Mining Action Intents in Web Search

Patrick Pantel

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

June 2012
### Price prediction

<table>
<thead>
<tr>
<th>Flights from Seattle, WA to Maui, HI</th>
<th>WEEKEND ESTIMATES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FROM</strong></td>
<td><strong>TO</strong></td>
</tr>
<tr>
<td>Seattle, WA (SEA) - Seattle</td>
<td>Kahului, HI (OGG) - Kahului</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Task completion**

- Flights to Maui

### Aggregate ratings

- Royal Lahaina Resort
  - Address: 2780 Kekaa Dr, Lahaina, HI 96761
  - Phone: (808) 661-9469
  - Rating: ★★★★☆ (12)

**Task completion**

- Search for Royal Lahaina Resort

**Sample rate:** $230

**Compare rates from major travel sites**

- Check in: 03/15/2012
- Check out: 03/17/2012

[Find rate](http://bing.com/travel/hotels/search?q=royal+lahaina+resort)
How old is Mariah Carey?

Mariah Carey, age 42

Mariah Carey was born on March 27, 1970 on Long Island, New York. She is an American singer, songwriter, record producer, and actress. She made her recording debut in 1990 under the guidance of Columbia Records executive Tommy Mottola, and released her self-titled debut studio album, Mariah Carey.

Born: March 27, 1970 (age 42), Huntington, New York
Net worth: US$ 500 million (2011)
Spouse: Nick Cannon (m. 2008), Tommy Mottola (m. 1993–1998)
Children: Monroe Cannon, Moroccan Scott Cannon

Song | Year | Album
--- | --- | ---
Without You | 1993 | Music Box
Hero | 1993 | Music Box
My All | 1997 | Butterfly
We Belong Together | 2006 | The Emancipation of Mimi
When You Believe | 1998 | #1's

Nick Cannon kidney failure: Mariah Carey's husband in hospital as ... www.dailymail.co.uk.../Nick-Cannon-kidney-failure-Mariah-Carey...
Jan 5, 2012 ... He's in a lot of pain'. Mariah Carey keeps bedside vigil as husband Nick ... The 31-year-old tweeted an update yesterday afternoon after being ...
Current Experience

[Image of a search result page showing a search for "jetbeam rrt-0" on Bing.]

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

**SEARCH HISTORY**
- Search more to see your history
- See all
- Clear all · Turn off

**ALL RESULTS**

- **JETBeam RRT-0 R5 with Infinite Ramping LED Flashlight**
  Going Gear JETBeam RRT-0 R5 with Infinite Ramping LED Flashlight [jetbeamrrt0r5] - The popular RRT-0 has been updated to be an even more versatile flashlight, ...
goinngear.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 didnt ...
  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

Recognize entity in query

Actions easily accessible

Web Images Videos Shopping News Maps More | MSN Hotmail

[Image]

Jetbeam RRT-0
Flashlight (Product)

JETBeam RRT-0 R5 with Infinite Ramping LED Flashlight...
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, Search & Rescue and General Outdoors The Rapid ...
products/RRT%2d0.html · Mark as spam

Jetbeam RRT-0
Raptor today. All I can say is to JETBeam! This thing is ‘... BEFORE MY HEAD EXPLODES!!! Ok... it didn’t ...

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

Active Objects
Better experience

Recognize entity in query

Actions easily accessible

Active Objects
Better experience

- Recognize entity in query
- Actions easily accessible
Politicians

Actions easily accessible

Recognize entity in query

Barack Obama
(President)

Obama for America | 2012
Barackobama.com is the official re-election campaign website of President Barack Obama. Visit the site for the latest updates from the Obama campaign, 2012 election news...

Official Obama Website

Other ideas: obama barack bio · president obama website

Obama - Wikipedia, the free encyclopedia
Barack Obama II (born August 4, 1961) is the 44th and current President of the United States. He is the first African American to hold the office. Obama ...

Watch Weekly Address
Follow Social Updates
Read Biography
View Approval Ratings
Support Candidacy

Related searches

Mark Obama and other U.S. politicians by office, state, and branch of government
Films

Recognize entity in query

Actions easily accessible

Pandorum (2009 Film)

Pandorum - Now Available on DVD & Blu-ray
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/tt1192789 - Mark as spam

Netflix
Amazon Instant Video
Zune Store
iTunes
more...
1) INTRODUCTION

2) ACTIONS IN WEB SEARCH

3) ACTIVE OBJECTS

4) MINING ENTITY TYPES
**Actions vs Intents**

**User Intents and Goals**
- **Query**: plan vacation, get in shape
- **Query Intent**: informational, navigational, transactional
- **Finer-grained Intents**: advice, locate, download, obtain, interact
- **Actions on Entities**: read reviews(hotel), add to Netflix queue(film), get address(landmark), 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”)

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

- 40% creative work
- 37% organization
- 9% product
- 8% person
- 3% event
- 3% other

*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>
</tr>
</thead>
<tbody>
<tr>
<td>1x Find Location(s) (on an Organization entity)</td>
</tr>
<tr>
<td>1x Find Lyrics (on a CreativeWork / MusicalTrack entity)</td>
</tr>
<tr>
<td>2x Find Recipe For (on a food)</td>
</tr>
<tr>
<td>1x Find Where to Buy (on a Product entity)</td>
</tr>
<tr>
<td>2x Get Contact Information (on an Organization entity)</td>
</tr>
<tr>
<td>1x Get Directions To (on an Organization / Location entity)</td>
</tr>
<tr>
<td>2x Get Domain Information (on a Website entity)</td>
</tr>
<tr>
<td>1x Get Event Details (on an Event entity)</td>
</tr>
<tr>
<td>2x Get Event Results (on an Event entity)</td>
</tr>
<tr>
<td>4x Product Detail (on a Product entity)</td>
</tr>
<tr>
<td>29x Learn (on any entity)</td>
</tr>
<tr>
<td>6x Learn / Educational (on a Person / Product / Organization entity)</td>
</tr>
<tr>
<td>1x Learn / Trivia (on any entity)</td>
</tr>
<tr>
<td>1x Operating Hours (on an Organization entity)</td>
</tr>
<tr>
<td>3x Read Articles (on a News / Magazine entity)</td>
</tr>
<tr>
<td>1x Read Guide (on a Product entity)</td>
</tr>
<tr>
<td>1x Read Help (on a Product entity)</td>
</tr>
<tr>
<td>8x Read News About (on any entity)</td>
</tr>
<tr>
<td>3x Read Reviews (Shopping on a CreativeWork / Product / Service entity)</td>
</tr>
<tr>
<td>1x Read Spoilers (on a CreativeWork)</td>
</tr>
<tr>
<td>8x Research (focused information gathering, on any entity)</td>
</tr>
<tr>
<td>8x Search Database of (e.g., obituaries, on an Organization / Website)</td>
</tr>
<tr>
<td>See Menu (on a Restaurant)</td>
</tr>
<tr>
<td>3x See Pictures (on a Person / Product / Organization entity)</td>
</tr>
<tr>
<td>Side Effects / Safety (on a Product entity)</td>
</tr>
<tr>
<td>Stock Price (on an Organization entity)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transactional (navigating to a web-mediated action)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1x Apply for Job (on a LocalBusiness / Organization entity)</td>
</tr>
<tr>
<td>4x Buy (Shopping on a Product entity)</td>
</tr>
<tr>
<td>3x Buy Tickets (on an Event / Product / Person entity)</td>
</tr>
<tr>
<td>3x Content Creation (on a Website entity)</td>
</tr>
<tr>
<td>1x Discuss Online (on any entity)</td>
</tr>
<tr>
<td>5x Download (on a CreativeWork or Software entity)</td>
</tr>
<tr>
<td>1x Listen to Music (on a CreativeWork or Website entity)</td>
</tr>
<tr>
<td>1x Manage Account (on a Local Business / Website / Org entity)</td>
</tr>
<tr>
<td>14x Pay Bill (on a Website / Organization entity)</td>
</tr>
<tr>
<td>1x Play Game (on a Game entity)</td>
</tr>
<tr>
<td>2x Rent (on a CreativeWork / Product entity)</td>
</tr>
<tr>
<td>2x Reservation (on a Hotel entity)</td>
</tr>
<tr>
<td>1x Schedule Appointment (on a LocalBusiness entity)</td>
</tr>
<tr>
<td>1x Sell (Shopping on a Product entity)</td>
</tr>
<tr>
<td>1x Use Service On (e.g., translate, on a Website)</td>
</tr>
<tr>
<td>6x Watch Video About (on any entity)</td>
</tr>
<tr>
<td>1x Web Chat</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>13x Shopping (category of actions including reviews and buying)</td>
</tr>
<tr>
<td>19x Various/Unknown</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

---

**Active Objects**
1) INTRODUCTION

2) ACTIONS IN WEB SEARCH

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

2,164,579 (query, host) pairs over 3 months
235,385 entities
21 types
129,088 contexts
58,123 hosts

get reviews ← Orlando hotel reviews
read biography
download software
watch shows online
Does Hope Solo have a boyfriend?
Free Winzip download
watch family guy online
**Model 1**

**Goal:** Define a theory for how actionable queries are generated.

The story for $p(\text{actionable query})$, or more formally

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

For each action $a$

$$\phi_a \sim \text{Dirichlet}(\beta) \quad \text{(action} \rightarrow \text{contexts die)}$$

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 $n \sim \text{Multinomial}(\phi_a)$
The story for $p(\phi, \theta, a, n | \alpha, \beta)$

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

“ebert Star Wars review" 
read reviews action

Model 1.02

Model 2
Clicked hosts matter...

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

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

For each action $a$

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

$\omega_a \sim \text{Dirichlet}(\iota)$ (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 $P(\phi, \theta, \tau, \omega, t, a, n, c \mid \alpha, \beta, \gamma, \lambda)$

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

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

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 \( a \)
\[
\phi_a \sim \text{Dirichlet}(\beta) \\
\omega_a \sim \text{Dirichlet}(i)
\]

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) \\
\text{type } t \sim \text{Multinomial}(\theta) \\
\text{action } a \sim \text{Multinomial}(\tau_t) \\
\text{entity } e \sim \text{Multinomial}(\psi_t) \\
\text{ngram } n_1 \sim \text{Multinomial}(\phi_a) \\
\text{ngram } n_2 \sim \text{Multinomial}(\phi_a) \\
\text{click } c \sim \text{Multinomial}(\omega_a)
The story for $P(\phi, \theta, \tau, \psi, \omega, \sigma, t, a, e, s, n, c | \alpha, \beta, \gamma, \eta, \iota, \epsilon)$

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

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 $c \sim \text{Multinomial}(\omega_a)$
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.

**Web Trigrams**

**Pattern Match**
- “want to (x)”
- “have to (x)”
- “you can (x)”
- “I can (x)”

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

**Filter noise**
- (the 25% with lowest $f_{frequency}$ unigram count)
  - e.g., “a”, “boy”

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

↑

Finds modern/web actions that older annotated corpora might miss.

↑

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

---

**Active Objects**
Web Action words from Trigrams

- 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

Cluster 19 (149.1%)
- divorce
- interview
- book
- show
- contact
- email
- film
- about
- update

Cluster 12 (141.9%)
- install
- download
- search
- setup
- beta

Cluster 24 (125.6%)
- store
- address
- about
- travel
- mortgage
- contact
- login
- search

Use as training data

en.wikipedia.org (19.787474%)
www.answers.com (4.219255%)
www.amazon.com (1.090042%)
www.absoluteastronomy.com (1.048446%)
www.huffingtonpost.com (1.017043%)
www.thesaurus.com (0.923841%)
answers.yahoo.com (0.916725%)
www.youtube.com (0.810244%)
www.biography.com (0.781476%)
www.freebase.com (0.736192%)
www.encyclopedia.com (0.688985%)

download.cnet.com (2.850659%)
www.google.com (1.941431%)
www.skype.com (1.654065%)
en.wikipedia.org (1.327014%)
earth.google.com (0.886125%)
www borrsoft.com (0.790033%)
www.softpedia.com (0.564006%)
en.softonic.com (0.447494%)
sourcesforge.net (0.443420%)
www.xionline.org (0.435577%)
www.maplandia.com (0.413988%)

en.wikipedia.org (11.28587%)
www.yelp.com (2.08095%)
www.answers.com (1.634436%)
www.manta.com (1.436725%)
www.indeed.com (1.219442%)
investing.businessweek.com (1.18738%)
www.supercamps.com (0.655248%)
www.lasso.com (0.552158%)

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

Model 2  Model 3  Model 4  Model 5  Model 6

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

Data

2,164,579 (query, host) pairs over 3 months
235,385 entities  21 types
129,088 contexts  58,123 hosts

+ Filter out Navigational Queries
### Evaluation Framework

#### Active Objects

**Combined set of actions over all models**

<table>
<thead>
<tr>
<th>Query 1: &quot;40/42 walmart address&quot;</th>
<th>Annotation Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Entity:</strong> walmart</td>
<td>perfect action</td>
</tr>
<tr>
<td><strong>Types:</strong> location/location</td>
<td>Exactly the explicit intent of the user as stated in the query.</td>
</tr>
<tr>
<td>Create account at (Walmart)</td>
<td>excellent action</td>
</tr>
<tr>
<td>Download (Walmart)</td>
<td>The presumed likely intent of the user as stated in the query.</td>
</tr>
<tr>
<td>Find address of (Walmart)</td>
<td>good action</td>
</tr>
<tr>
<td>Find coupons for (Walmart)</td>
<td>Likely to be interesting to the user, although not the stated intent.</td>
</tr>
<tr>
<td>Find map of (Walmart)</td>
<td>fair action</td>
</tr>
<tr>
<td>Follow sports teams of (Walmart)</td>
<td>Possibly of interest to some users who issue the query.</td>
</tr>
<tr>
<td>Get coupons for (Walmart)</td>
<td>bad action</td>
</tr>
<tr>
<td>Get lyrics of (Walmart)</td>
<td>Unlikely to be of interest to any user who issues this query.</td>
</tr>
<tr>
<td>Get tickets for (Walmart)</td>
<td></td>
</tr>
<tr>
<td>Get weather in (Walmart)</td>
<td></td>
</tr>
<tr>
<td>Install (Walmart)</td>
<td></td>
</tr>
<tr>
<td>Login to (Walmart)</td>
<td></td>
</tr>
<tr>
<td>Read reviews of (Walmart)</td>
<td></td>
</tr>
<tr>
<td>Read reviews of (Walmart)</td>
<td></td>
</tr>
<tr>
<td>See map of (Walmart)</td>
<td></td>
</tr>
<tr>
<td>See pictures of (Walmart)</td>
<td></td>
</tr>
<tr>
<td>See schedule of (Walmart)</td>
<td></td>
</tr>
<tr>
<td>Torrent (Walmart)</td>
<td></td>
</tr>
<tr>
<td>View map of (Walmart)</td>
<td></td>
</tr>
<tr>
<td>Watch movies with (Walmart)</td>
<td></td>
</tr>
</tbody>
</table>

... 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.

---

**x 7**

UHRS Fair κ agreement
Performance on HEAD vs. TAIL vs. Type-Balanced queries

Tail is dominated by “People” type
Action Discovery, Diversity

The graph illustrates the total probability of actions given type across different cluster ranks for three models: Model 4, Model 5, and Model 6. The x-axis represents the cluster rank, while the y-axis shows the total probability of action given type. The keywords 'biography' and 'download' are prominently featured, indicating the focus of the analysis on these topics.
### 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</td>
<td>3. Read news</td>
<td>3. Follow sports</td>
<td>3. Follow sports</td>
<td>3. Find map of</td>
</tr>
<tr>
<td>pictures of</td>
<td>about</td>
<td>teams of</td>
<td>teams of</td>
<td>about</td>
</tr>
<tr>
<td>5. See picture of</td>
<td>of</td>
<td>5. Apply for jobs</td>
<td>5. Apply for jobs</td>
<td>about</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>6. See cost of</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7. See ranking of</td>
</tr>
</tbody>
</table>

*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 networks</td>
<td>1. Download</td>
</tr>
<tr>
<td>2. Login to</td>
<td>2. Play games</td>
<td>2. Login to</td>
<td>2. Download</td>
<td>2. Find reviews of</td>
</tr>
<tr>
<td>5. Hack</td>
<td>5. Create account at</td>
<td>5. Create account at</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Find games with</td>
<td>Torrent</td>
<td>6. Torrent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Watch movies with</td>
<td>Read biography of</td>
<td>7. Read biography of</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
+ 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>HEAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TAIL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_{FB}</td>
<td>0.71</td>
<td>0.60</td>
<td>0.45</td>
<td>0.30</td>
<td>0.73</td>
<td>0.64</td>
<td>0.49</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Guo'09</td>
<td>0.79↑</td>
<td>0.71↑</td>
<td>0.62↑</td>
<td>0.51↑</td>
<td>0.80↑</td>
<td>0.73↑</td>
<td>0.66↑</td>
<td>0.52↑</td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>0.79↑</td>
<td>0.72↑</td>
<td>0.65↑</td>
<td>0.52↑</td>
<td>0.82↑</td>
<td>0.75↑</td>
<td>0.67↑</td>
<td>0.57↑</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>0.83↑</td>
<td>0.76↑</td>
<td>0.72↑</td>
<td>0.61↑</td>
<td>0.81↑</td>
<td>0.74↑</td>
<td>0.67↑</td>
<td>0.55↑</td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td>0.87↑</td>
<td>0.82↑</td>
<td>0.77↑</td>
<td>0.73↑</td>
<td>0.80↑</td>
<td>0.72↑</td>
<td>0.66↑</td>
<td>0.52↑</td>
<td></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
  – q₁ = “jamestown ymca ny” → IM correctly classified “ymca” as a place
  – q₂ = “ymca palomar” → IM correctly classified “ymca” as an educational_institution
Wrap-Up
A hodgepodge of related strings
Only actionable through search
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 · Mark as spam

Entertainment News: hannah montana
Billy Ray Cyrus: ‘Hannah Montana’ destroyed family
NASHVILLE, Tenn. (AP) -- Billy Ray Cyrus says the Disney TV show “Hannah Montana” destroyed his family, causing his divorce and sending daughter Miley Cyrus spinning out of... Yahoo Finance · 10 hours ago
Miley Cyrus Angry With Dad Over Hannah Montana Slur Entertainmentwise
Popservations: Lady Gaga at the Grammys, HBO shows and Hannah Montana Florida Times-Union

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 ... en.wikipedia.org/wiki/Hannah_montana · Mark as spam

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 · Mark as spam

Hannah Montana Wigs
15,000 Costumes - All Sizes & Styles. Low Prices & Same Day Shipping!
www.BuyCostumes.com
See your message here
Big Wins

- Typed relations
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
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
- Active Objects as an entity-centric UI