

CSC421 Natural Language Processing

Instructor & Textbooks

- Jung-jae Kim
 - ◆ Email: jungjae.kim@ntu.edu.sg
 - ◆ Office N4 2b 44
 - ◆ <http://www.ntu.edu.sg/home/jungjae.kim/>
 - ◆ Natural language processing, text mining, bioinformatics, sentiment analysis, Semantic Web, ...

Outline for today

- Course expectation & Evaluation
- What is NLP?/ Why take NLP course?
 - ◆ Example application

One Sentence about NLP

- Teaches key theory and methods for NLP: word-level analysis, parsing, information extraction, machine translation, semantics, etc.
 - ♦ Learn techniques which can be used in **practical**, robust systems that can (partly) understand human language
- This is not a language course
 - ♦ **Computational methods** of processing natural languages

Preparation

- Pre-requisite
 - ◆ Basic algorithm and data structure analysis, e.g., dynamic programming
 - ◆ Basic probability concepts, e.g., conditional probability
 - ◆ Ability to program
 - ◆ Some exposure to basic computational logic, e.g., first-order logic
- Not everyone has the same skills
 - ◆ Assumes some ability to learn missing knowledge

Course Evaluation

- About 35% project
- About 15% mid-term quiz
 - ♦ Submit MC if you cannot make it
- About 50% final exam
- Objective: spread evaluation over the whole course, not just one exam or one report

Our Expectations

- You are final year students; you are willing to learn NLP
- We expect you to participate.
- We expect you will read
 - ◆ Lecture slides for reference only
 - ◆ Read J&M book and other references
 - ◆ Enjoy project and programming!
- We expect to get your feedbacks

Goals of the field of NLP

- Computers would be a lot more useful if they could handle our email, do our library research, chat to us...
- But someone has to work on the hard problems!
 - ◆ How can we tell computers about language?
 - ◆ Or help them learn it as kids do?
- In this course we seek to identify many of the open research problems in natural language processing

What/where is NLP?

- Goals can be very far reaching ...
 - ♦ True text understanding
 - ♦ Real-time participation in spoken dialogs
- Or very down-to-earth ... (the Web, business documents)
 - ♦ Finding the price of products on the web
 - ♦ Sentiment detection about products or stocks
 - ♦ Extracting facts or relations from documents
- These days, the latter predominate (as NLP becomes increasingly practical, it is increasingly engineering oriented)

Commercial World

Lots of exciting stuff going on in industry

Google™

YAHOO!®

Microsoft®

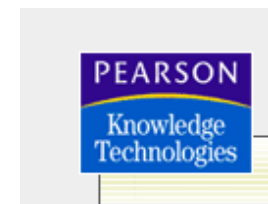
Powerset



THOMSON REUTERS



Nielsen
BuzzMetrics



Example of down-to-earth Applications

- some deployed applications
 - ◆ Machine translation: Chinese < == > English
 - ◆ Question answering: Yahoo! Answer, Baidu Zhidao
 - ◆ Information extraction: Extracting product information from the Web
 - ◆ Text analytics: Sentiment Analysis

Google Translate

الجزيرة نت

http://www.aljazeera.net/NR/exeres/8FD54E7F-56C5-49A0-B60A-89A67426F3B3.htm

Speech and ...of Contents Book Schedule University o...uter Science James Marti... Home Page The Daily Camera The

نحضر لك الأخبار الساخنة أينما تكون
اشترك الآن

ملخص الخبر

انتقل الأمين العام للأمم المتحدة الذي يزور لبنان لمقر قوة المراقبة الدولية ببلدة الناقورة الجنوبية حيث وضع إكليلًا من الزهور على نصب لاربعة مراقبين قتلوا خلال الحرب الأخيرة. وفيما ينتظر تفقد أنان للقرى المدمرة تنظم إيطاليا اليوم احتفالًا لتوديع قواتها المرسلة إلى لبنان.

أنا في الناقورة لتفقد الخط الأزرق والمناطق المدمرة

29/8/2006 م (آخر تحديث) الساعة 17:44 (مكة المكرمة)، 14:44 (غرينتش)

بدر امتثال وزير المالية الفلسطينية و13 ثلثًا من حماس

استشهاد فلسطينيين وإصابة تسعة في غارات بالصفقة والقطاع

أصيب تسعة فلسطينيين بينهم مدنيون في غارة جوية إسرائيلية على حي الشجاعية في قطاع غزة. يأتي ذلك مباشرة بعد استشهاد قائدتين بارزتين من كتائب شهداء الأقصى في عملية لقوات الاحتلال الإسرائيلي نفذها سلاح الجو وقوات المشاة في مخيم بلاطة بالضفة الغربية.

البشير يلتقي فريزر ومجلس الأمن لن يفرض قوات بدارفور

من المقرر أن يلتقي الرئيس السوداني عمر البشير جينداي فريزر مساعد وزير الخارجية الأمريكية التي تحاول في الخرطوم إقناع المسؤولين السودانيين بنشر قوة أممية بدارفور. من جانبه قال السفير الأمريكي في الأمم المتحدة إنه لا نية لمجلس الأمن بفرض قوات في الإقليم.

رمسفيلد وتشيني يصران على إبقاء القوات الأميركية بالعراق

دعا وزير الدفاع الأمريكي دونالد رمسفيلد الأميركيين إلى التحلي بالصبر بخصوص العراق. وانتقد ديك تشيني نائب الرئيس دعوات الديمقراطيين لسحب القوات الأميركية من العراق للربط بين الانسحاب المبكر واحتمال وقوع هجمات داخل الولايات المتحدة.

مقتل مدنيين وإصابة ضابط بهجوم انتحاري في أفغانستان

أعلنت القوة الدولية للمساعدة على إرساء الأمن (إيساف) مقتل مدنيين وإصابة ضابط أفغاني في هجوم استهدف قافلة لقوات الأطلسي بجنوب أفغانستان. وفي العاصمة كابل انفجرت قنبلة يدوية الصنع لدى مرور

الصفحة برعاية

المتفوقون الحقيقيون لا يلحقون بالآخرين بل يقودون الطريق

القطرية QATAR AIRWAYS

اسم البرنامج: بلا حدود

عنوان الحلقة: يوسف ندا.. اتهامات بدعم الإرهاب

الأربعاء/ مباشر

19:05 غرينتش 22:05 مكة

frontiers@aljazeera.net

التفصيلية الاقتصادية الرياضية

Killing Palestinians and wounding nine in the raids Sector

Nine Palestinians were wounded among civilians in an Israeli air raid in the neighborhood result in the Gaza Strip. This comes immediately after the killing of two prominent Al-Aqsa Martyrs Brigades in the Israeli occupying forces carried out air and infantry forces in the Balata camp in the West Bank.



Bashir meets Fraser, the Security Council will not impose forces Darfur

Is scheduled to meet with Sudanese President Omar al-Bashir Jenday Fraser Assistant Minister for Foreign Affairs of the American attempt to persuade officials in Khartoum, Sudanese Darfur deployment of the nationalities. For his part, US Ambassador to the United Nations that it has no intention of the Security Council to impose its forces in the province.



Rmsfield and Cheney insist on keeping the American forces in Iraq

Called American Defense Minister Donald Rmsfield Americans to show patience on Iraq. I take Vice President Dick Cheney calls Democrats withdrawal of American forces from Iraq link and the possibility of early withdrawal of attacks inside the United States.



Killing civilians and wounding officer suicide attack in Afghanistan

The international force to help establish security (ISAF) killed civilians and the wounding of an officer in an attack against Afghan forces convoy south Atlantic Afghanistan. In the capital Kabul, a hand grenade exploded at the passage of manufacture French patrol was not reported injuries or damage.



Web Q/A



what is the population of singapore?

Search

About 73,000,000 results (0.43 seconds)

Web

Images

Maps

Videos

News

Shopping

More

Show search tools

Population, Singapore



www.google.com/publicdata

5,183,700 - 2011

Source: World Bank

[Disclaimer](#)

Singapore - Wikipedia, the free encyclopedia

en.wikipedia.org/wiki/Singapore


There are slightly over 5 million people in **Singapore**, of which 2.91 million were born locally. The **population** is highly diverse; the majority are Chinese, with ...



[Durian](#) - [History of Singapore](#) - [Battle of Singapore](#) - [Stamford Raffles](#)

Community based QnA

YAHOO! ANSWERS

[HOME](#) [BROWSE CATEGORIES](#) [MY ACTIVITY](#) [ABOUT](#)

 **Ask**

 **Answer**  **Discover**

[Home](#) > Search Results for "What is the cost of living in Singapore?"

Search

 [Advanced Search](#)

Refine search

Keyword match in

Category

Location

Question Status

Number of Answers


Date Submitted


Sort by: [Relevance](#) | [Newest](#) | [Most Answers](#)

Sponsor Results

Cost of Living Calculator
Side-by-Side City Comparison of Categories: Food, Taxes, Housing
[Bankrate.com/CostOfLiving](#)

Singapore Job Search
Post Your Resume 100% Free. And Get Hired.
[SingaporeJobs77.com](#)

 **singapore living cost for family of three?**
...relocate to **singapore** from india in september. my family consist of me, my wife and small baby boy. my job place is 180 Cecil Street #14-01 Bangkok Bank Building **Singapore** 069546. kindly suggest me..., food all the exp for living. and what should be minimum salary to live...
☆ In Singapore - Asked by deepak - 2 answers - 2 years ago

 **What is the cost of living in Singapore and reasonable salary for 9yrs IT experience?**
... for a job opportunity in **Singapore**. I want to know the cost of living there and reasonable package that I...
2 ☆ In Singapore - Asked by giriishere - 3 answers - 3 years ago

Extracting Job Openings from the Web

OPUS International, Inc., an executive search firm focusing on the Food Science industry. - Microsoft Internet Explorer

File Edit View Favorites

Back Forward Stop

Address http://www.foodscience.com/jobs_midwest.html

Links AMEX Rewards Time DogHouse

Welcome

About OPUS

Executive Staff

Job Listings

Résumé Form

Job Hunt Hints

Academic Links

Science Fair Help

Industry Assocs.

FAQs

Contact Us

Site Map

e-mail

OPUS INTERNATIONAL

About | Staff | Job

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home

Address http://www.foodscience.com/jobs_midwest.html

Links AMEX Rewards Time DogHouse

Welcome

About OPUS

Executive Staff

Job Listings

Résumé Form

Job Hunt Hints

Academic Links

Science Fair Help

Industry Assocs.

FAQs

Contact Us

Site Map

e-mail

Test Kit Consumables

Major food products for Chicago food processors, recipes, marketing a cross-section of BS in Food Science, a minimum of 2 years experience. Contact Susan at 1-800-488-2611

Ice Cream Guru

If you dream of cold creamy chocolate or gooey, gooey cookie, there's a great opportunity for you to maintain and expand this major corporation's high-end ice cream brand. Will be based in the Upper Midwest for about a year. After that, California here I come! Requires a BS in Food Science or dairy, plus ice cream formulation experience. Will consider entry level with an M.S. and an internship. Contact Susan at 1-800-488-2611

foodscience.com-Job2

JobTitle: Ice Cream Guru

Employer: foodscience.com

JobCategory: Travel/Hospitality

JobFunction: Food Services

JobLocation: Upper Midwest

Contact Phone: 800-488-2611

DateExtracted: January 8, 2001

Source: www.foodscience.com/jobs_midwest.html

OtherCompanyJobs: foodscience.com-Job1

Job Openings:

Category = Food Services

Keyword = Baker

Location = Continental U.S.


FlipDog
Fetch Your Next Job Here™

[Home](#)
[Find Jobs](#)
[Your Account](#)
[Resource Center](#)

[Return to Results](#) | [Modify Search](#) | [New Search](#)


The University Alliance
A RISK EDUCATION NETWORK
Degrees Online

Learn While You Earn
MBA, BA, AA Degrees
Online & Project Mgt.

[Click here to e-mail your resume to 1000's of Head Hunters with ResumeZapper.com](#)



how to easily DOUBLE your chances when applying FOR JOBS!

Breakthrough ebook
shows why most people are WRONG about how to apply for jobs.

1 - 25 of 47 jobs shown below

[GO!](#)
[Search tips](#)

Show Jobs Posted:

[View: Brief](#) | [Detailed](#)

Web Jobs: FlipDog technology has found these jobs on thousands of employer Web sites.

Food Pantry Workers at Lutheran Social Services	October 11, 2002	Archbold, OH
Cooks at Lutheran Social Services	October 11, 2002	Archbold, OH
Bakers Assistants at Fine Catering by Russell Morin	October 11, 2002	Attleboro, MA
Baker's Helper at Bird-in-Hand	October 11, 2002	United States
Assistant Baker at Gourmet To Go	October 11, 2002	Maryland Heights, MO
Host/Hostess at Sharis Restaurants	October 10, 2002	Beaverton, OR
Cooks at Alta's Rustler Lodge	October 10, 2002	Alta, UT
Line Attendant at Sun Valley Coporation	October 10, 2002	Huntsville, UT
Food Service Worker II at Garden Grove Unified School District	October 10, 2002	Garden Grove, CA
Night Cook / Baker at SONOCO	October 10, 2002	Houma, LA
Cooks/Prep Cooks at GrandView Lodge	October 10, 2002	Nisswa, MN
Line Cook at Lone Mountain Ranch	October 10, 2002	Big Sky, MT
Production Baker at Whole Foods Market	October 08, 2002	Willowbrook, IL
Cake Decorator/Baker at Mandalay Bay Hotel and Casino	October 08, 2002	Las Vegas, NV
Shift Supervisors at Brueggers Bagels	October 08, 2002	Minneapolis, MN

Text Analytics

- Data-mining of weblogs, microblogs, discussion forums, message boards, user groups, and other forms of user generated media
 - ◆ Product marketing information
 - ◆ Political opinion tracking
 - ◆ Social network analysis
 - ◆ Buzz analysis (what's hot, what topics are people talking about right now).



Live Search

iphone



Products

See also: [Web](#), [Images](#), [Video](#), [News](#), [Maps](#), [More](#) ▼

Apple iPhone Smart Phone

\$499 - \$849 [Compare prices](#) (6)★★★★☆ [User reviews](#) (1233)★★★★☆ [Expert reviews](#) (1)

Quad Band - GSM 800, GSM 900, GSM 1800, GSM 1900 - Bluetooth, Wi-Fi - EDGE, GPRS - Polyphonic - 16.7 Million Colors - 4GB - Bar - Smartphone

[User reviews](#) | [Product details](#) | [Expert reviews](#) | [Compare prices](#)[All user reviews](#)

General Comments (497 comments)

85% positive

[Ease Of Use](#) (217 comments)

93% positive

[Features](#) (132 comments)

93% positive

[Affordability](#) (126 comments)

52% positive

[Screen](#) (70 comments)

80% positive

[Appearance](#) (64 comments)

95% positive

General Comments

View by: **Positive comments** (85%) | [Negative comments](#) (15%)The Iphone is the best cell phone I had so far and i think that if you buy it you will love it. [More...](#)ruben_93 [catalog.ebay.com](#) 8/23/2008I have decided to buy it because I like it very much and good product the apple iphone 16GB, has very good quality [More...](#)davidxerach [catalog.ebay.com](#) 5/30/2008We are very happy with the item, it is really a great product. [More...](#)3703guns [catalog.ebay.com](#) 5/7/2008Overall this is great item and I highly recommend this device to travelers. [More...](#)clink1381 [catalog.ebay.com](#) 4/26/2008

Sponsored sites

[iPhone 3G - Official Site](#)

Watch the new **iPhone** TV ads featuring the App store today.[www.apple.com/iphone](#)

[Save on iPhone](#)

Shop for **iPhone** and save at Live Search cashback![Search.Live.com/cashback](#)

Live Search cashback

[Iphone' at Amazon](#)

Low prices on **Iphone'**. Qualified orders over \$25 ship free[Amazon.com](#)

[iphone](#)

Find Low Prices and Multiple Offers On **iphone**[shopping.yahoo.com](#)

[By Iphone](#)

Explore 5,000+ Cell Phones & Gear. Save On By **Iphone!**[CellPhones.Shopzilla.com](#)[See your message here...](#)

The hidden structure of language

- We're going beneath the surface...
 - ◆ Not just string processing
 - ◆ Not just keyword matching in a search engine
 - ◆ Search Google on "tennis racquet" and "tennis racquets" or "laptop" and "notebook" and the results are quite different ... though these days Google does lots of subtle stuff beyond keyword matching itself
- We want to recover and manipulate at least some aspects of language structure and meaning

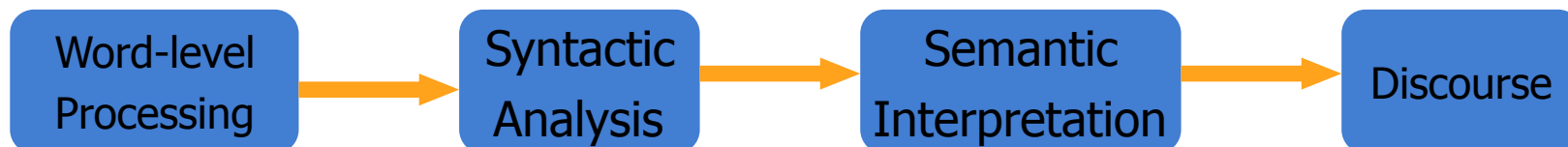
Major Topics to cover

Methodology

- Word-level processing:
 - ♦ Regular expression, Spelling correction, segmentation
 - ♦ Language models, POS tagging
- Syntax: knowledge of structural relation between words
 - ♦ Parsing
- Semantics: Meaning
- Discourse: context

Applications:

- ♦ Information extraction, Named entity recognition
- ♦ Sentiment Analysis
- ♦ Machine translation



Example tasks (1)

- Word-level processing
 - ◆ Task 1: Locate all verbs and verbs only
 - E.g. the tower collapsed as a **result** of safety violations (*Is 'result' here a noun or a verb?*)
- Syntactic processing
 - ◆ Task 2: Answer “Who killed John?”
 - E.g. “Mary killed John.”
 - E.g. “John was killed by Mary.”
 - E.g. “The guy who loved Mary killed John.”
 - E.g. “Mary is not sure of who killed John.”
 - *Hint: find subject of 'killed' whose object is 'John'*

Example tasks (2)

- Semantic processing
 - ♦ Task 3: Answer “Who killed John?”
 - E.g. Mary **assassinated** John.
 - ♦ Task 4: Answer “Who snores?”
 - E.g. **Everyone who** smokes snores and **John** smokes.
- Discourse analysis
 - ♦ Task 5: Answer “Who killed John?”
 - E.g. Mary threw **John** into sea. **He** drowned.

Learning Objective

- To learn the basic concepts and techniques of natural language processing on the levels of **words, syntax, and semantics**.
- To apply the techniques to **real-world problems** and evaluate the applications.

Caveat

Why NLP is difficult? NLP has an **AI** aspect to it.

- ◆ The language is hugely **ambiguous**
- ◆ We don't often come up with exact solutions/ algorithms

Exercise: Ambiguity

- Find at least 5 meanings of this sentence:
 - ♦ I made her duck

Ambiguity is Pervasive

- “duck” (**Lexical category**)
 - ♦ can be a noun or verb
- “her” (**Lexical category**)
 - ♦ can be a possessive (“of her”) or dative (“for her”) pronoun
- “make” (**Lexical semantics**)
 - ♦ can mean “create” or “cook”, and about 100 other things as well

More about Projects

- Group based project: 4-5 members in a group
- You can choose among given topics, and also can propose your own topics. **keep it simple**
- Project guidelines (wk-3)
- What you need to deliver:
 - ♦ Proposal (End of wk-7) – not evaluated
 - ♦ Final report (End of wk-12)
 - ♦ Final presentation (Presentation wk-13)
- Evaluation: Two options:
 - Group based– similar contribution
 - Individual– if you wish

Textbook

- SPEECH and LANGUAGE ROCESSING by [Daniel Jurafsky](#) and [James H. Martin](#), Prentice-Hall 2009
 - ♦ A copy is kept in Lee Wee Nam Library, in the reserve section
- Recommended reading
 - ♦ Pierre M. Nugues. *An Introduction to Language Processing with Perl and Prolog*, Springer; 2006.
 - available as e-version
 - ♦ Christopher D. Manning and Hinrich Schütze. *Foundations of Statistical Natural Language Processing*. 1st Edition, MIT Press, 1999.

Course Web Page

The course web page can be found at
edventure

It will have the lecture notes,
announcements, etc.

- ◆ We will not put everything on slides. We will leave out some content on purpose
- ◆ Slides cannot replace the textbook. They are at most a guideline.