

CS544: Named Entity Discrimination

January 27, 2011

Zornitsa Kozareva
USC/ISI

Marina del Rey, CA
kozareva@isi.edu
www.isi.edu/~kozareva

Who is Jerry Hobbs?



Search Engine

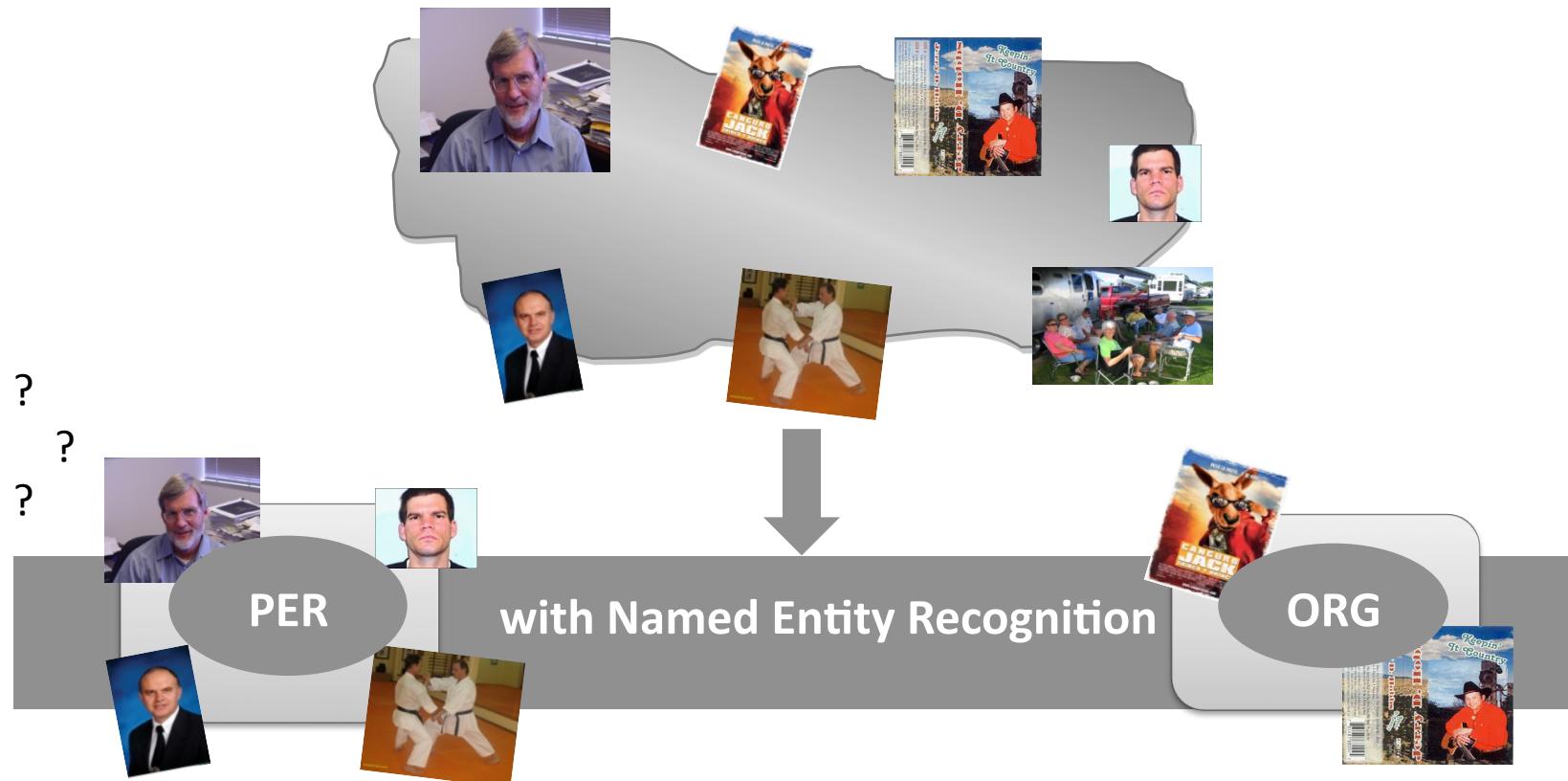
Jerry R. Hobbs. Address: USC/ISI 4676 Admiralty Way ... *Jerry R. Hobbs hobbs@isi.edu*

AMW.com **Jerry Hobbs**-Fugitive America's Most Wanted is a long-running American TV

Alleged Killer Dad Denied Bond, Services Set for Two Little Girld. A judge had denied

Jerry Hobbs- Wikipedia, the free encyclopedia Dr. *Jerry R. Hobbs* (born 25 January

194 World Martial Arts Games Rank: 3rd Dan, Discipline: Ju-Jutsu Issued By: **Jerry Hobbs**
10th Dan Rank: 1st



NE Recognition vs. NE Discrimination

- NE Recognition = detection & classification of entity mentions into a predefined set of categories.
⇒ achieves only a partial disambiguation of names
- NE Discrimination = finding the actual entity denoted by a particular name occurrence in text.





Jerry Hobbs

Search

Advanced Search

Web Show options...

[Jerry R. Hobbs](#)

Oct 14, 2003 ... **Jerry R. Hobbs**. Address: USC/ISI 4676 Admiralty Way ... [Jerry R. Hobbs](#)
hobbs@isi.edu. USC/ISI, 4676 Admiralty Way, Marina del Rey, CA 90292 ...
www.isi.edu/~hobbs/ - [Cached](#) - [Similar](#)

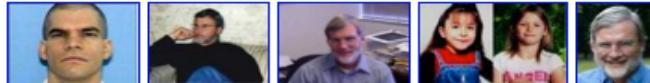
[AMW | Fugitives | Jerry Hobbs | Brief](#)

May 10, 2005 ... Fugitives | **Jerry Hobbs** - Brief - Father Denied Bail Awaits Trial For Children
s Murders Jerry Branton Hobbs accused of the stabbing deaths ...
www.amw.com/fugitives/brief.cfm?id=31767 - [Cached](#) - [Similar](#)

[Jerry Hobbs - Wikipedia, the free encyclopedia](#)

Dr. **Jerry R. Hobbs** (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial ...
en.wikipedia.org/wiki/Jerry_Hobbs - [Cached](#) - [Similar](#)

[Image results for Jerry Hobbs](#) - Report images



[Bluhm Blog: DNA test results suggest that Jerry Hobbs's confession ...](#)

Nov 14, 2008 ... Who could forget this picture? Could **Jerry Hobbs** be innocent of capital
murder? More than three years ago, on Mother's Day, residents or ...
biogaw.northwestern.edu/.../dna-test-results-suggest-that-jerry-hobbs-confession-is-false.html - [Cached](#) - [Similar](#)

[Alleged Killer Dad Denied Bond - CBS News](#)

May 11, 2005 ... **Jerry Hobbs**, who is accused of killing his 9-year-old daughter and her best ...
On Wednesday, a judge denied bail for **Jerry Hobbs**, 34, ...
www.cbsnews.com/stories/2005/05/11/.../main694398.shtml - [Cached](#) - [Similar](#)

[Jerry Hobbs](#)

www.ai.sri.com/~hobbs/ - [Similar](#)

[Jerry Hobbs News](#)

Articles covering **Jerry Hobbs**: Lake County judge denies bail request for **Jerry Hobbs**, suspect in ... / **Jerry Hobbs** Hearing / DNA clears suspect **Jerry Hobbs**, ...
jerry-hobbs-news.newslib.com/ - [Cached](#) - [Similar](#)

[Orange Tangerine: What should be done with Jerry Hobbs](#)

May 11, 2005 ... **Jerry Hobbs** is the rage-filled, domestic-abusing career criminal who killed his
8-year-old daughter and her 9-year-old friend, with scarcely ...
orangetangerine.blogspot.com/.../what-should-be-done-with-jerry-hobbs.html -
[Cached](#) - [Similar](#)

[Wayne Cryts - One Man With Courage](#)

Jerry Hobbs, Author. A fifth generation farmer, Wayne Cryts finished harvesting his crop in the
fall of 1960 and hauled more than 32000 bushels of soybeans ...
www.waynecryts.com/ - [Cached](#) - [Similar](#)

[Guest Speaker: Jerry R. Hobbs](#)

Bio: Dr. **Jerry R. Hobbs** is a prominent researcher in the fields of computational linguistics,
discourse analysis, and artificial intelligence. ...
www.cs.umn.edu/medlang/speakers/hobbs.html - [Cached](#)

Ideally, we want



Dr. *Jerry R. Hobbs* is a prominent researcher
in the fields of computational linguistic

Jerry R. Hobbs. Address: USC/ISI 4676
Admiralty Way .



Jerry Hobbs is the rage-filled, domestic-
abusing career criminal who killed his

Jerry Hobbs, who is accused of killing his
8-year-old daughter and her best



Jerry Hobbs, a fifth generation farmer,
Wayne Cryts finished harvesting his crop

U.S. Census Bureau states 90,000 names are shared by 100,000,000 people

More details for Jerry Hobbs: [Find Email](#) [Hidden Profiles](#) [Address History](#)



Jerry Hobbs male, 68 years old

Yale University, menlo park, abduction, SRI International, Information Sciences Institute, Knowledge representation... [Jerry R. Hobbs](#)

Dr. Jerry R. Hobbs (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial...



Jerry male, 58 years old (Marina del Rey, California, United States)

single, Aquarius



Jerry Hobbs female, 20 years old (Scottsboro, Alabama, United States)

single, high school, Pisces, White / Caucasian, A few extra pounds, straight

'Remove Profile'.



Jerry male, 35 years old (Homestead, Florida, United States)

single, security officer, gay, virgo, Grad / professional school, White / Caucasian

I am easy going I like to drive fast meet friendly people and hang-out I am looking for friends I can hangout with and go places with so let me know.



Jerry Hobbs male, 42 years old (Las Vegas, Nevada, United States)

single, gemini, slim, Grad / professional school, White / Caucasian, From Las Vegas

Born to race!



Jerry Hobbs male, 26 years old (Grove City, Pennsylvania, United States)

single, leo



Jerry Hobbs male (Orlando, Florida, United States)



Jerry Hobbs male, 23 years old (United States)

gemini, in a relationship, straight, likes Not a big fan of them (book), likes look at interests music, likes only adult swim (TV show)

i am a high person (that is about my personality). not in to homosexual males i will cus fags out. lezbians are f***ing amsome.



Jerry Hobbs female (United Kingdom)

voluptuous, Sagittarius, White / Caucasian, in a relationship, straight, From BASILDON

Hi well my name is jerry, im 23 and live in good old basildon, essex. i have a partner of 2 years and a little boy who is 8 months old, im on here...



Jerry Hobbs male (Bloomington, Indiana, United States)

single, Libra

The first debate was in Oxford.

Oxford, Mississippi - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W http://en.wikipedia.org/wiki/Oxford,_Mississippi

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo! RealPlayer Windows Marketplace Windows Media Yahoo! Answers http://tmsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... Oxford, Mississippi - Wikipedia, t...

Help us improve Wikipedia by supporting it financially.

article discussion edit this page history

Oxford, Mississippi

From Wikipedia, the free encyclopedia

WIKIPEDIA
The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Transferring data from upload.wikimedia.org...

start

Oxford - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W http://en.wikipedia.org/wiki/Oxford

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo! RealPlayer Windows Marketplace Windows Media Yahoo! Answers http://tmsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... W Oxford - Wikipedia, the free enc...

Help us improve Wikipedia by supporting it financially.

Log in / create account

article discussion edit this page history

Oxford

From Wikipedia, the free encyclopedia

WIKIPEDIA
The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

This article may require cleanup to meet Wikipedia's quality standards.
Please improve this article if you can. (May 2008)

This article is about the city of Oxford in England. For other cities and other meanings, see Oxford (disambiguation).

Oxford is currently bidding for the 2010 Wikimania Conference.

Oxford (pronounced /'ɔksfərd/, listen (help·info)) is a city, and the county town of Oxfordshire, in South East England. It has a population of 151,000. The rivers Cherwell and Thames run through Oxford and meet south of the city centre. For a distance of some 10 miles (16 km) along the river, in the vicinity of Oxford, the Thames is known as The Isis.

Oxford is home to the University of Oxford, the oldest university in the English-speaking world.

Buildings in Oxford reflect every English architectural period since the arrival of the Saxons, including the mid-18th century Radcliffe Camera, the hub of the city. Oxford is known as the "city of dreaming spires", a term coined by poet Matthew Arnold in reference to the harmonious architecture of Oxford's university buildings.

Oxford, Alabama - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W http://en.wikipedia.org/wiki/Oxford,_Alabama

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo! RealPlayer Windows Marketplace Windows Media Yahoo! Answers http://tmsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... W Oxford, Alabama - Wikipedia, the...

Your continued donations keep Wikipedia running!

Log in / create account

article discussion edit this page history

Oxford, Alabama

From Wikipedia, the free encyclopedia

WIKIPEDIA
The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Done

Contents [hide]

Oxford, Alabama

From Wikipedia, the free encyclopedia

Coordinates: 33°59'10.5" N 85°38'8.8" W

Oxford, Alabama

Seal

Nickname(s): Crossroads To The Future

Location in Alabama

Coordinates: 33°35'49.578"N 85°00'19.972"W

Country United States
State Alabama
Counties Calhoun, Talladega
Government - Mayor Leon Smith

Geography

Oxford is located at 33°35'49.578"N, 85°00'19.972"W (33.97105, -85.838881).
According to the U.S. Census Bureau, the city has a total area of 18.3 square miles (47.4 km²), of which, 18.2 square miles (47.2 km²) of it is land and 0.1 square miles (0.2 km²) of it (0.38%) is water.

Oxford lies among the foothills at the southern end of the Blue Ridge mountains. Nearby Mount Cheaha is Alabama's highest point and offers expansive views of the surrounding wilderness and the city below. Much of the city's southern border is shared with the Talladega National Forest.

Major bodies of water include Oxford Lake, Lake Hillabee, and Choccolocco Creek, which bisects the city.

The first debate was in **Oxford**. A draw for Obama would be considered a win.

Oxford, Mississippi - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W http://en.wikipedia.org/wiki/Oxford,_Mississippi

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo! RealPlayer Windows Marketplace Windows Media Windows Yahoo! Answers http://tmsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... Oxford, Mississippi - Wikipedia, t...

Help us improve Wikipedia by supporting it financially.

article discussion edit this page history

Oxford, Mississippi

From Wikipedia, the free encyclopedia

WIKIPEDIA The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Transferring data from upload.wikimedia.org...

start

Oxford - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W http://en.wikipedia.org/wiki/Oxford

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo! RealPlayer Windows Marketplace Windows Media Windows Yahoo! Answers http://tmsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... W Oxford - Wikipedia, the free enc...

Help us improve Wikipedia by supporting it financially.

Log in / create account

article discussion edit this page history

Oxford

From Wikipedia, the free encyclopedia

Coordinates: 51°45'07"N 1°15'28"W

This article may require cleanup to meet Wikipedia's quality standards. Please improve this article if you can. (May 2008)

This article is about the city of Oxford in England. For other cities and other meanings, see Oxford (disambiguation).

Oxford is currently bidding for the 2010 Wikimania Conference

Oxford (pronounced /'oks-fərd/, listen) is a city, and the county town of Oxfordshire, in South East England. It has a population of 151,000. The rivers Cherwell and Thames run through Oxford and meet south of the city centre. For a distance of some 10 miles (16 km) along the river, in the vicinity of Oxford, the Thames is known as The Isis.

Oxford is home to the University of Oxford, the oldest university in the English-speaking world.

Buildings in Oxford reflect every English architectural period since the arrival of the Saxons, including the mid-18th century Radcliffe Camera, the hub of the city. Oxford is known as the "city of dreaming spires", a term coined by poet Matthew Arnold in reference to

Oxford, Alabama - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W http://en.wikipedia.org/wiki/Oxford,_Alabama

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo! RealPlayer Windows Marketplace Windows Media Windows Yahoo! Answers http://tmsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... W Oxford, Alabama - Wikipedia, the...

Your continued donations keep Wikipedia running!

Log in / create account

article discussion edit this page history

Oxford, Alabama

From Wikipedia, the free encyclopedia

Coordinates: 33.597105, -85.838881

Oxford, Alabama

Seal

Nickname(s): Crossroads To the Future

Location in Alabama

Coordinates: 33°35'49.578"N 85°50'19.972"W (33.597105, -85.838881)

Country United States

State Alabama

Counties Calhoun, Talladega

Government Mayor Leon Smith

Oxford is a city in Calhoun and Talladega counties in the U.S. state of Alabama. The population was 14,592 at the 2000 census. Oxford is one of two principal cities of and included in the Anniston-Oxford Metropolitan Statistical Area.

Contents [hide]

- Geography
- History
- Government
- Education
- Economy
- Demographics
- Points of interest
- Media
- References
- External links

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Done

start

Georgia (country) - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo RealPlayer Windows Marketplace Windows Media Windows Yahoo! Answers http://timsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... W Georgia (country) - Wikipedia, th... W John McCarthy - Wikipedia, the free en... Ke hoach trinh hieu truong - Span - Ya...

Log in / create account

Georgia (country)

From Wikipedia, the free encyclopedia

This article is about the Eurasian country. For the US state, see [Georgia \(U.S. state\)](#). For other uses, see [Georgia \(disambiguation\)](#).

Georgia ([dʒɔdʒə] (help·info); Georgian: საქართველო, *Sakartvelo*) is a trans-[Georgian \(U.S. state\)](#) Caucasus region, situated at the meeting line between Europe and Asia^[2]. Georgia is bordered to the north by the [Russian Federation](#), to the east by Azerbaijan, to the south by Armenia and to the southwest by [Turkey](#).^[3] Georgia covers a territory of 69,700 km²; its population, excluding Abkhazia and South Ossetia (the Tskhinvali region), is 4.4 million, with nearly 84% ethnic Georgians.^[4]

The history of Georgia can be traced back to the ancient Kingdom of Iberia, that was one of the first countries to adopt Christianity as an official religion early in the 4th century. At the beginning of the 19th century Georgia became part of the Russian Empire. After a brief period of independence following the Russian Revolution of 1917 Georgia was forcibly incorporated into the Soviet Union in 1922.

Independence was restored in 1991. Like many post-communist countries Georgia suffered from the economic crisis and civil unrest during 1990s. After the bloodless Rose Revolution, however, the new leadership has established efficient government institutions^[5], reformed the economy and guided the country through a period of the fastest economic growth in its history.

Georgia is a representative democracy, organized as a secular, unitary, presidential republic; however the idea to restore the constitutional monarchy is popular in certain circles, most notably in the Georgian Orthodox Church.^[10] It is currently a member of the United Nations, the Council of Europe, the World Trade Organization, the Organization of the Black Sea Economic Cooperation, and GUAM Organization for Democracy and Economic Development. The country seeks to join NATO and, in the longer term, accession to the European Union.^[6]

In August 2008, Georgia engaged in an armed conflict with Russia and separatist groups from South Ossetia and Abkhazia. In the aftermath of the conflict Russia recognized Georgian regions of South Ossetia and Abkhazia as independent states, but the West insists that they are still Osetia "Russia-occupied".^[7] As of 2009, the Parliament of Georgia has passed a resolution declaring Abkhazia and South Ossetia as independent states.

Georgia (U.S. state) - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo RealPlayer Windows Marketplace Windows Media Windows Yahoo! Answers http://timsach.com.vn... Yahoo! Downloads

Dai hoc Ton Duc Thang - Ton Duc Than... W Georgia (U.S. state) - Wikipedia, ... W WikipediaRedirect - Wikipedia, the free en... Ke hoach trinh hieu truong - Span - Ya...

Log in / create account

Georgia (U.S. state)

From Wikipedia, the free encyclopedia

Coordinates: 33° -83.5

This article is about the U.S. state. For the country, see [Georgia \(country\)](#). For other uses, see [Georgia \(disambiguation\)](#).

The [State of Georgia](#) ([dʒɔdʒə] (help·info)) is a state in the United States and was one of the original Thirteen Colonies that revolted against British rule in the American Revolution. It was the last of the Thirteen Colonies to be established, in 1733. It was the fourth state to ratify the [United States Constitution](#), on January 2, 1788. It seceded from the Union on January 21, 1861 and was one of the original seven [Confederate States](#). It was the last state to be readmitted to the Union, on July 15, 1870. Georgia is the ninth-largest state in the nation by population, with an estimated 9,544,750 residents as of July 1, 2007.^[1] It is also the third-fastest-growing state in terms of numeric gain and fifth in terms of percent gain, adding 202,670 residents at a rate of 2.2 percent. From 2006 to 2007, Georgia had 18 counties among the nation's 100 fastest-growing counties, the most of any state. Georgia is also known as the *Peach State* and the *Empire State of the South*. Atlanta is the most populous city, and the capital.

Georgia is bordered on the south by [Florida](#); on the east by the [Atlantic Ocean](#) and [South Carolina](#); on the west by [Alabama](#) and by [Florida](#) in the extreme southwest; and on the north by [Tennessee](#) and [North Carolina](#). The northern part of the state is in the [Blue Ridge Mountains](#), a mountain range in the vast mountain system of the [Appalachians](#). The central piedmont extends from the foothills to the fall line, where the rivers cascade down in elevation to the continental coastal plain of the southern part of the state. The highest point in Georgia is [Brasstown Bald](#), 4,784 feet (1,458 m); the lowest point is sea level.

With an area of 59,424 square miles (153,909 km²), Georgia is ranked 24th in size among the 50 U.S. states. Georgia is the largest state east of the [Mississippi River](#) in terms of land area, although it is the fourth largest (after [Michigan](#), [Florida](#), and [Wisconsin](#)) in total area, a term which includes expanses of water claimed as state territory.^[4]

Contents [hide]

- 1 Geography
- 1.1 Boundaries
- 1.2 Geology and terrain
- 1.3 Flora and fauna
- 1.4 Climate
- 1.5 Natural resources

Flag of Georgia Seal of Georgia

Nickname(s): Peach State, Empire State of the South Motto(s): Wisdom, Justice, and Moderation

Map of the United States showing the location of Georgia.

Official language(s)	English
Demonym	Georgian
Capital	Atlanta
Largest city	Atlanta
Largest metro area	Atlanta metro area

He is seen as a national hero by those who live in **Georgia** after the independence from the USSR.

Importance of NE Discrimination

- Queries about NEs constitute significant portion of Web queries:
 - 11-17% contain person name*
 - 4% are about a person name*
- Ideally, search results should be clustered such that each cluster corresponds to the same individual
 - faster fact extraction
 - more accurate information retrieval

* study by Javier Artiles, 2009

On the Web ...

- Nobody knows how many senses (meanings) are there for a given person name
- It is impossible to estimate and trace the most frequent sense
 - the task is time consuming and tedious for humans
 - new Web pages constantly appear
 - old Web pages might be deleted over time

Why is it called “Discrimination”?

Disambiguation

- the total number of senses is known
- the meaning of each sense is known
- the order is based on the frequency

bank



meaning 1:

the slope beside a body of water



meaning 2:

depository financial institution

...

Discrimination

- the total number of senses is unknown
- the meaning of each sense is unknown
- no specific mapping of cluster& sense

Jerry Hobbs

group1

group 2

group 3

Today

Carrot Search Results Clustering Engine
<http://search.carrot2.org/stable/search>

Jerry Hobbs

Tree Visualization

All Topics (100)

- Time Ontology (8)
- Author (5)
- David Martin (5)
- Discourse Analysis (5)
- DBLP (1)
- Jerry Branton Hobbs (4)
- Killed (4)
- Music (4)
- Information Extraction (3)
- Information Sciences Institute (3)

[more | show all](#)

Top 100 results of about 79100 for Jerry Hobbs

- 1 [Jerry R. Hobbs](#)
- Oct 14, 2003 ... **Jerry R. Hobbs**. Address: USC/ISI 4676 Admiralty Way ... **Jerry R. Hobbs** hobbs@isi.edu. USC/ISI, 4676 Admiralty Way, Marina del Rey, CA 90292. <http://www.isi.edu/~hobbs/> [Ask, Bing, Bluewin, Cuil, Entireweb, Exalead, Google, Yahoo]
- 2 [AMW | Fugitives | Jerry Hobbs | Brief](#)
- May 10, 2005 ... Fugitives | **Jerry Hobbs** - Brief - Father Denied Bail Awaits Trial For Children's Murders **Jerry Branton Hobbs** accused of the stabbing deaths ... <http://www.amw.com/fugitives/brief.cfm?id=31767> [Ask, Bing, Bluewin, Exalead, Google, Yahoo]
- 3 [Jerry Hobbs - Wikipedia, the free encyclopedia](#)
- Dr. **Jerry R. Hobbs** (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial ... http://en.wikipedia.org/wiki/Jerry_Hobbs [Ask, Bing, Bluewin, Entireweb, Exalead, Wikipedia, Yahoo, Google]
- 4 [Bluhm Blog: DNA test results suggest that Jerry Hobbs's confession ...](#)
- Nov 14, 2008 ... Who could forget this picture? Could **Jerry Hobbs** be innocent of capital murder? More than three years ago, on Mother's Day, residents of ... <http://blog.law.northwestern.edu/bluhm/2008/11/dna-test-results-suggest-that-jerry-hobbss-confession-is-false.html> [Ask, Bluewin, Exalead, Google, Yahoo]
- 5 [Alleged Killer Dad Denied Bond - CBS News](#)
- May 11, 2005 ... **Jerry Hobbs**, who is accused of killing his 9-year-old daughter and her best ... On Wednesday, a judge denied bail for **Jerry Hobbs**, 34, ... <http://www.cbsnews.com/stories/2005/05/11/national/main694398.shtml> [Ask, Bluewin, Yahoo, Google]
- 6 [Jerry Hobbs](#)
- <http://www.ai.sri.com/~hobbs/> [Bing, Cuil, Google, Yahoo]
- 7 [Jerry Hobbs News](#)
- Articles covering **Jerry Hobbs**: Lake County judge denies bail request for **Jerry Hobbs**, suspect in ... / **Jerry Hobbs** Hearing / DNA clears suspect **Jerry Hobbs**, ... <http://jerry-hobbs-news.newslib.com/> [Ask, Bluewin, Google]
- 8 [Orange Tangerine: What should be done with Jerry Hobbs](#)
- May 11, 2005 ... **Jerry Hobbs** is the rage-filled, domestic-abusing career criminal who killed his 8-year-old daughter and her 9-year-old friend, with scarcely ... <http://orangetangerine.blogspot.com/2005/05/what-should-be-done-with-jerry-hobbs.html> [Ask, Bluewin, Exalead, Google]
- 9 [DBLP: Jerry R. Hobbs](#)
- Jerry R. Hobbs**, Vladik Kreinovich: Optimal choice of granularity in commonsense estimation: Why half-orders of magnitude? Int. J. Intell. Syst. ... <http://dblp.uni-trier.de/search/author?author=Jerry%20R.%20Hobbs> [Ask, Bluewin, Google]

Today

Carrot Search <http://search.carrot-search.com/carrot2-webapp/search>

Jerry Hobbs

Tree Visualization

All Topics (100)

- Jerry R. Hobbs (25)**    
 - [Discourse Analysis \(5\)](#)
 - [David \(5\)](#)
 - [Feng \(4\)](#)
 - [Semantic \(4\)](#)
 - [DBLP \(4\)](#)
 - [Scientific Commons \(3\)](#)
 - [Srinivasa Narayanan \(3\)](#)
 - [Science \(3\)](#)
 - [Microsoft Academic Search \(2\)](#)
 - [Formal Theories of the Commonsense World \(2\)](#)
 - [more | show all](#)
- [John Hobbs \(9\)](#)
- [Murder \(9\)](#)
- [David Martin \(5\)](#)
- [Denied Bond \(4\)](#)
- [Music \(4\)](#)
- [Videos \(4\)](#)
- [Laura Hobbs \(3\)](#)
- [Facebook \(3\)](#)
- [8-year-old Daughter \(3\)](#)
- [LinkedIn \(3\)](#)
- [Public \(4\)](#)
- [Lake County Judge \(3\)](#)
- [Media \(3\)](#)
- [Roy Jerry Hobbs \(2\)](#)
- [Own Daughter \(2\)](#)
- [Mathematics Genealogy Project \(2\)](#)
- [Daml-time Mailing List Archive \(2\)](#)
- [Information Extraction \(2\)](#)
- Information Sciences Institute (2)**    
 - [Coherence \(2\)](#)
 - [Abduction \(2\)](#)
 - [Finite-state \(2\)](#)
 - [References \(2\)](#)
 - [Write Fiction \(2\)](#)
 - [Design \(2\)](#)
 - [Interpretation \(2\)](#)
 - [Other topics \(14\)](#)

Cluster Jerry R. Hobbs with 25 documents ([search for more like this](#))

- 1 [Jerry R. Hobbs](#)   

Oct 14, 2003 ... **Jerry R. Hobbs**. Address: USC/ISI 4676 Admiralty Way ... **Jerry R. Hobbs** hobbs@isi.edu. USC/ISI, 4676 Admiralty Way, Marina del Rey, CA 902
<http://www.isi.edu/~hobbs/> [Ask, Bing, Bluewin, Cuil, Entireweb, Exalead, Google, Yahoo]
- 3 [Jerry Hobbs - Wikipedia, the free encyclopedia](#)   

Dr. **Jerry R. Hobbs** (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial ...
http://en.wikipedia.org/wiki/Jerry_Hobbs [Ask, Bing, Bluewin, Entireweb, Exalead, Wikipedia, Yahoo, Google]
- 11 [DBLP: Jerry R. Hobbs](#)   

Jerry R. Hobbs, Vladik Kreinovich: Optimal choice of granularity in commonsense estimation: Why half-orders of magnitude? Int. J. Intell. Syst. ...
<http://dblp.uni-trier.de/search/author?author=Jerry%20R.%20Hobbs> [Ask, Bluewin, Google]
- 14 [Publications in Discourse Analysis](#)   

Hobbs, Jerry R., and Rutu Mukar-Mehta, 2008. ``Using Abduction for Video-Text Coreference'', Proceedings, BOEMIE Workshop, Koblenz, Germany, December 20
<http://www.isi.edu/~hobbs/discourse-references/discourse-references.html> [Ask, Bluewin]
- 15 [Feng Pan's Homepage - Feng Pan](#)   

Feng Pan and **Jerry R. Hobbs**. 2006. "Temporal Arithmetic Mixing Months and ... **Jerry R. Hobbs** and Feng Pan. 2004. "An Ontology of Time for the Semantic Web
<http://fengpan.net/> [Ask, Bluewin, Google]
- 18 [Time Ontology in OWL](#)   

Jerry R. Hobbs, Feng Pan. **Jerry R. Hobbs**, Feng Pan, Time Ontology in OWL, World Wide Web Consortium, Working Draft WD-owl-time-20060927, September 21
<http://dret.net/biblio/reference/owltime> [Ask, Cuil]
- 22 [Guest Speaker: Jerry R. Hobbs](#)   

Bio: Dr. **Jerry R. Hobbs** is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial intelligence. ...
<http://www.cs.umn.edu/medlang/speakers/hobbs.html> [Ask, Bluewin]
- 23 [AI Center :: People](#)   

SRI International's Artificial Intelligence Center (AIC) is one of the world's ... Pan, Feng and **Hobbs, Jerry R.** Temporal Aggregates in OWL-Time, in Proceedings, ...
<http://www.ai.sri.com/people/hobbs> [Bing, Yahoo]
- 29 [DBLP: Jerry R. Hobbs](#)   

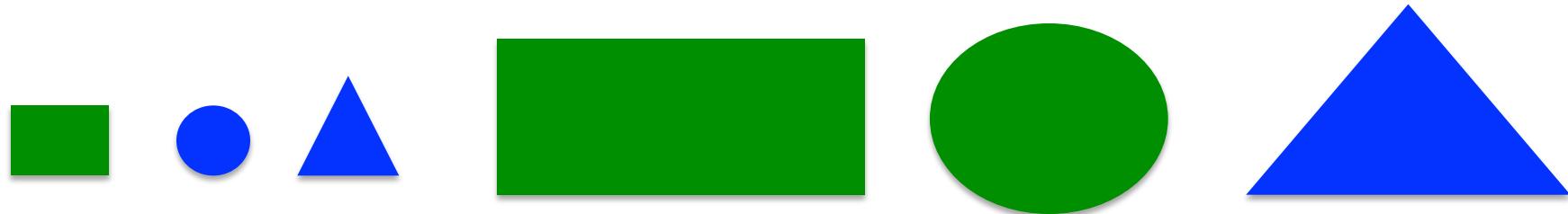
Coauthor Index
http://www.informatik.uni-trier.de/~ley/db/indices/a-tree/h/Hobbs:Jerry_R=.html [Bing]
- 34 [Scientific Commons: Jerry R. Hobbs](#)   

Clustering

Unsupervised Learning

- Labels must be automatically discovered
- Method: clustering
- Tasks which can be resolved: Named Entity Disambiguation, any type of Text Classification task

Clustering



- Are there any “groups” in the data ?
- What is each group ?
- How many ?
- How to identify them?

What is Clustering?

- Clustering is the process of grouping a set of objects into classes of similar objects
- A common and important task that finds many applications in Science, Engineering among others
 - group genes that perform the same function
 - group individuals that has similar political view
 - identify similar objects from pictures
 - categorize documents of similar topics
 - disambiguate named entities (in our case)

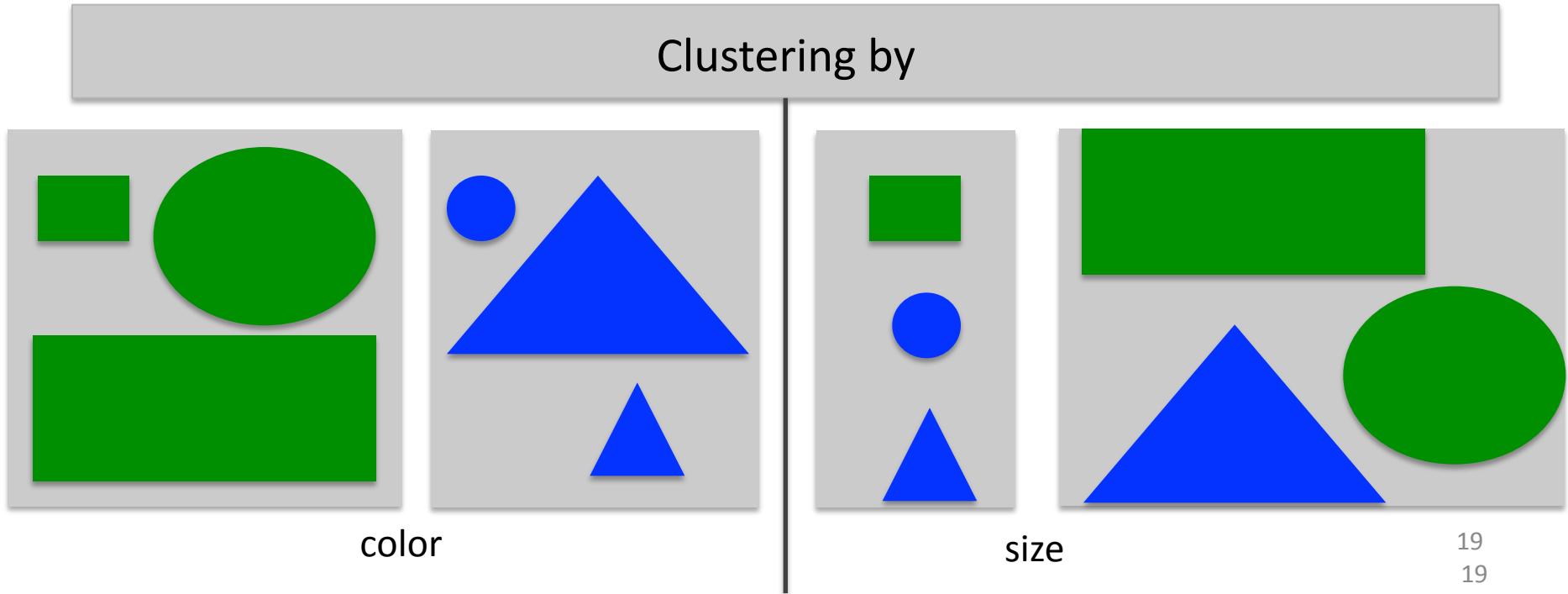
Issues in Clustering

- What is a natural grouping among these objects?
- What makes objects “related”?
 - “similarity/distance” metric
- Representation for objects
 - through vector space
- How many clusters?
 - fixed a-priory
 - completely data driven
 - avoid “trivial” clusters - too large or small
- Clustering Algorithms
 - Partition algorithms
 - Hierarchical algorithms

What is the natural grouping
among these objects?



What is the natural grouping among these objects?



What is Similarity?



- Hard to define, but we know it when we see it.
- Easier to think in terms of the distance between vectors

Properties of distance measure

- $D(A,B) = D(B,A)$ *Symmetry*
- $D(A,A) = 0$ *Constancy of Self-Similarity*
- $D(A,B) = 0$ if $A = B$ *Positivity Separation*
- $D(A,B) \leq D(A,C) + D(B,C)$ *Triangular Inequality*

Properties of distance measure

- $D(A,B) = D(B,A)$ *Symmetry*
Otherwise you could claim that “Alex looks like Bob, but Bob looks nothing like Alex”
- $D(A,A) = 0$ *Constancy of Self-Similarity*
- $D(A,B) = 0$ if $A = B$ *Positivity Separation*
- $D(A,B) \leq D(A,C) + D(B,C)$ *Triangular Inequality*

Properties of distance measure

- $D(A,B) = D(B,A)$ *Symmetry*
Otherwise you could claim that “Alex looks like Bob, but Bob looks nothing like Alex”
- $D(A,A) = 0$ *Constancy of Self-Similarity*
Otherwise you could claim that “Alex looks more like Bob, than Bob does”
- $D(A,B) = 0$ if $A = B$ *Positivity Separation*
- $D(A,B) \leq D(A,C) + D(B,C)$ *Triangular Inequality*

Properties of distance measure

- $D(A,B) = D(B,A)$ *Symmetry*
Otherwise you could claim that “Alex looks like Bob, but Bob looks nothing like Alex”
- $D(A,A) = 0$ *Constancy of Self-Similarity*
Otherwise you could claim that “Alex looks more like Bob, than Bob does”
- $D(A,B) = 0$ if $A = B$ *Positivity Separation*
Otherwise there are objects in your world that are different, but you cannot tell apart
- $D(A,B) \leq D(A,C) + D(B,C)$ *Triangular Inequality*

Properties of distance measure

- $D(A,B) = D(B,A)$ *Symmetry*
Otherwise you could claim that “Alex looks like Bob, but Bob looks nothing like Alex”
- $D(A,A) = 0$ *Constancy of Self-Similarity*
Otherwise you could claim that “Alex looks more like Bob, than Bob does”
- $D(A,B) = 0$ if $A = B$ *Positivity Separation*
Otherwise there are objects in your world that are different, but you cannot tell apart
- $D(A,B) \leq D(A,C) + D(B,C)$ *Triangular Inequality*
Otherwise you could claim that “Alex is very like Bob and Alex is very like Carl, but Bob is very unlike Carl”

Distance measures

- Given two objects x and y both with n values

$$x = (x_1, x_2, \dots, x_n)$$
$$y = (y_1, y_2, \dots, y_n)$$

calculate the Euclidean distance as

$$d(x, y) = \sqrt{2} \sqrt{\sum_{i=1}^p |x_i - y_i|^2}$$

Bottom-up Clustering

- Begin with each element in a separate cluster
- Merge clusters into successively large cluster
- Repeat until one cluster is left

Recommended reading:

Chapter 14 on Clustering from the book of Manning & Schütze

Top-down Clustering

- Begin with all elements in a whole cluster
- Divide clusters into successively smaller cluster
- Repeat until all elements are in singleton clusters

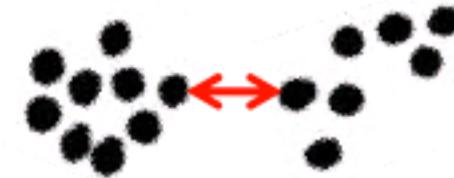
Recommended reading:

Chapter 14 on Clustering from the book of Manning & Schütze

Cluster Proximity Estimate

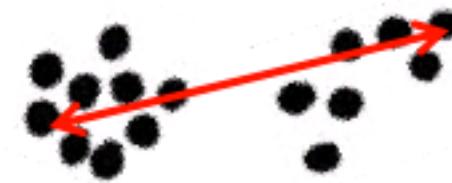
- Single-Link

- Nearest Neighbor: the closest members



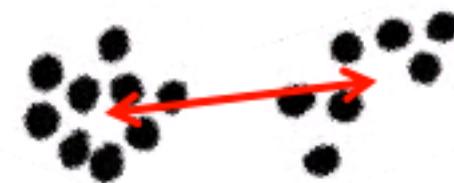
- Complete-Link

- Furthest Neighbor: the furthest members



- Centroid

- Centers of gravity



Partitioning Clustering

- Constructs a partition of n objects into a set of K clusters
- K-means algorithm:

Input: Desired number of clusters, k

Initialize: the k cluster centers (random if necessary)

Iterate:

1. Decide the class memberships of the N objects by assigning them to the nearest cluster centroids (mean)
2. Re-estimate the k clusters, by assuming the membership found above are correct

$$\vec{\mu}_k = \frac{1}{c_k} \sum_{i \in C_k} \vec{x}_i$$

Terminate:

If none of the N objects changed membership in the last iteration, exit

Back to NE Discrimination

Text Snippet

- A small fragment of text that contains from one to three sentences

Dr. **Jerry R. Hobbs** (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial

Jerry Hobbs is the rage-filled, domestic-abusing career criminal who killed his 8-year-old daughter and her 9-year-old friend, with scarcely ...

Problem Formulation

- Input:
 - N *text snippets* that mention a particular proper name (it can be person, organization or location)
- Output:
 - K clusters, where each cluster has *text snippets* that are similar to each other and different from the *snippets* in the rest of the clusters

Input

- Dr. **Jerry R. Hobbs** (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial
- **Jerry Hobbs** is the rage-filled, domestic-abusing career criminal who killed his 8-year-old daughter and her 9-year-old friend, with scarcely ...
- **Jerry Hobbs**, Author. A fifth generation farmer, Wayne Cryts finished harvesting his crop in the fall of 1980 and hauled more than 32000 bushels of soybeans ...
- **Jerry Hobbs**, who is accused of killing his 8-year-old daughter and her best ... On Wednesday, a judge denied bail for **Jerry Hobbs**, 34, ...
- Fugitives | **Jerry Hobbs** - Brief - Father Denied Bail Awaits Trial For Children s Murders Jerry Branton Hobbs accused of the stabbing deaths ...
- **Jerry R. Hobbs**. Address: USC/ISI 4676 Admiralty Way ... **Jerry R. Hobbs** hobbs@isi.edu. USC/ISI, 4676 Admiralty Way, Marina del Rey, CA 90292

Output

- **Cluster 1:**

- Dr. **Jerry R. Hobbs** (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial
- **Jerry R. Hobbs**. Address: USC/ISI 4676 Admiralty Way ... **Jerry R. Hobbs** hobbs@isi.edu. USC/ISI, 4676 Admiralty Way, Marina del Rey, CA 90292

- **Cluster 2:**

- **Jerry Hobbs** is the rage-filled, domestic-abusing career criminal who killed his 8-year-old daughter and her 9-year-old friend, with scarcely ...
- **Jerry Hobbs**, who is accused of killing his 8-year-old daughter and her best ... On Wednesday, a judge denied bail for **Jerry Hobbs**, 34, ...
- Fugitives | **Jerry Hobbs** - Brief - Father Denied Bail Awaits Trial For Children s Murders Jerry Branton Hobbs accused of the stabbing deaths ...

- **Cluster 3:**

- **Jerry Hobbs**, Author. A fifth generation farmer, Wayne Cryts finished harvesting his crop in the fall of 1980 and hauled more than 32000 bushels of soybeans

Underlying Premise*

- You shall know a word by the company it keeps
 - Firth, 1957 (Studies in Linguistic Analysis)
- Meanings of words are determined by their distributional patterns (Distributional Hypothesis)
 - Harris, 1968 (Mathematical Structures of Language)
- Words that occur in similar contexts will have similar meanings (Strong Contextual Hypothesis)
 - Miller and Charles, 1991 (Language and Cognitive Processes)

* This slide is adapted from a tutorial of Ted Pedersen

Dr. Jerry R. Hobbs is a prominent researcher in
the fields of computational linguistics

Jerry Hobbs, a fifth generation farmer,
Wayne Cryts finished harvesting his crop

Jerry Hobbs is the rage-filled, domestic-
abusing

Jerry Hobbs, who is accused of killing his 9-
year-old daughter and her best

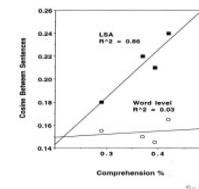
Jerry R. Hobbs. Address: USC/ISI 4676
Admiralty Way .

Collection of text snippets
containing the name of interest

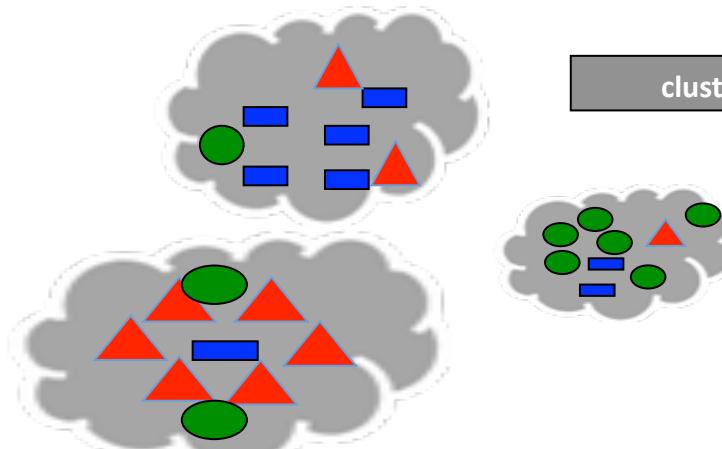
text snippet representation

	s1	s2	...	sn
teach	2	0	...	7
kill	10	2	...	3
child	1	3	...	0

snippet similarity



cluster snippets



Text Snippet Representation

- The context of each snippet is represented by a vector with k dimensions
- Each dimension indicates whether a particular feature occurred in the context
 - the value can be binary, frequency count etc.
- The features capture the characteristics of the context to be clustered
- Intuitively, vectors/contexts that share the same features will be similar to each other

Contexts (input text snippets)

- Cnt1: Dr. **Jerry R. Hobbs** (born 25 January 1942) is a prominent researcher in the fields of computational linguistics, discourse analysis, and artificial
- Cnt2: **Jerry Hobbs** is the rage-filled, domestic-abusing career criminal who killed his 8-year-old daughter and her 9-year-old friend, with scarcely ...
- Cnt3: **Jerry Hobbs**, Author. A fifth generation farmer, Wayne Cryts finished harvesting his crop in the fall of 1980 and hauled more than 32000 bushels of soybeans ...
- Cnt4: **Jerry Hobbs**, who is accused of killing his 9-year-old daughter and her best ... On Wednesday, a judge denied bail for **Jerry Hobbs**, 34, ...

Text Snippet Features (1)

- Unigram – a single word that occurs more than a given number of times

binary values

	kill	artificial	researcher	...	daughter
Cnt1:	0	1	1		0
Cnt2:	1	0	0		1
Cnt3:	0	0	0		0
Cnt4:	1	0	0		1

Text Snippet Features (1)

- Unigram – a single word that occurs more than a given number of times

• kill	1000]	frequency estimated from corpus
• artificial	500		
• researcher	200		
...			
• daughter	100		

frequency values

	kill	artificial	researcher	...	daughter
Cnt1:	0	500	200		0
Cnt2:	1000	0	0		100
Cnt3:	0	0	0		0
Cnt4:	1000	0	0		100

Text Snippet Features (2)

- Bigram – an ordered pair of words that occur together more often than expected by chance

binary values

	kill his	prominent researcher	criminal who	...	8-year-old daughter
Cnt1:	0	1	0		0
Cnt2:	1	0	1		1
Cnt3:	0	0	0		0
Cnt4:	1	0	0		1

Text Snippet Features (2)

- Bigram— an ordered pair of words that occur together more often than expected by chance

- kill his 21.2
- prominent researcher 102.9
- criminal who 68.5
- ...
- 8-year-old daughter 35.9

$\left. \begin{array}{l} \cdot \text{kill his} & 21.2 \\ \cdot \text{prominent researcher} & 102.9 \\ \cdot \text{criminal who} & 68.5 \\ \dots \\ \cdot \text{8-year-old daughter} & 35.9 \end{array} \right\}$ - $\log P(w_1 | w_0)$, log-likelihood scores based on frequency estimated from corpus

frequency weights

	kill his	prominent researcher	criminal who	...	8-year-old daughter
Cnt1:	0	102.9	0		0
Cnt2:	21.2	0	68.5		35.9
Cnt3:	0	0	0		0
Cnt4:	21.2	0	0		35.9

Text Snippet Clustering

- group text snippets by similar meaning
- snippet similarity is calculated as $sim(Cnt_1, Cnt_2) = \sum_{i=1}^n w_{1i} * w_{2i}$

	kill	artificial	researcher	daughter
Cnt1:	0	1	1	0
Cnt2:	1	0	0	1
Cnt3:	0	0	0	0
Cnt4:	1	0	0	1

$$sim(Cnt1, Cnt2) = (0*1) + (1*0) + (1*0) + (0*1) = 0$$

$$sim(Cnt1, Cnt3) = (0*0) + (1*0) + (1*0) + (0*0) = 0$$

$$sim(Cnt1, Cnt4) = (0*1) + (1*0) + (1*0) + (0*1) = 0$$

$$sim(Cnt2, Cnt3) = (1*0) + (0*0) + (0*0) + (0*0) = 0$$

$$sim(Cnt2, Cnt4) = (1*1) + (0*0) + (0*0) + (1*1) = 2$$

$$sim(Cnt3, Cnt4) = (0*1) + (0*0) + (0*0) + (0*1) = 0$$

Final Output

- Each cluster consists of a certain number of *text snippets*, i.e. small text fragments
- For each cluster assign cluster labels:
 - top 10 most significant bigrams of each cluster act as a descriptive label
 - top 10 most unique bigrams for each cluster act as discriminating label

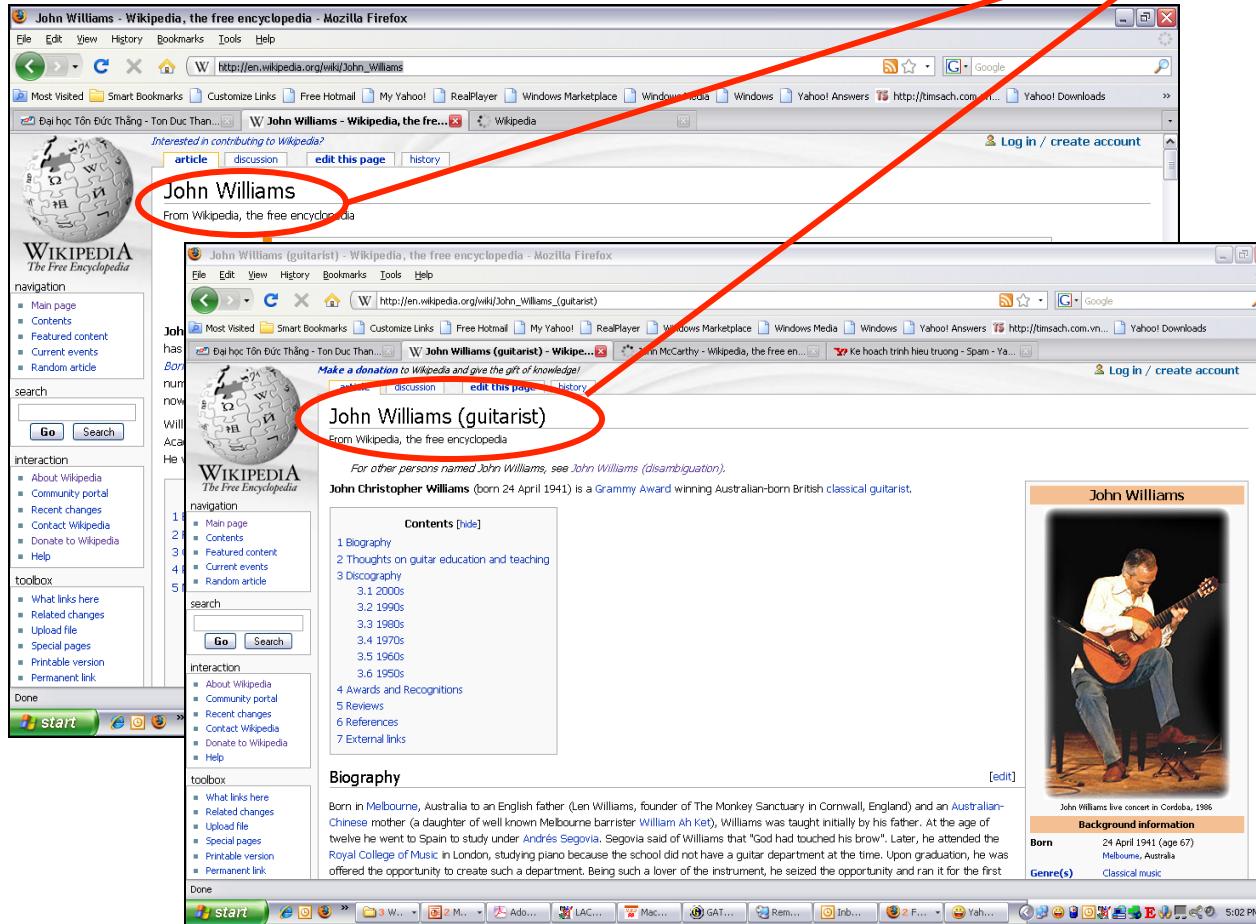
Cluster Evaluation

- Internal criterion
 - intra-class high similarity
 - inter-class low similarity
 - the quality depends on the object representation and the similarity measure used
- External criterion for clustering quality
 - measure the ability to discover the named entity groups in the gold standard data
 - assess the clustering with respect to ground truth

What other features can one use?

Ideas for more features if we use structured text like Wikipedia

Page Titles (ID)



Ideas for more features if we use structured text like Wikipedia

The image displays three separate Mozilla Firefox windows, each showing a different Wikipedia page:

- Top Window:** Shows the Wikipedia page for "Georgia (country)". The title "Georgia (country)" is circled in red.
- Middle Window:** Shows the Wikipedia page for "Oxford, Mississippi". The title "Oxford, Mississippi" is circled in red.
- Bottom Window:** Shows the Wikipedia page for "John McCarthy (computer scientist)". The title "John McCarthy (computer scientist)" is circled in red.

A large red arrow originates from the circled title in the top window and points to the circled title in the bottom window. Another red arrow originates from the circled title in the middle window and also points to the circled title in the bottom window. To the right of the bottom window, the text "Disambiguation text" is written in red.

Georgia (country) - Wikipedia, the free encyclopedia - Mozilla Firefox
http://en.wikipedia.org/wiki/Georgia_(country)

Oxford, Mississippi - Wikipedia, the free encyclopedia - Mozilla Firefox
http://en.wikipedia.org/wiki/Oxford,_Mississippi

John McCarthy (computer scientist) - Wikipedia, the free encyclopedia - Mozilla Firefox
http://en.wikipedia.org/wiki/John_McCarthy_(computer_scientist)

Ideas for more features if we use structured text like Wikipedia

The image shows a dual-monitor setup. The top monitor displays a Mozilla Firefox window with the Wikipedia page for "John McCarthy (computer scientist)". The bottom monitor also displays a Mozilla Firefox window showing the same Wikipedia page. A red arrow originates from the word "Category" in red text on the right side of the slide and points to the "Categories" section at the bottom of the page on the bottom monitor. A large red circle highlights the entire list of categories listed under the "Categories" section.

Category

Categories: 1927 births | Members of the National Academy of Sciences | Living people | Artificial intelligence researchers | American Jews | Jewish American scientists | Lisp programming language | Programming language designers | Usenet people | American computer scientists | Computer pioneers | Turing Award laureates | People from Boston, Massachusetts | Irish-Americans | National Medal of Science laureates | Formal methods people | Stanford University faculty | California Institute of Technology alumni | Fellows of the Association for Computing Machinery | History of artificial intelligence | Fellows of the American Association for Artificial Intelligence

This page was last modified on 7 October 2008, at 12:37. All text is available under the terms of the GNU Free Documentation License. (See [Copyrights](#) for details.) Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a U.S. registered 501(c)(3) tax-deductible nonprofit charity. Privacy policy About Wikipedia Disclaimers

Ideas for more features if we use structured text like Wikipedia

The image displays three separate screenshots of a Wikipedia-like website interface, each showing a redirect page. Red arrows and circles highlight specific elements to illustrate a feature proposal.

- Top Screenshot:** Shows a redirect page for "US". The title "US" is circled in red. A red arrow points from this circle to the text "Redirect page titles" on the right.
- Middle Screenshot:** Shows a redirect page for "USA". The title "USA" is circled in red. A red arrow points from this circle to the text "Redirect page titles" on the right.
- Bottom Screenshot:** Shows a redirect page for "United states of america". The title "United states of america" is circled in red.

Redirect page titles

From Wikipedia, the free encyclopedia
Redirect page

↳ United States

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

article discussion edit this page history

From Wikipedia, the free encyclopedia
Redirect page

↳ United States

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

article discussion edit this page history

From Wikipedia, the free encyclopedia
Redirect page

↳ United States

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

article discussion edit this page history

Ideas for more features if we use structured text like Wikipedia

Hyperlinks

John McCarthy (computer scientist) - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W http://en.wikipedia.org/wiki/John_McCarthy_(computer_scientist)

Most Visited Smart Bookmarks Customize Links Free Hotmail My Yahoo! RealPlayer Windows Marketplace Windows Media Windows Yahoo! Answers http://timsach.com.vn... Yahoo! Downloads

Đại học Tân Đức Thắng - Ton Duc Than... W Oxford, Mississippi - Wikipedia, the free encyclopedia W John McCarthy (computer scientist) Kế hoạch trình hiệu trưởng - Spam - Ya...

Log in / create account

Make a donation to Wikipedia and give the gift of knowledge!

article discussion edit this page history

John McCarthy (computer scientist)

From Wikipedia, the free encyclopedia

John McCarthy (born September 4, 1927, in Boston, Massachusetts), is an American computer scientist and cognitive scientist who received the Turing Award in 1971 for his major contributions to the field of Artificial Intelligence [1]. He was responsible for the coinage of the term "Artificial Intelligence" in his 1955 proposal for the 1956 Dartmouth Conference and is the inventor of the Lisp programming language.

Contents [hide]

- 1 Life
- 2 Work
- 3 Major Publications
- 4 See also
- 5 References
- 6 Further reading
- 7 External links

Life

[edit]

John McCarthy was born in Boston on September 4, 1927 to two immigrants, John Patrick and Ida Glatt McCarthy. The family was forced to move frequently during the depression, until McCarthy's father found work as an organizer for the Amalgamated Clothing Workers in Los Angeles, California. McCarthy showed an early aptitude for mathematics; in his teens he taught himself mathematics by studying the textbooks used at the nearby California Institute of Technology (Caltech). As a result, when he was accepted into Caltech the following year, he was able to skip the first two years of mathematics. [2]

Receiving a B.S. in Mathematics in 1948, McCarthy initially continued his studies at Caltech. He received a Ph.D. in Mathematics from Princeton University in 1951 under Solomon Lefschetz. After short-term appointments at Princeton, Stanford, Dartmouth, and MIT, he became a full professor at Stanford in 1962, where he remained until his retirement at the end of 2000. He is now a Professor Emeritus.

Doctoral students

- Ruzena Rajcsy
- Randall Davis
- Claude Green
- Barbara Liskov
- Robert Moore
- Francis Morris
- Daniel Padua

Web People Search Challenge

- The first challenge was organized in 2007
- WePS focuses on person and organization name disambiguation of Web pages
- For each ambiguous name, the system must return the documents and the attributes which are relevant for the different senses of the name
- Last such challenge was on 1st of July 2010
- More information at: <http://nlp.uned.es/weps/>

Name Discrimination Demo

- SenseClusters by Ted Pedersen

<http://marimba.d.umn.edu/cgi-bin/SC-cgi/index.cgi>

- The software can be used for:
 - proper name discrimination
 - word sense discrimination
 - e-mail clustering
 - synonym finding

What would you do with clustering?

What would you do with clustering?

- E-mail spam classification
- Organize the documents into multiple categories like Google news
- Trace what two people would talk about
- Sentiment analysis
- Tweet trend identification
- Tweet topic detection
- ...