

## Tsz Ting, Chung

**EDUCATION** **The Hong Kong University of Science and Technology**  
Doctor of Philosophy in Computer Science and Engineering 2021 - Present

**The Chinese University of Hong Kong**  
Bachelor of Science (Hons) in Computer Science [1st Hons, ELITE Stream] 2017 - 2021

**WORKING EXPERIENCE** **Pattern Recognition Center, WeChat AI, Tencent**  
Research Intern May 2025 - Present

**Tencent AI Lab**  
Research Intern Nov 2023 - Sept 2024

**Hospital Authority AI Lab**  
Research Assistant (search engines with QA system) Jan 2021 - July 2021

**Stanley Ho Big Data Decision Analytics Research Centre**  
Research Assistant (ASR model) Jun 2020- Sept 2020

**AWARDS & SCHOLARSHIP**

2021-Now	Hong Kong Ph.D. Fellowship, Hong Kong Research Grants Council
2024-2025	Outstanding PG Teaching Assistant Honorable Mention, HKUST
2021-2022	Professor Samuel Chanson Best PGTA Award, HKUST
2021-2022	RedBird Ph.D. Scholarship, HKUST

**RESEARCH** **Cross-Domain Contrastive Training of Embedding Models for Insight-Guided Agentic Reasoning**  
**Tsz Ting Chung**, Mo Yu\*, Jie Zhou\*, Dit-Yan Yeung - Preprint.

- (Agentic Retrieval) To address the problem of poor retrieval from queries to procedural guidance, we introduce a new embedding model trained from preference pairs that generalizes across reasoning and agentic domains.

**PRELUDE: A Benchmark Designed to Require Global Comprehension and Reasoning over Long Contexts**  
Mo Yu\*, **Tsz Ting Chung\***, Chulun Zhou\*, Tong Li\*, Rui Lu\*, Jiangnan Li\*, Liyan Xu\*, Haoshu Lu, Ning Zhang, Jing Li, Jie Zhou – *In Submission*.

- (Measure of AGI) Introduce a long-context benchmark requiring global comprehension and deep reasoning. Experiments show ICL, RAG, SFT, and DeepResearch systems fall behind humans and even more in reasoning.

**Many-Shot CoT-ICL: Making In-Context Learning Truly Learn**

**Tsz Ting Chung**, Lemaou Liu, Mo Yu, Dit-Yan Yeung – *ICML 2026*.

- (Many-shot CoT-ICL) It scales unreliably, similarity retrieval fails, and order effects grow with more demos. We reframe CoT-ICL as in-context test-time learning with 2 principles and propose CDS to improve reasoning.

**DivLogicEval: A Framework for Benchmarking Logical Reasoning Evaluation in Large Language Models**

**Tsz Ting Chung**, Lemaou Liu, Mo Yu, Dit-Yan Yeung – *EMNLP 2025*.

- (Logic Evaluation) Introduce a new benchmark to assess LLMs' logical reasoning while minimizing external influences, address data distribution bias, and propose a metric to reduce evaluation bias and uncertainty.

**Unified Triplet-Level Granularity Hallucination Evaluation for Vision Language Models**

Junjie Wu\*, **Tsz Ting Chung\***, Kai Chen\*, Dit-Yan Yeung – *TMLR 2025*.

- (LVLM Hallucination) Introduce a new framework to evaluate LVLMs' hallucination on the triplet level, with a benchmark dataset for evaluation and a mitigation method proposed based on the paper's findings.

**The Stochastic Parrot on LLMs Shoulder: A Summative Assessment of Physical Concept Understanding**

Mo Yu\*, Lemaou Liu\*, Junjie Wu\*, **Tsz Ting Chung\***, Shunchi Zhang\*, Jiangnan Li, Dit-Yan Yeung, Jie Zhou – *NAACL 2025 (Oral)*.

- (Measure of AGI) Investigate the stochastic parrot phenomenon and propose a task that alleviates the memorization issue via the usage of grid-format inputs that abstractly describe physical phenomena.

**Selection-p: Self-Supervised Task-Agnostic Prompt Compression for Faithfulness and Transferability**

**Tsz Ting Chung**, Leyang Cui, Lemaou Liu, Xinting Huang, Shuming Shi, Dit-Yan Yeung – *EMNLP 2024*.

- (Token Compression) With simple tuning and small additional parameters, LLMs can achieve a better or similar level of performance in natural language understanding tasks with compressed demonstrations.