

Subject Description Form

Subject Code	COMP 5412
Subject Title	Fundamentals of Chinese Computing
Credit Value	3
Level	5
Pre-requisite/Exclusion	Nil
Objectives	<p>The objectives of this subject are to enable students to acquire:</p> <ol style="list-style-type: none"> 1. the basic understanding of Chinese computing; 2. the basic skill in operating within a Chinese computing environment; 3. the basic vocabulary for effective communication when computer representation and processing of Chinese is involved; 4. the basic programming skill for Chinese computing; 5. the basic skill to deal with algorithmic problems posed by the large Chinese alphabet; 6. the basic knowledge for the strategic design of software for the different language communities.
Intended Learning Outcomes	<p>Upon completion of the subject, students will be able to:</p> <ol style="list-style-type: none"> a) be aware of the different Chinese character sets and their encoding schemes; b) identify and convert the character codes generated by one encoding scheme to another; c) demonstrate the ability to design and implement internationalized software and localization components; and d) design and implement effective input and output systems.
Subject Synopsis/ Indicative Syllabus	<ul style="list-style-type: none"> • Characteristics of Chinese Language: Historical Development, Geographical Variations (Dialects) and Linguistic Descriptions (Character, Morph, Word, Phrase and Sentence). • Representation of Chinese Character Sets: Mathematical description of representation, character set organization, encoding schemes (ISO2022 and UTF), discussion of (defacto) standard character sets (for PC, workstation and network), character decoding techniques and character conversion problems. • Output Processing of Chinese: Typesetting terminology and text rendering process, Bitmap fonts (representation, compression and scaling problems), Vector/Outline fonts (Limn algorithm), X-Window Fonts (BDF and Postscript) and font related operations (Installation, Specification, Extraction and User-Defined Glyph addition), Automatic Glyph Construction.

	<ul style="list-style-type: none"> • Input Processing of Chinese: Introduction to Chinese input processing by pen, image, speech and keystrokes, Shape-based keystroke input method, Phonetic-based input method, Shape-phonetic based input method, Evaluation metrics for keystroke input method, Phrase-based input method, Sentence-based input method, Microsoft Windows Input Method Architecture, X-Windows Input Method Architecture. • Software Development for Chinese Computing: Internationalization, Localization, ANSI-C model, Windows programming for Chinese Computing (Microsoft- and X-Windows). • Algorithmic Problems in Chinese Computing: Character set selection (NP-complete), Hashing functions for Chinese character sets, String searching (KMP, BM or Sunday), Dictionary lookup for Chinese computing, String-set searching (Aho and Corasick Algorithm). 																																	
Teaching/Learning Methodology	<p>Characteristics of Chinese language are taught before computer representation and processing of Chinese. For tutorials, exercises and computer software demonstrations will be given to consolidate understanding in class. At least one assignment will be given to develop a realistic Chinese computing application.</p> <p>39 hours of class activities including - lecture, tutorial, lab, workshop seminar where applicable</p>																																	
Assessment Methods in Alignment with Intended Learning Outcomes	<table border="1" data-bbox="579 1137 1361 1417"> <thead> <tr> <th rowspan="2">Specific Assessment Methods/Tasks</th> <th rowspan="2">% weighting</th> <th colspan="5">Intended subject learning outcomes to be assessed</th> </tr> <tr> <th>a</th> <th>b</th> <th>c</th> <th>d</th> <th></th> </tr> </thead> <tbody> <tr> <td>Assignments, Tests & Projects</td> <td>55</td> <td>✓</td> <td>✓</td> <td>✓</td> <td>✓</td> <td></td> </tr> <tr> <td>Final Examination</td> <td>45</td> <td>✓</td> <td>✓</td> <td>✓</td> <td></td> <td></td> </tr> <tr> <td>Total</td> <td>100</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	Specific Assessment Methods/Tasks	% weighting	Intended subject learning outcomes to be assessed					a	b	c	d		Assignments, Tests & Projects	55	✓	✓	✓	✓		Final Examination	45	✓	✓	✓			Total	100					
Specific Assessment Methods/Tasks	% weighting			Intended subject learning outcomes to be assessed																														
		a	b	c	d																													
Assignments, Tests & Projects	55	✓	✓	✓	✓																													
Final Examination	45	✓	✓	✓																														
Total	100																																	
Student study effort expected	<table border="1" data-bbox="579 1462 1361 1626"> <tr> <td colspan="2">Class Contact:</td> </tr> <tr> <td>Class activities (lecture, tutorial, lab)</td> <td>39 hours</td> </tr> <tr> <td colspan="2">Other student study effort:</td> </tr> <tr> <td>Assignments, Quizzes, Projects, Exams</td> <td>65 hours</td> </tr> <tr> <td>Total student study effort</td> <td>104 hours</td> </tr> </table>	Class Contact:		Class activities (lecture, tutorial, lab)	39 hours	Other student study effort:		Assignments, Quizzes, Projects, Exams	65 hours	Total student study effort	104 hours																							
Class Contact:																																		
Class activities (lecture, tutorial, lab)	39 hours																																	
Other student study effort:																																		
Assignments, Quizzes, Projects, Exams	65 hours																																	
Total student study effort	104 hours																																	
Reading list and references	<ol style="list-style-type: none"> (1). 胡裕樹, 現代漢語, 三聯書局 2011 (2). Berry, K. and Hargraves, K.A. GNU Font utilities (Limn algorithm), on WWW. (3). Hopcroft, J.E. and Ullman, J.D., 2008, Introduction to Automata, Theory and Languages, Addison-Wesley. (4). Kano, N., 2003, Developing International Software, Microsoft Press. (5). Lunde, K., 2009, Understanding Japanese Information Processing, O Reilly & Associates. 																																	

	<p>Communications of COLIPS Journal and Conference Proceedings of Computer Processing of Oriental Languages Journal of Chinese Information Processing</p>
--	---