Christopher R. Collins

chris@crcollins.com • (678) 632-8759
http://github.com/crcollins
http://crcollins.com

Education

Carnegie Mellon University — **Pittsburgh, PA**
Ph.D. in Theoretical/Computational Chemistry — *2014 – 2019 Expected*
Thesis Topic: Using Machine Learning and Molecular Similarity to Predict Chemical Properties

University of North Georgia — **Dahlonega, GA**
B.S. in Chemistry with Chemical Physics Concentration — *2010 – 2014*
Seminar Topic: Application of Machine Learning to Predict the Optoelectronic Properties of Benzobisazoles

Research

Carnegie Mellon University — **Pittsburgh, PA**
Graduate Student; Advisor: David Yaron — *Aug. 2014 – Present*

- Developed parallel and distributed code for optimizing INDO parameters (50x speedup on 60 core cluster)
- Achieved state-of-the-art performance for predicting properties of polymers with Deep Learning
- Constructed new molecular descriptors for machine learning in chemistry
- Developed method for predicting atomic forces by differentating Deep Learning models

University of North Georgia — **Dahlonega, GA**
Undergraduate Researcher; Advisor: Aimée Tomlinson — *Aug. 2011 – May 2014*

- Researched Benzobisazole Cruciform structures using DFT for use as efficient organic solar cells
- Built set of tools to automate research process (reduced hours of work to minutes)
- Formulated machine learning model for predicting Benzobisazole properties (orders of magnitude faster than comparable methods for the same accuracy)

Carnegie Mellon University — **Pittsburgh, PA**
Undergraduate Researcher; Advisor: David Yaron — *May 2013 – Aug. 2013*

- Researched decomposition of Thiaphospholes using quantum computational methods
- Wrote MATLAB code to automate running Gaussian/AMPAC calculations and parsing results

Work Experience

Google — **Pittsburgh, PA**
Software Engineering Intern; Host: John Tantalo — *Summer 2017*

Tripwire, Inc. — **Alpharetta, GA**
Engineering Intern — *May 2012 – May 2014*

- Rewrote entire test suite for Benchmark product using Selenium and PyTest (Python)
- Aided in migration of critical IP360 product and wrote automated test suite for API
- Expanded automated installation program for CCM product
- Wrote script to test for CSRF and XSS vulnerabilities in IP360 web application

Technical Skills

Proficient Languages: Python, MATLAB, Octave, Javascript, C, C++ , Bash, Mathematica, LaTeX, HTML/HTML5, SQL
Familiar Languages: Go, CUDA C, TI-BASIC, CSS, Maple, Haskell, ASM (for my CPU), x86 ASM, Ruby, Java
Software: *nix (Linux, Centos, FreeBSD, AIX, Solaris), Vim, Gaussian/AMPAC, Git, SVN, SSH, rsync, coreutils, Make, Logisim, Vagrant, PostgreSQL, Varnish, nginx, Apache, Torque, ab, LabVIEW
Libraries: NumPy, SciPy, Matplotlib, Scikit-Learn, Django, CVXPY, CVX, Keras, PyBrain, Caffe, TensorFlow, Theano, Selenium, Paramiko, PyTest, Twitter Bootstrap, jQuery, jQueryUI, Swig, SDL, OpenMP, OpenMPI

Side Projects

MolML
(https://github.com/crcollins/molml)

- Library to extract features from molecules to interface with Scikit-Learn
  - Extracts 4 classes of atom-wise and 4 classes of molecule-wise features (~20-30 descriptors)
  - Parallel feature generation
- Full automated testing suite (100% code coverage)

Chemtools/Chemtools-Webapp
(http://gauss.crcollins.com)

- Implemented a set of tools for use in research
  - Gaussian log output parser to collect molecular properties
  - Machine learning model to predict polymer properties from oligomers
  - Benzobisazole structure generator and machine learning predictor for optoelectronic properties
  - Supercomputer job submitter and curator
- Full automated testing suite (over 95% code coverage on 12k LOC)
- Python/Django Application with a Bootstrap front end
- Vagrant VM build system

pyOS
(https://github.com/crcollins/pyos)

- Implementation of a *nix like operating system using Python
- Created shell similar to Bash Shell
- Reimplemented standard unix programs (cp, mv, rm, ls, etc)
- Included permissions, pipes, multiple users, and a virtual file system

Other

- Terminal Text Editor (C)
- 8 bit CPU (Logisim)
- Mandelbrot/Buddhabrot Generator (Python, Javascript, C, C++ , CUDA C, TI-BASIC)
- Quadtree Visualizations (Python, C, C++ , Javascript)
- Cellular Automata Based Electronic Simulation (Javascript)
- Object Relation Management Library (Python/SQL)
- University Course Scraper (Python)
- Cryptography Library (Python, Logisim)
- Automated Peer Review System Web Application (Python/Django)
- Virtual Cluster and Server Architecture Vagrant Builds

Courses

(UNG) Linear Algebra
(UNG) Intro to Cryptography
(UNG) Computational Methods in Physics
(Udacity) Intro to Parallel Programming
(CMU) Introduction to Machine Learning
(CMU) Statistical Machine Learning
(CMU) Graduate Artificial Intelligence
(CMU) Quantum Chemistry
(CMU) Special Topics in Computational Quantum Chemistry
(CMU) Computational Chemistry
(CMU) Convex Optimization
(CMU) Topics in Deep Learning
(CMU) Deep Reinforcement Learning and Control

Publications

Collins, C. R. ; Li, H.; Gordon, G. J.; Yaron, D. J. Continuous Bag of Bonds: A Size-Consistent Way of Representing Molecules. *Submitted*.

Collins, C. R. ; Gordon, G. J.; von Lilienfeld, O. A.; Yaron, D. J. Constant Size Molecular Descriptors For Use With Machine Learning. *Submitted*.

Collins, C. R. ; Tomlinson, A. L. Application of Machine Learning to Predict the Optoelectronic Properties of Benzobisazoles. *In Progress*.

Collins, C. R. ; Yaron, D. J. Developing Coarse-Grained Site Models for Excited Electronic States of Conjugated Polymers. *SPIE* **2015**.

Qiu, Y.; Worch, J. C.; Chirdon, D. N.; Kaur, A.; Maurer, A. B.; Amsterdam, S.; **Collins, C. R.** ; Pintauer, T.; Yaron, D.; Bernhard, S.; Noonan, K. J. T. Tuning Thiophene with Phosphorus: Synthesis and Electronic Properties of Benzobisthiaphospholes. *Chem. Eur. J.* **2014**.

Tlach, B. C.; Tomlinson, A. L.; Morgan, K. D.; **Collins, C. R.** ; Jeffries-EL, M. Evaluation of the Impact of Extended Conjugation on the Optoelectronic Properties Benzo[1,2-d:4,5-d']bisoxazole Polymers. *Aust. J. Chem.* **2013**.

Teaching

- Teaching Assistant
  - 09-221 Laboratory I: Introduction to Chemical Analysis
  - 09-214 Physical Chemistry
  - 09-101 Introduction to Experimental Chemistry (2 times)
  - 09-103 Atoms, Molecules, and Chemical Change
  - 09-231 Mathematical Methods for Chemists
  - 09-106 Modern Chemistry II
- Weekly Lecture Series in Theory Suite (Organizer)
  - Intro to Python
  - Intro to Bash Scripting (2 lectures)
  - Intro to Object Oriented Programming
  - Intro to NumPy (2 lectures)

Presentations and Posters

Accelerating the (Augmented) Roothaan-Hall Method in Solving the Density Functional Theory Problem — **Pittsburgh, PA**
CMU 10-725 Convex Optimization Term Project — *2016*

Constant Size Molecular Descriptors For Use With Machine Learning — **Valencia, PA**
CMU Chemistry Department Retreat Poster Session — *2016*

Constant Size Molecular Descriptors For Use With Machine Learning — **Pittsburgh, PA**
Midwest Theoretical Chemistry Conference — *2016*

Constant Size Molecular Descriptors For Use With Machine Learning — **Zürich, Switzerland**
CECAM Chemical Space Workshop — *2016*

Using Data to Accelerate Quantum Chemical Calculations by Getting Better at Guessing — **Pittsburgh, PA**
CMU 15-780 Graduate Artificial Intelligence Term Project — *2016*

Using Machine Learning and Molecular Similarity to Predict Chemical Properties — **Pittsburgh, PA**
CMU Progress Report — *2016*

Using Machine Learning and Molecular Similarity to Predict Chemical Properties — **Pittsburgh, PA**
CMU Chemistry Department Poster Session — *2016*

Predicting Chemical Properties Using Machine Learning Methods — **Farmington, PA**
CMU Chemistry Department Retreat Poster Session — *2015*

The Influence of Chemical Representations on the Efficiency of Molecular Screening — **Pittsburgh, PA**
CMU Graduate Seminar — *2015*

A Mission to Multivariate Adaptive Regression Splines (MARS) and Its Otherworldly Neighbors — **Pittsburgh, PA**
CMU 10-702 Statistical Machine Learning Term Project — *2015*

Predicting Chemical Properties Using Machine Learning Methods — **Pittsburgh, PA**
CMU 10-701 Intro to Machine Learning Term Project — *2014*

Application of Machine Learning to Predict the Optoelectronic Properties of Benzobisazoles — **Dahlonega, GA**
UNG Senior Seminar — *2014*

Machine Learning in Chemical Compound Space — **Dahlonega, GA**
UNG Junior Seminar — *2013*

Donor-Acceptor Behavior of Benzobisazole Cruciformic Polymers — **Atlanta, GA**
American Chemical Society Southeastern Regional Meeting — *2013*

The Impact of Conjugation Length on Benzobisazole Cruciforms **Raleigh, NC**
American Chemical Society Southeastern Regional Meeting *2012*

– Runner-up Undergraduate Poster in Organic Chemistry