of</td>
<td>1. Find address</td>
</tr>
<tr>
<td>2. Read biography</td>
<td>2. Read biography</td>
<td>2. See map of</td>
<td>2. See map of</td>
<td>2. See pictures of</td>
</tr>
<tr>
<td>3. Find adult pictures of</td>
<td>3. Read news about</td>
<td>3. Follow sports teams of</td>
<td>3. Follow sports teams of</td>
<td>3. Find map of</td>
</tr>
<tr>
<td>5. See picture of</td>
<td>5. Apply for jobs at</td>
<td>5. Apply for jobs at</td>
<td>5. Apply for jobs at</td>
<td>5. Apply for jobs at</td>
</tr>
<tr>
<td>7. Apply for jobs at</td>
<td>7. See videos with</td>
<td>7. See rankings of</td>
<td>7. See rankings of</td>
<td>7. See ranking of</td>
</tr>
</tbody>
</table>

+ User Model?

Models 4, 5, 6 automatically generate reasonable actions for this query
Examples

Query: download Skype
Entity: Skype
Context: (“download”, “”)
Types: /computer/software,/business/employer,/business/business_operation

<table>
<thead>
<tr>
<th>Model 2 (context)</th>
<th>Model 3 (+click)</th>
<th>Model 4 (+type)</th>
<th>Model 5 (+entity)</th>
<th>Model 6 (+switch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Download</td>
<td>1. Download</td>
<td>1. Download</td>
<td>1. Find on social</td>
<td>1. Download</td>
</tr>
<tr>
<td>2. Login to</td>
<td>2. Play games</td>
<td>2. Login to</td>
<td>networks</td>
<td>2. Find reviews of</td>
</tr>
<tr>
<td>4. Watch videos</td>
<td>5. Create account</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with</td>
<td>at</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Hack</td>
<td>6. Torrent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Find games with</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Watch movies</td>
<td>7. Read biography</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with</td>
<td>of</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Again, we can now automatically generate reasonable actions for queries!
1) INTRODUCTION

2) ACTIONS IN WEB SEARCH

3) ACTIVE OBJECTS

4) MINING ENTITY TYPES
Inferring Type Distributions

• Entity types are modeled as latent variables, jointly with the intended actions.

• Extensions:
  – Given admissible types from a KB such as Freebase, learn their priors and contextual disambiguation
  – Given a new term, induce the types of the term
  – Automatically induce type list and admissible types for arbitrary entities
Generative process for entity bearing queries.

For each query $q$

- entity $e \sim \text{Multinomial}(\psi)$
- type $t \sim \text{Multinomial}(\tau_e)$
- ngram $n_1 \sim \text{Multinomial}(\phi_t)$
- ngram $n_2 \sim \text{Multinomial}(\phi_t)$
+ Empty Switch

Model M0

Active Objects
+ Empty Switch + Click
+ Empty Switch + Click + Action
Experimental Setting

• Training
  – Queries from 3 months of US Bing search usage logs
  – Entities from 73 Freebase types, accounting for 50% query traffic in US market
  – Model parameters trained using 2-step learning over 100 EM iterations, 2 folds per model

• Testing
  – Query-weighted random sample of 500 HEAD and 500 TAIL entity-bearing queries
  – 7 paid independent annotators identified all applicable Freebase types to the entities in the queries
    • Two annotators per query
    • Fleiss’ $\kappa$ was 0.445, moderate agreement
Performance Analysis

<table>
<thead>
<tr>
<th></th>
<th>nDCG</th>
<th>MAP</th>
<th>MAP_W</th>
<th>Prec@1</th>
<th></th>
<th>nDCG</th>
<th>MAP</th>
<th>MAP_W</th>
<th>Prec@1</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_FB</td>
<td>0.71</td>
<td>0.60</td>
<td>0.45</td>
<td>0.30</td>
<td></td>
<td>0.73</td>
<td>0.64</td>
<td>0.49</td>
<td>0.35</td>
</tr>
<tr>
<td>Guo'09</td>
<td>0.79†</td>
<td>0.71†</td>
<td>0.62†</td>
<td>0.51†</td>
<td></td>
<td>0.80†</td>
<td>0.73†</td>
<td>0.66†</td>
<td>0.52†</td>
</tr>
<tr>
<td>M0</td>
<td>0.79†</td>
<td>0.72†</td>
<td>0.65†</td>
<td>0.52†</td>
<td></td>
<td>0.82†</td>
<td>0.75†</td>
<td>0.67†</td>
<td>0.57†</td>
</tr>
<tr>
<td>M1</td>
<td>0.83‡</td>
<td>0.76‡</td>
<td>0.72‡</td>
<td>0.61‡</td>
<td></td>
<td>0.81†</td>
<td>0.74†</td>
<td>0.67†</td>
<td>0.55†</td>
</tr>
<tr>
<td>IM</td>
<td>0.87‡</td>
<td>0.82‡</td>
<td>0.77‡</td>
<td>0.73‡</td>
<td></td>
<td>0.80†</td>
<td>0.72†</td>
<td>0.66†</td>
<td>0.52†</td>
</tr>
</tbody>
</table>

• † indicates statistical significance over B_FB, and ‡ over both B_FB and Guo’09.
• Bold indicates statistical significance over all non-bold models in the column.
• M1 (empty context + click signal) significantly outperforms baseline and Guo’09, on HEAD.
• IM significantly better over all models and across all metrics
  — Biggest gains in first position of its ranking (Prec@1 metric).
Switch Parameter Analysis

- Switch improves performance across all models
- More expressive models benefit more from switch
Discussion

• Why performance on TAIL lower than expected?
  – TAIL is skewed towards the PEOPLE types
  – Latent actions are over-expressive and they do not help in differentiating PEOPLE types
    • Inspection of latent Action parameter in IM shows that most PEOPLE types have all their mass distributed to three generic and common intents (see pictures of, find biographical information about, and see video of)

• Success case in the TAIL
  – “ymca” -> {song, place, educational_institution}
    • Marginalizing out the context words gives the following intent priors: 0.63, 0.29, 0.08
  – $q_1 = \text{“jamestown ymca ny”} \rightarrow \text{IM correctly classified “ymca” as a place}$
  – $q_2 = \text{“ymca palomar”} \rightarrow \text{IM correctly classified “ymca” as a educational_institution}$
Wrap-Up
hannah montana

Hannah Montana Forever | Disney Channel
Explore the Hannah Montana Forever home page for videos, full-length episodes, character bios, games, photos, downloads, music, and more!
tv.disney.go.com/disneychannel/hannahmontana

Videos of hannah montana

Hannah Montana - Wikipedia, the free encyclopedia
Production · Cast · Episodes · Films
Hannah Montana is an American television series, which debuted on March 24, 2006 on Disney Channel, and also airs Saturday mornings on ABC as part of the ABC Kids block.
en.wikipedia.org/wiki/Hannah_montana

Character Bios | Hannah Montana Forever | Disney Channel
Meet Miley, Lilly, Oliver, Rico, and the rest of the gang from Disney Channel's Hannah Montana Forever!
tv.disney.go.com/disneychannel/hannahmontana/characters/index.html
A hodgepodge of related strings
• Only actionable through search
Big Wins

- Typed relations
- User Interface accesses structured data
Big Wins

- Typed relations
- User Interface accesses structured data
- Click through experience can now leverage strongly-typed identifier
Active Objects

162

Big Wins

- Typed relations
- User Interface accesses structured data
- Click through experience can now leverage strongly-typed identifier
- Brokered Actions (one click conversions)
To Do

- Fixed actions
- Model tasks
- Annotate URLs/Apps with actions
- Think of UI