Prof. Samuel W.K. CHAN
Curriculum Vitae

Contact:

Corresponding address:

Rm 934, Cheng Yu Tung building,
The Chinese University of Hong Kong
Shatin, N.T.
Hong Kong SAR

Telephone: (852) 3943 8562 (office)
Fax: (852) 2603 5104 (office)
Email: swkchan@cuhk.edu.hk

Education/Academic Qualifications:

- The University of New South Wales, Australia

Ph.D. in Computer Science & Engineering
Department of Artificial Intelligence

- The Chinese University of Hong Kong

M.Phil. in Computer Science

- The Chinese University of Hong Kong

Postgraduate Diploma in Education

- Victoria University of Manchester, United Kingdom

M.Sc. in Computer Science

Research Interests:

Text Analytics; Sentiment Analysis; Machine Learning; Text/Data Stream Mining; Computational Linguistics; Chinese Text in Information Systems; Information Retrieval / Extraction and their applications

Academic Experience:

- Associate Professor, Department of Decision Sciences & Managerial Economics, The Chinese University of Hong Kong. [August, 2003 – Present]

  Teaching areas: Data Mining, Web/Mobile Applications Development, Information Systems Auditing, Business Processes Analysis, Internet Programming, Information Systems and E-Commerce (Undergraduate and postgraduate levels)

- Assistant Professor, Department of Decision Sciences & Managerial Economics, The Chinese University of Hong Kong. [August, 2000 – July, 2003]

  Teaching areas: System Development Methodology, Information Systems Analysis and Design, Internet Programming, Object-Oriented Programming.

- Assistant Professor, Department of Computer Science, City University of Hong Kong. [Nov, 1996 – June, 2000]

  Teaching areas: Artificial Intelligence, Human Computer Interaction, Computer Programming, Software Engineering, Databases.

Awards and Honors:

- *First Place, Best Paper Award*, in The 7th International Conference of Computational Linguistics and Intelligent Text Processing, (CICLing), Mexico City, Mexico.
- *First in the open track of the corpus from Academia Sinica*, in The Fourth International Chinese Language Processing Bakeoff & the First CIPS Chinese Language Processing Evaluation.
- *Second in the closed track of the corpus from City University of Hong Kong*, in The Fourth International Chinese Language Processing Bakeoff & the First CIPS Chinese Language Processing Evaluation.
- *Third in two closed tracks of the corpora from State Language Commission of P.R.C. and Academia Sinica*, in The Fourth International Chinese Language Processing Bakeoff & the First CIPS Chinese Language Processing Evaluation.
- *Outstanding Student Paper Award*, in IEEE International Conference on Systems, Man, and Cybernetics. Beijing, China.
- *Australian Postgraduate Research Award*, by the Australian Federal Government.
- *Ansett Postgraduate Scholarship*, by the Australia Ansett Airline.
- *Faculty of Engineering Scholarship*, by the University of New South Wales.

Research Grants (on a competitive basis):

Sources	Project Title	Capacity	Period
GRF, Research Grants Council	*Encoding Supra-Syntactic Clues in the Parsing of Chinese Sentences*	Principal Investigator (PI)	2014 – 2016
GRF, Research Grants Council	*Super-Sense Tagging of Chinese Unknown Words*	Principal Investigator (PI)	2010 – 2013
Public Policy Research, Research Grants Council	*Analyzing Stakeholders' Views on Environmental Issues through Text Analysis - An Interdisciplinary Approach Using Computational Content Analysis*	Co-Investigator (Co-I)	2013 – 2015
GRF, Research Grants Council	*Chinese Noun Phrases Parsing in Content Sensitive Analysis*	Principal Investigator (PI)	2008 - 2011
GRF, Research Grants Council	*Learning Shallow Natural Language Patterns: Toward a Direct Inference from Chinese Financial News Articles*	Principal Investigator (PI)	2006 - 2008
GRF, Research Grants Council	*A Natural Language Front-End: Synergy between a Chinese Treebank and Verb Dictionaries in a Connectionist Model*	Principal Investigator (PI)	2004 - 2007
GRF, Research Grants Council	*A Connectionist-Based Chinese Hypermedia Document Classifier: Using Text Structure and Image as Complementary Media*	Principal Investigator (PI)	2001- 2004
GRF, Research Grants Council	*Encoding the Connectivity of Events: A Dynamic Connectionist Approach to Gist Preservation*	Principal Investigator (PI)	2000 - 2003
GRF, Research Grants Council	*Lexical Cohesion and Boundary Detection in Chinese Discourse: A Computational Approach*	Principal Investigator (PI)	1999 - 2002
Strategic Research Grant, City University of HK	*A Computational Approach to Analyzing Lexical Cohesion in Chinese Discourse Structure*	Principal Investigator (PI)	1998 - 2000

Partial List of Publications (Principal author of 60+ refereed journals articles and conference papers)

Chan, S.W.K., Chong, M.W.C., & Lai, T.B.Y. (2015). Unknown Chinese Composite Words Tagging using a Selective Back-Off Smoothing. *Journal of Chinese Linguistics*.

Chan, S.W.K. (2013). Generating context templates for word sense disambiguation. *Proceedings of the 26th Australasian Joint Conference on Artificial Intelligence*, New Zealand. Also appeared in *Lecture Notes in Computer Science*, v.8272, 466-477, Springer-Verlag.

Chan, S.W.K., & Chong, M.W.C. (2013a). Predicting part-of-speech tags and morpho-syntactic relations using similarity-based technique. *Proceedings of 1st International Conference on Statistical Language and Speech Processing*. Tarragona, Spain. Also appeared in *Lecture Notes in Artificial Intelligence*, v.7978, 71-82, Springer-Verlag.

Chan, S.W.K., & Chong, M.W.C. (2013b). Recursive part-of-speech tagging using word structures. *Proceedings of the Sixteenth International Conference on Text, Speech and Dialogue*, Czech Republic. Also appeared in *Lecture Notes in Artificial Intelligence*, v.8082, 419-425, Springer-Verlag.

Chan, S.W.K. & Franklin, J. (2011). A Text-based decision support system for financial sequence prediction. *Decision Support Systems*, 52, 189-198.

Chan, S.W.K., Chong, M.W.C., & Cheung, L.Y.L. (2011). An analysis of tree topological features in classifier-based unlexicalized parsing. *Proceedings of 12th International Conference on Computational Linguistics and Intelligent Text Processing*. Also appeared in *Lecture Notes in Computer Science*, v.6608, 155-170, Springer-Verlag.

Chan, S.W.K. (2009). Shallow Semantic Labeling using Two-Phase Feature-Enhanced String Matching. *Expert Systems with Applications*, 36, 9729-9736.

Chan, S.W.K., Cheung, L.Y.L., Chong, M.W.C. (2010). Tree Topological Features for Unlexicalized Parsing. *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010) Main Conference Poster Session*, 117-125.

Chan, S.W.K., Cheung, L.Y.L., Chong, M.W.C. (2010). A machine learning parser using an unlexicalized distituent model. *Proceedings of 11th International Conference on Computational Linguistics and Intelligent Text Processing*. Also appeared in *Lecture Notes in Computer Science*, v.6008, 121-136, Springer-Verlag.

Chan, S.W.K., Cheung, L.Y.L., Chong, M.W.C. (2009). Level-based shallow parsing using supervised learning. *Proceedings of the Chinese Information Processing Society on Sentence Parsing Evaluation* (*CIPS-ParsEval-2009*), 4-September 2009, Beijing.

Chan, S.W.K. (2009). Knowledge discovery from financial text. *Proceedings of the 9th International Conference on Electronic Business*, 30 November – 4 December 2009, Macau.

Chan, S.W.K. (2008). Finding approximate language patterns. *Proceedings of the International Conference on Bio-Inspired Signals and Systems*, 28-31 January, Madeira, Portugal, 295-301.

Chan, S.W.K. & Mickey W.C. Chong (2008). An agent-based approach to Chinese word segmentation. *The Sixth SIGHAN Workshop on Chinese Language Processing (SIGHAN6), The Third International Joint Conference on Natural Language Processing*, 7-12 January, Hyderabad, India, 112-114.

Chan, S.W.K. (2007). An edit distance approach to shallow semantic labelling. In H. Yin, P. Tino, & X. Yao (Ed.), *Proceedings of the 8th International Conference on Intelligent Data Engineering and Automated Learning* (IDEAL'07). Also appeared in *Lecture Notes in Computer Science*, v.4881, 57-66. Berlin, Germany: Springer-Verlag.

Chan, S.W.K. (2007). A decision tree approach to sentence chunking. In M.A. Orgun & J. Thornton (Ed.), *Proceedings of the Twentieth Australian Joint Conference on Artificial Intelligence* (AI'07). 2-6 December. Also appeared in *Lecture Notes in Artificial Intelligence*, v.4830, 425-434. Berlin, Germany: Springer-Verlag.

Chan, S.W.K. (2006). Beyond Keyword and Cue-Phrase Matching: A Sentence-based Abstraction Technique for Information Extraction. *Decision Support Systems*, 44, 2, 759-777.

Chan, S.W.K. (2006). Shallow case role annotation using two-stage feature-enhanced string matching. In A. Gelbukh (Ed.), *Proceedings of the 7th International Conference on Computational Linguistics and Intelligent Text Processing, (CICLing 2006)*, Mexico City, Mexico. Also appeared in *Lecture Notes in Computer Science*, v. 3878, 263-274. Berlin, Germany: Springer-Verlag.

Chan, S.W.K. (2005). Multi-Attributes Image Analysis for the Classification of Web Documents using Unsupervised Technique. In M. Gallagher, J. Hogan, F. Maire (Eds.), *Lecture Notes in Computer Science*, v. 3578, 78-85. Berlin, Germany: Springer-Verlag.

Chan, S.W.K., & Chong, M.W.C. (2004). Unsupervised clustering for non-textual Web document classification. *Decision Support Systems*, 37, 3, 377-396.

Chan, S.W.K. (2004). Automatic Discourse Structure Detection using Shallow Textual Continuity. *International Journal of Human-Computer Studies,* 61, 1, 138-164.

Chan, S.W.K. (2004). A Treebank-Based Case Role Annotation Using an Attributed String Matching. In M. Bramer, F. Coenen, and T. Allen (Eds.), *Research and Development in Intelligent Systems XXI,* 117-129. London, U.K.: Springer-Verlag.

Chan, S.W.K. (2004). Extraction of Shallow Language Patterns: an Approximation of Data Oriented Parsing. In G.I. Webb and X. Yu (Eds.), *Lecture Notes in Artificial Intelligence*, v.3339, 574-586. Berlin, Germany: Springer-Verlag.

Chan, S.W.K. (2004). Extraction of Textual Salient Patterns: Synergy between Lexical Cohesion and Contextual Coherence. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 34, 2, 205-218.

Chan, S.W.K., & Franklin, J. (2003). Dynamic context generation for natural language understanding: A multifaceted knowledge approach. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* 33, 1, 23-41.

Chan, S.W.K., T'sou, B.K., & Lai, T.B.Y. (2002). Textual summarization: A comparative study. *Journal of Applied Systems Studies,* 3, 1, 153-71. Cambridge International Science Publishing.

Chan, S.W.K. & T'sou, B.K. (published in 2001). Semantic inference for anaphora resolution: Toward a framework in machine translation. *Machine Translation*, 14,3/4, 163-190. Kluwer Academic.

Chan, S.W.K. (2001). Integrating linguistic primitives in learning context-dependent representation. *IEEE Transactions on Knowledge and Data Engineering: Special Issue on Connectionist Models for Learning in Structured Domains,* 13, 2, 157-175.

Chan, S.W.K. (2000). Using heterogeneous linguistic knowledge in local coherence identification for information retrieval. *Journal of Information Science*, 26, 5, 313-328.

Chan, S.W.K., T'sou, B.K., & Lai, T.B.Y. (2000). An investigation of salient information in textual summarization. *International Journal of Computer Processing of Oriental Languages*, 13, 4, 265-290. World Scientific.

Chan, S.W.K. & Franklin, J. (1998). Symbolic connectionism in natural language disambiguation. *IEEE Transactions on Neural Networks: Special Issue on Neural Networks and Hybrid Intelligent Models*, 9, 5, 739-755.

T'sou, B.K., Lai, T.B.Y., Chan, S.W.K., & Wang, W. S-Y. (1998). *Quantitative and Computational Studies on the Chinese Language*. City University of Hong Kong.

T'sou, B.K., Lin, H-L, Lai, T.B.Y., & Chan, S.W.K. (1998). Human judgment as a basis for evaluation of discourse-connective-based full-text abstraction in Chinese. *Computational Linguistics and Chinese Language Processing*, 3, 1, 101-116.

Chan, S.W.K., Leung, K.S., & Wong, W.S.F. (1996). An expert system for the detection of cervical cancer cells using knowledge-based image analyzer. *Artificial Intelligence in Medicine*, 8, 67-90, Elsevier Science.

Chan, S.W.K., Leung, K.S., & Wong, W.S.F. (1996). Object-oriented knowledge based system for image diagnosis. *Applied Artificial Intelligence*, 10, 407-438, Taylor & Francis.

Other Research Experience:

1. Chief architect, *HanMosaic* (`HanMosaic.baf.cuhk.edu.hk`), an open analytics platform in processing Chinese texts in real time.
2. Regular journal paper reviewer, *Decision Support Systems*, 2010 - Present. (Topics involved: Text Analytics and Machine Learning)
3. External Examiner, Dept. of Information Systems, Lingnan University.
4. External Examiner, Dept. of Computing, Hong Kong Polytechnics University.
5. Journal paper reviewer in, *Expert Systems in Applications, IEEE Transactions of Systems, Man, and Cybernetics, IEEE Transactions of Knowledge and Data Engineering*.
6. Committee member, program committee of *Pacific Asia Conference on Language, Information and Computation, Round Table Conference: Quantitative and Computational Studies on the Chinese Language*