

Patrick Xia

PERSONAL INFORMATION	paxia@cs.jhu.edu (973) 960-5954 github.com/pitrack cs.jhu.edu/~paxia <i>Efficient Information Extraction</i>	Hackerman 226 3400 N Charles Street Baltimore, MD 21218
EDUCATION	Johns Hopkins University <i>Advisor: Benjamin Van Durme</i> Center of Language and Speech Processing PhD, Computer Science M.S.E., Computer Science	2016 - 2022 (expected) 2018
	Carnegie Mellon University	2016
	B.S. Computer Science, B.S. Mathematics, Minor in Language Technologies	
REFEREED PUBLICATIONS	◦ Michelle Yuan, <u>Patrick Xia</u>, Chandler May, Benjamin Van Durme, Jordan Boyd-Graber. Adapting Coreference Resolution Models through Active Learning ACL 2022. Dublin. ◦ <u>Patrick Xia</u>, Richard Shin. Pruning Pretrained Encoders with a Multitask Objective ENLSP NeurIPS Workshop 2021. Virtual ◦ Shubham Toshniwal*, <u>Patrick Xia</u>*, Sam Wiseman, Kevin Gimpel, Karen Livescu. On Generalization in Coreference Resolution CRAC 2021. Punta Cana. Best short paper ◦ <u>Patrick Xia</u>, Benjamin Van Durme. Moving on from OntoNotes: Coreference Resolution Model Transfer EMNLP 2021. Punta Cana. ◦ <u>Patrick Xia</u>*, Guanghui Qin*, Siddarth Vashishtha, Yunmo Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron Steven White, Benjamin Van Durme. LOME: Large Ontology Multilingual Extraction EACL System Demonstrations. 2021. Virtual. ◦ <u>Patrick Xia</u>, Shijie Wu, Benjamin Van Durme. Which *BERT? A Survey Organizing Contextualized Encoders EMNLP 2020. Virtual. ◦ <u>Patrick Xia</u>, João Sedoc, Benjamin Van Durme. Incremental Neural Coreference Resolution In Constant Memory EMNLP 2020. Virtual. ◦ Abhinav Singh, <u>Patrick Xia</u>, Guanghui Qin, Mahsa Yarmohammadi, Benjamin Van Durme. <i>CopyNext</i>: Explicit Span Copying and Ailgnment in Sequence to Sequence Models SPNLP 2020. Virtual ◦ Seth Ebner*, <u>Patrick Xia</u>*, Ryan Culkin, Kyle Rawlins, Benjamin Van Durme. Multi-Sentence Argument Linking ACL 2020. Virtual. ◦ Arya D McCarthy, Christo Kirov, Matteo Grella, Amrit Nidhi, <u>Patrick Xia</u>, Kyle Gorman, Ekaterina Vylomova, Sabrina J. Mielke, Garrett Nicolai, Miikka Silfverberg, Timofey Arkhangelskiy, Nataly Krizhanovsky, Andrew Krizhanovsky, Elena Klyachko, Alexey Sorokin, John Mansfield, Valts Ernštreits, Yuval Pinter, Cassandra L Jacobs, Ryan Cotterell, Mans Hulden, David Yarowsky UniMorph 3.0: Universal Morphology LREC 2020. Virtual. ◦ Alex Wang, Jan Hula, <u>Patrick Xia</u>, Raghavendra Pappagari, Thomas R. McCoy, Roma Patel, Najoung Kim, Ian Tenney, Yinghui Huang, Katherin Yu, Shuning Jin, Berlin Chen, Benjamin Van Durme, Edouard Grave, Ellie Pavlick, Samuel R. Bowman. Can You Tell Me How to Get Past Sesame Street? Sentence-Level Pretraining Beyond Language Modeling ACL 2019. Florence.	

- J. Edward Hu, Huda Khayrallah, Ryan Culkin, [Patrick Xia](#), Tongfei Chen, Matt Post, Benjamin Van Durme. Improved Lexically Constrained Decoding for Translation and Monolingual Rewriting NAACL 2019. Minneapolis.
- Najoung Kim, Roma Patel, Adam Poliak, Alex Wang, [Patrick Xia](#), R. Thomas McCoy, Ian Tenney, Alexis Ross, Tal Linzen, Benjamin Van Durme, Sam Bowman, Ellie Pavlick. Probing What Different NLP Tasks Teach Machines about Function Word Comprehension *SEM 2019. Minneapolis. Best Paper.
- Ian Tenney, [Patrick Xia](#), Berlin Chen, Alex Wang, Adam Poliak, R Tom McCoy, Najoung Kim, Ben Van Durme, Sam Bowman, Dipanjan Das, Ellie Pavlick. What do you learn from context? Probing for sentence structure in contextualized word representations ICLR. 2019. New Orleans.
- Christo Kirov, Ryan Cotterell, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, [Patrick Xia](#), Manaal Faruqui, Sabrina J. Mielke, Arya McCarthy, Sandra Kübler, David Yarowsky, Jason Eisner, and Mans Hulden. UniMorph 2.0: Universal Morphology LREC 2018. Miyazaki.
- [Patrick Xia](#) and David Yarowsky. Deriving Consensus for Multi-Parallel Corpora: an English Bible Study IJCNLP 2017. Taipei.

UNREFEREED

- Yunmo Chen, Seth Ebner, Tongfei Chen, [Patrick Xia](#), Elias Stengel-Eskin, Tzu-Ray Su, J Edward Hu, Nils Holzenberger, Ryan Culkin, Craig Harman, Max Thomas, Thomas Lippincott, Aaron Steven White, Kyle Rawlins, Benjamin Van Durme. NIST TAC SM-KBP 2019 System Description: JHU/UR Pipeline TAC 2019. Gaithersburg.
- [Patrick Xia](#), Elias Stengel-Eskin, Tongfei Chen, Seth Ebner, Nils Holzenberger, Ryan Culkin, Pushpendre Rastogi, Xutai Ma, Benjamin Van Durme. NIST TAC SM-KBP 2018 System Description: JHU/UR Pipeline TAC 2018. Gaithersburg.
- Ryan Cotterell, Christo Kirov, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, [Patrick Xia](#), Manaal Faruqui, Sandra Kübler, David Yarowsky, Jason Eisner, and Mans Hulden. CoNLL-SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection in 52 Languages In Proceedings of the CoNLL SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection. 2017. Vancouver
- Philip Massey, [Patrick Xia](#), David Bamman, Noah A. Smith. Annotating Character Relationships in Literary Texts *arXiv* abs/1512.00728. 2015.

WORK

- | | |
|--|------|
| Research Intern, Microsoft Semantic Machines (Host: Richard Shin) | 2021 |
| ◦ Multitask model compression via pruning | |
| Research Intern, Google (Host: William Cohen) | 2019 |
| ◦ Used knowledge graphs to improve understanding and ground natural language | |
| Software Engineer Intern, Facebook | 2015 |
| ◦ Improved parser and semantic analyzer for an internal query language | |
| ◦ Created monitoring service for backend data pipelines | |
| Software Development Engineer Intern, Microsoft | 2014 |
| ◦ Designed and implemented new UX features for Microsoft Office | |

PRESENTATIONS

- | | |
|--|------------|
| ◦ LOME: Large Ontology Multilingual Extraction
EACL/Virtual | April 2021 |
| ◦ Multilingual and Cross-Domain Coreference Resolution
Johns Hopkins University | March 2021 |

	◦ Incremental Coreference Resolution in Constant Memory SustainNLP@EMNLP/Virtual	Nov 2020
	◦ Which *BERT? A Survey on Contextualized Encoders EMNLP/Virtual	Nov 2020
	◦ (Very) Recent Ideas in Contextualized Sentence Representations Johns Hopkins University	Oct 2019
	◦ Probing for Sentence Structure in Contextualized Word Representations Johns Hopkins University	Apr 2019
	◦ Deriving Consensus for Multi-Parallel Corpora: an English Bible Study International Joint Conference on Natural Language Processing Taipei, Taiwan.	Nov 2017
	◦ Translating into Morphologically Rich Languages with Word-Level LSTMs Senior Honors Thesis (<i>Advised by: Chris Dyer</i>) Carnegie Mellon University	May 2016
TEACHING	Teaching Assistant, Johns Hopkins University	
	◦ Declarative Methods	Spring 2017
	Teaching Assistant, Carnegie Mellon University	
	◦ Great Theoretical Ideas of Computer Science (15-251)	Spring 2015 - Spring 2016
	◦ Parallel and Sequential Data Structures and Algorithms (15-210)	Fall 2014
	◦ Principles of Imperative Programming (15-122)	Fall 2013
ADVISING	◦ Z. Song Classifying Reddit discourse acts. Poster at MASC 2018 ◦ M. Harwani Abstractive summarization using Reddit autotldr data	
OTHER TECHNICAL PROJECTS	Puflantu	2019
	◦ <i>Puflantu</i> is a conlang created for Galactic Puzzlehunt with over 1K expressed word types ◦ Contributed to phonetics and typology decisions, vocabulary creation, glossing, and translation	
	Codenames	2016
	◦ AI using word vectors to find optimal clue phrases for the board game <i>Codenames</i>	
	Algorithms with a Purpose, Carnegie Mellon University	2013 - 2015
	◦ Designed algorithmic problems for an annual day-long competitive event ◦ Implemented test solutions and starter code for those problems	
	TeX Team, TartanHacks	2014
	◦ Crowdsourced and gamified translation of handwritten math into LaTeX ◦ Awarded Microsoft Prize	
	Classity, PennApps	2013
	◦ Real-time virtual learning environment with interaction between lecturers and students ◦ Awarded "Best Cloud-Connected Hack", "Best Hack That Makes Life So Easy"	
REFERENCES	Benjamin Van Durme, Associate Professor at Johns Hopkins University vandurme@cs.jhu.edu. Relationship: PhD advisor.	
	João Sedoc, Assistant Professor at New York University Stern jsedoc@stern.nyu.edu. Relationship: Research collaborator.	
	Richard Shin, Research Scientist at Microsoft Semantic Machines richard.shin@microsoft.com. Relationship: Internship mentor.	

MISC.

Programming Languages: Python, C++, C

Frameworks/libraries: TensorFlow, PyTorch, AllenNLP, Transformers

U.S. Permanent Resident

- Lived in Ireland, Australia (citizen), Hong Kong, and US
- Conversational in Mandarin Chinese; familiar with French

Puzzlehunts

- Puzzle writing, organization, and testsolving
 - teammate hunt (online, once/year) 2020 - present
 - Galactic Puzzlehunt (online, once/year) 2017 - present
 - CMU Puzzlehunt (on-site, twice/year) 2012 - 2017
- 1st at Microsoft College Puzzle Challenge (400 teams of 4 students) 2016