

# William Chan

✉ williamchan@google.com ☎ +1 650 450-9455 🌐 williamchan.ca

## Education

<b>George Brown College</b> Culinary Arts	2019 - current
<b>Carnegie Mellon University</b> PhD Electrical and Computer Engineering Thesis: <i>End-to-End Speech Recognition Models</i> Research Areas: <i>Machine Learning, Deep Learning, Speech Recognition</i>	2011 - 2016
<b>University of Waterloo</b> BASc Computer Engineering, Management Sciences Option	2006 - 2011
<b>National University of Singapore</b> Computer Engineering Exchange Student	2010

## Professional Experiences

<b>Google - Google Brain Research Scientist</b>	2017 - current
<ul style="list-style-type: none"><li>• <i>Bytes are All You Need: End-to-End Multilingual Speech Recognition and Synthesis with Bytes</i></li><li>• <i>Optimal Completion Distillation for Sequence Learning</i></li><li>• <i>InferLite: Simple Universal Sentence Representations from Natural Language Inference Data</i></li><li>• <i>Illustrative Language Understanding: Large-Scale Visual Grounding with Image Search</i></li><li>• TensorFlow PinToHostOptimizer, &gt; 15% improvement in Transformer training speed</li><li>• Manager: Geoffrey Hinton</li><li>• Research Areas: machine learning, deep learning, natural language processing</li></ul>	
<b>Google - Google Brain Research Intern</b>	2016
<ul style="list-style-type: none"><li>• <i>Latent Sequence Decompositions</i></li><li>• <i>Very Deep Convolutional Networks for End-to-End Speech Recognition</i></li><li>• Mentors: Navdeep Jaitly and Quoc Le</li><li>• Research Areas: machine learning, deep learning, speech recognition</li></ul>	
<b>Google - Google Brain Research Intern</b>	2015
<ul style="list-style-type: none"><li>• <i>Listen, Attend and Spell</i></li><li>• Mentors: Navdeep Jaitly, Quoc Le and Oriol Vinyals</li><li>• Research Areas: machine learning, deep learning, speech recognition</li></ul>	
<b>Google - Google Brain Research Intern Part-Time</b>	2015
<ul style="list-style-type: none"><li>• Cooking secret sauces part-time while continuing PhD @CMU</li><li>• Various code optimizations, ??% efficiency in CPU and memory over baseline DistBelief implementation</li><li>• Research Areas: machine learning, deep learning</li></ul>	
<b>Google - Natural Language Processing Research Intern Part-Time</b>	2014
<ul style="list-style-type: none"><li>• Continuation of summer internship but now working 1 day / week while continuing PhD @CMU</li><li>• ?x production inference code optimization under certain scenarios</li><li>• Research Areas: machine learning, deep learning, natural language processing</li></ul>	

- Google** - Natural Language Processing Research Intern 2014
- SAFT Machine Intelligence
  - Deep Learning, Natural Language Processing and Multilingual Coreference Resolution
  - 7% increase in coref performance on the CoNLL dataset over baseline system
  - Research Areas: machine learning, deep learning, natural language processing
- Google** - Machine Learning Ads Intern 2012
- Display Campaign Optimizer
  - Replaced a heuristic ranking model with a machine learning ranking model, 22% increase in Click Through Rate and 22% increase in customer return on investment in certain scenarios (patent pending)
  - Added various Map-Reduce analysis scripts
  - Business Impact: higher ROI for customers, higher revenue for Google and more satisfied customers
- TD Securities** - Quantitative Financial Engineer Intern 2010
- TD Bank's trade floor in the Quantitative Research Group
  - Optimized Monte Carlo Simulation through dividend events consolidation, 500%-1700% performance gain on dividend heavy deals
  - Added a new mixed Dividend Pricing Model, allows pricing of most exotic derivatives using an inter-blended discrete dollar and proportional dividend schedule
  - Architected new 2D Interpolation Infrastructure and added Thin Plate Spline Interpolation for Local Volatility Surfaces
  - Cholesky decomposition optimizations for FX/Equity correlation matrices by bootstrapping similar matrices used in Monte Carlo Simulations
  - Business Impact: faster execution, more accurate pricing, happier clients, better risk management and hedging
- Amazon** - Software Engineer Intern 2009
- Amazon Prime AdAttribution
  - Re-architected database schemas (Oracle and MySQL)
  - Modified data mining infrastructure to more efficiently attribute Amazon Prime signups to specific ads for statistical analysis
  - Business Impact: faster and more effective analysis of customer behaviour
- Google** - Disk Platform Engineer Intern 2009
- Research and analysis on disk latency and bandwidth
  - Data analysis on disk performance across entire Google's server fleet
  - Added a new Linux Kernel Power Capping Module to output detailed CPU usage to assist in power management across Google's server fleet
  - Business Impact: maximize utility, lower capital and operating costs, negotiation leverage with disk vendors
- Intel** - Mobile Platform Architecture Engineer Intern 2008
- Adaptive Power Algorithms Research
  - Designed and prototyped several dynamic frequency algorithms for CPU/GPU to reduce dynamic power, +55% CPU power reduction with no performance penalty in certain apps/games
  - Multi-Chip Package Power Controller Integral Algorithm
  - New methodology to increase CPU and GPU performance while maintaining same thermal design power envelope
  - Statistical Data Analysis on Power Traces - analyze and isolate power events for simulation and forecasting next generation mobile platform architectures

- Business Impact: additional product value with zero capital and marginal unit cost

#### **NVIDIA** - Driver Software Engineer Intern

2007

- Driver Optimizations
- Various driver optimizations including +20% performance on Call of Duty 2 on G80
- Business Impact: additional product utility without additional capital or marginal costs, product differentiation against competitor

#### **AMD** - Hardware Engineer Intern

2007

- High Bandwidth Video Playback Research and Development
- Developed custom video system for 1920x1080 @ 120Hz playback with throughput of over 6Gb/s and designed to scale over 240Hz
- Developed (Verilog) FPGA for a custom proprietary board for converting a DVI/HDMI signals to LVDS for a variety of LCD panels
- Floating-point Optimizations (SSE2 and SSE3)
- Business Impact: capital cost savings (> 100k CAD) with custom in-house solutions vs commercial purchases

### **Publications**

1. Bo Li, Yu Zhang, Tara Sainath, Yonghui Wu, **William Chan**, "Bytes are All You Need: End-to-End Multilingual Speech Recognition and Synthesis with Bytes," in arXiv 2018.
2. Sara Sabour, **William Chan** and Mohammad Norouzi, "Optimal Completion Distillation for Sequence Learning," in ICLR 2019.
3. Jamie Ryan Kiros and **William Chan**, "InferLite: Simple Universal Sentence Representations from Natural Language Inference Data," in EMNLP 2018.
4. Jamie Ryan Kiros\*, **William Chan\*** and Geoffrey Hinton, "Illustrative Language Understanding: Large-Scale Visual Grounding with Image Search," in ACL 2018.
5. Takaaki Hori, Shinji Watanabe, Yu Zhang and **William Chan**, "Advances in Joint CTC-Attention based End-to-End Speech Recognition with a Deep CNN Encoder and RNN-LM," in INTERSPEECH 2017.
6. **William Chan**, Yu Zhang, Quoc Le and Navdeep Jaitly, "Latent Sequence Decompositions," in ICLR 2017.
7. Yu Zhang, **William Chan** and Navdeep Jaitly, "Very Deep Convolutional Networks for End-to-End Speech Recognition," in ICASSP 2017.
8. **William Chan** and Ian Lane, "On Online Attention-based Speech Recognition and Joint Mandarin Character-Pinyin Training," in INTERSPEECH 2016.
9. Guan-Lin Chao, **William Chan** and Ian Lane, "Speaker-Targeted Audio-Visual Models for Speech Recognition in Cocktail-Party Environments," in INTERSPEECH 2016.
10. **William Chan**, Navdeep Jaitly, Quoc Le and Oriol Vinyals, "Listen, Attend and Spell: A Neural Network for Large Vocabulary Conversational Speech Recognition," in ICASSP 2016. **Best Student Paper Award.**
11. **William Chan** and Ian Lane, "Deep Recurrent Neural Networks for Acoustic Modelling," in arXiv 2015.
12. **William Chan\***, Nan Rosemary Ke\* and Ian Lane, "Transferring Knowledge from a RNN to a DNN," in INTERSPEECH 2015.
13. **William Chan** and Ian Lane, "Deep Convolutional Neural Networks for Acoustic Modeling in Low Resource Languages," in ICASSP 2015.
14. **William Chan** and Ian Lane, "Distributed Asynchronous Optimization of Convolutional Neural Networks," in INTERSPEECH 2014.
15. Le Nguyen, Pang Wu, **William Chan**, Wei Peng and Joy Zhang, "Predicting Collective Sentiment Dynamics from Time-series Social Media," in ACM SIGKDD WISDOM 2012.
16. **William Chan** and Jason Lohn, "Spike Timing Dependent Plasticity with Memristive Synapse in Neuromorphic Systems," in IEEE IJCNN 2012. **Travel Fellowship for Spiking Neural Networks.**

## Service

- Reviewer (past and/or present) for:
  - IEEE International Conference on Acoustics, Speech and Signal Processing
  - IEEE Signal Processing Letters
  - IEEE Transactions on Evolutionary Computation
  - IEEE Transactions on Neural Networks and Learning Systems
  - International Conference on Artificial Intelligence and Statistics (AISTATS)
  - International Conference on Learning Representations (ICLR)
  - International Conference on Machine Learning (ICML)
  - INTERSPEECH
  - Neural Information Processing Systems (NIPS)

## References

- Chris Dyer (Carnegie Mellon University and DeepMind)
- Bhiksha Raj (Carnegie Mellon University)
- Florian Metze (Carnegie Mellon University)
- Ian Lane (Carnegie Mellon University)
- Navdeep Jaitly (Google Brain)
- Oriol Vinyals (DeepMind)
- Quoc Le (Google Brain)