Form 2B

City University of Hong Kong

Information on a Course
offered by Department of Linguistics and Translation
with effect from Semester A in 2014 / 2015

Part I

Course Title: Computational Linguistics

Course Code: LT5411

Course Duration: One semester

Credit Units: 3

Level: P5

Medium of Instruction: English

Prerequisites: Nil

Precursors: Nil

Equivalent Courses: CTL5411 Computational Linguistics

Exclusive Courses: Nil

Part II

Course Aims

This course aims at introducing students with various academic backgrounds to the quantitative aspects of language and the basic concepts of the application of computational techniques in language processing. The course will cover topics selected from a wide range of issues and challenges of the field, including but not limited to multilingual computing, using computers in linguistic studies, natural language processing, and the use of large corpora in natural language applications. This course will also provide students with a foundation for further research in relevant areas.

Course Intended Learning Outcomes (CILOs)
Upon successful completion of this course, students should be able to:

No.	CILOs	Weighting (if applicable)
1.	Identify key computational linguistic concepts and issues in major areas of study in the field.	30%
2.	Explain the theoretical basis underlying major computational approaches to handling these issues, with particular reference to English and Chinese.	40%
3.	Review the development of one or more major area of research in computational linguistics, and devise innovative solutions to the critical issues.	30%

Teaching and Learning Activities (TLAs)
(Indicative of likely activities and tasks designed to facilitate students' achievement of the CILOs. Final details will be provided to students in their first week of attendance in this course)

CILO No.	TLAs	Hours/week (if applicable)
1, 2	Lectures to explain the major issues and concepts in computational linguistics and introduce different approaches to their solution. Interaction between instructor and students is expected.	2
1, 2	Instructor-facilitated tutorial discussions on the theories and algorithms, and/or hands-on practical exercises involving the use of computational tools for various language processing tasks, to reinforce the concepts covered in lectures.	1
1, 2, 3	Independent studies on a selected topic in a major area of research in computational linguistics	
1, 2, 3	Class presentation on a selected topic in a major area of research in computational linguistics to encourage peer learning, with comments from instructor.	

Assessment Tasks/Activities
(Indicative of likely activities and tasks designed to assess how well the students achieve the CILOs. Final details will be provided to students in their first week of attendance in this course)

CILO No.	Type of Assessment Tasks/Activities	Weighting (if applicable)	Remarks
1, 2	Class participation, discussion and practical exercises on computational techniques used in language processing	30%	
1, 2	Quiz(zes) on concepts and issues in computational linguistics, and theoretical basis underlying some computational approaches to their solutions	30%	
1, 2, 3	• Independent studies on a selected topic in a major area of research in computational linguistics • Term essay and class presentation (3,000 – 4,000 words and 15 minutes for presentation)	40%	

Grading of Student Achievement:
Refer to Grading of Courses in the Academic Regulations for Taught Postgraduate Degrees.

Grading pattern: Standard (A+, A, A-...F). Grading is based on student performance in the assessment tasks/activities.

Excellent A+ A A-	Good B+ B B-	Adequate C+ C C-	Marginal D
1. Excellent knowledge of major issues and approaches in NLP. 2. Excellent review and innovative analysis on a selected topic. 3. Very active participation.	1. Good knowledge of major issues and approaches in NLP. 2. Good review and critical analysis on a selected topic. 3. Active participation.	1. Adequate knowledge of major issues and approaches in NLP. 2. Fair review and adequate analysis on a selected topic. 3. Adequate participation.	1. Basic familiarity with the subject matter. 2. Marginal ability to review and analyse a selected topic. 3. Marginal participation.

Part III

Keyword Syllabus

Natural language processing: Tokenisation, Morphological analysis, Part-of-speech tagging, Context-free rules, Parsing, Semantic representation, Disambiguation, Rule-based methods, Corpus-based methods, Statistical methods

Linguistic computing: Frequency counts, Quantitative methods in linguistic studies, Linguistic corpora, Text markup, Corpus annotation, Concordance

Natural language applications: Machine translation, Information retrieval, Information extraction, Natural language generation

Multilingual computing: Character encoding, Input and display, Internationalization and localization

Recommended Reading
Text(s)

Allen, J. (1995) *Natural Language Understanding*. Redwood City, CA: Benjamin/Cummings.

Grishman, R. (1986) *Computational Linguistics: An Introduction*. Cambridge, UK: Cambridge University Press.

Indurkhya, N. and Damerau, F.J. (2010) *Handbook of Natural Language Processing*. Boca Raton, FL: Chapman & Hall.

Jurafsky, D. and Martin, J.H. (2009) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (2nd Edition).* Upper Saddle River, NJ: Pearson Prentice Hall.

Manning, C.D. and Schutze, H. (1999) *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press.

McEnery, T. and Wilson, A. (1996) *Corpus Linguistics*. Edinburgh, UK: Edinburgh University Press.

Mitkov, R. (2003) *Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press.

黃昌寧、李涓子 (2002) 《語料庫語言學》 北京：商務印書館

俞士汶 (2003) 《計算語言學概論》 北京：商務印書館

Online Resources

Journal and conference papers at ACL Anthology

http://www.aclweb.org/anthology-index/