

Name: Mr David Goh Zhe Kai **Gender:** M **HP:** +65-92357049
E-mail: david.gohzhakai@gmail.com **Website:** davidgohzk.github.io
LinkedIn: <https://www.linkedin.com/in/david-goh-zhe-kai-3806a11bb/>

Experiences

Mar 2023 - *Fabrica.ai*, Intern - Robotics Software Engineer

Jun 2023 <https://www.fabrica.ai/>

- Interned as a Robotics Software Engineer and actively contributing to the development of an automatic tile grouting robot.
- Developed a robust **ROS2** node utilising BasicMicro Roboclaw PID controller to create a debugger for the robot's wheel motors.
- Optimised the operation flow and reliability of the robot by updating **MakeFile** and **Docker** commands.
- Enhanced the navigation logic for the new robot version using **Behavior Trees** and **ROS2** Nodes, validated through **Gazebo** simulation.
- Improved DevOps flow by implementing **pre-commit** hooks formatters and linters.
- Participated in technical interviews for internship applicants and provided valuable insights during company presentations.

Sep 2021 - *Better.sg, Heart2Heart*, Software Developer

Dec 2022 <https://better.sg/>

- Volunteered at Better.sg, a tech-for-good organisation, to develop Heart2Heart, an online platform fostering open discussions on societal topics.
- Provided expertise and guidance on **DevOps** and **Cloud Architecture** to optimise the project's infrastructure and deployment processes.
- Designed and implemented a Proof-Of-Concept **Machine Learning** algorithm to promote diversity within groups of 8 individuals by considering various attributes.

Sep 2020 - *Gracet.ai*, Intern - Full Stack Software Developer

Jan 2021

- Collaborated in a multi-national team to develop a HR management product.
- Introduced and executed a comprehensive testing flow for **GraphQL** resolvers, using **Jest** and **Node.js** to ensure the reliability and quality of the product.
- Implemented **GraphQL** resolvers for attendance on **Express**, utilising **Mongoose** to establish communication with the **MongoDB** server.
- Implemented input verification for resolvers by leveraging the **Yup** library to ensure data integrity and enhancing the overall system's security.
- Orchestrated the end-to-end development of the email verification feature by employing **React** for the frontend, **Apollo** for backend communication and **GraphQL** for efficient data handling.
- Introduced a proposed **Role-Based Access Control (RBAC)** framework for authorisation.
- Designed a comprehensive error handling flow for the backend, effectively communicating errors to the frontend, and ensuring a smooth user experience.
- Developed and integrated the Contact Us page, establishing seamless integration with the backend to facilitate effective communication with users and improve customer support.

- Dec 2020 *GovTech, GeekOut, Participant*
Chope (Hackathon)
 Chope is a device which allows those with mobility difficulties to have easy access to critical services.
- Led a team of six to create a hardware and website POC.
 - Assigned roles and responsibilities to team members.
 - Facilitated **Sprint** meetings for collaborative discussions and progress tracking.
 - Presented the project to senior developers, effectively pitching its value.
 - Developed communication between devices using **Arduino**, **ESP8266**, and **MQTT** over the internet.
- Aug 2020 *NUSH Hacks '20, Backend Developer*
nush.link (Hackathon)
 nush.link is a URL shortener website for students and teachers from NUS High School to use to shorten their URLs.
- Managed the project's development and ideation process, overseeing the creation of a POC.
 - Established a **MySQL** database and successfully connected it to the backend infrastructure.
 - Utilised **Express.js** on Node.js to develop the backend and create robust **APIs**.
- Jan 2020 – *NUS High School, Mentor Representative*
 Dec 2020 • Collaborated closely with teachers to efficiently manage class administration tasks.
- Sharing with Year 1 and Year 2 juniors*
- Organised and motivated the class to arrange a sharing session for lower-year students.
 - Led a team of twelve students in facilitating the discussion and sharing during the session.
 - Engaged with multiple stakeholders to coordinate and organize the sharing event.
- Jan 2019 – *DSO National Laboratories, Intern – Artificial Intelligence Researcher*
 Jan 2020 *Information Extraction from Air Traffic Control Communications*
 A research project which aims to use Machine Learning in Speech to Text to extract out key information from Air Traffic Control Communications
- Provided leadership and oversight to ensure the successful completion of the project.
 - Leveraged **machine learning** and **neural network** techniques to enhance the language model and entity extraction in the **speech-to-text** system.
 - Implemented FRAGE, a Long Short-Term Memory (**LSTM**) **neural network**, as the language model for improved performance.
 - Developed **Python** code to parse lattices, enabling efficient data processing.
 - Upgraded language scores within lattices using a **LSTM** language model, resulting in enhanced accuracy compared to traditional N-gram language models.
 - Implemented **beam search** algorithm to optimise the search process within lattices, improving overall system performance

- Nov 2019 *NUS High School, AppVenture, Introduction to Python Workshop, Facilitator*
- Facilitated a 3-hour workshop for 20 students, ranging from Year 4 to Year 6, as assigned by the school's Computer Science interest group.
 - Created and developed engaging workshop slides to deliver the content effectively.
 - Prepared and set up questions on Coursemology, an educational platform, for participants to practice their skills during the workshop.
 - Provided participants with foundational knowledge in **Python** and programming, empowering them to further explore the subject.
 - Actively facilitated and guided the session during the workshop, ensuring a productive and engaging learning experience for the students.
- Jan 2018 – Jan 2019 *DSO National Laboratories, Intern – Artificial Intelligence and Computer Security Researcher*
- Decompilation of Assembly Language using Machine Learning*
A research project which used machine learning to reverse engineer assembly language for better analysis of the program
- Directed and organised project meetings, ensuring effective communication and progress.
 - Utilised **IDA Pro**, a disassembler, to analyze and process assembly code as part of the project tasks.
 - Developed a **Python** program utilizing **CParser**, a Python module, to parse and standardize **C** code.
 - Decompiled and evaluated Open Source projects such as **OpenSSL** and **Redis** for further analysis and research purposes.
 - Demonstrated high scores on synthetic datasets, showcasing the capability of neural networks in decompiling simple **C** functions.
 - Successfully decompiled short **C** functions from real-world projects, highlighting the effectiveness of the developed model.
- Jun 2017 – Jan 2018 *DSO National Laboratories, Intern – Artificial Intelligence and Computer Security Research*
- Exploring Decompilation of Assembly Language using Pointer-Generator Networks*
A research project which used machine learning to reverse engineer assembly language for better analysis of the program
- Deployed Pointer-Generator Networks, a **Natural Language Processing** neural network, on a **Linux** server for experimentation.
 - Developed **Java** code to generate synthetic and simplified **C** functions.
 - Utilised **GCC** compiler to compile **C** functions and generate corresponding assembly code.
 - Combined **C** functions with their corresponding assembly code and fed them into the neural network for analysis.
 - Evaluated the output of the neural network using **ROGUE** metrics, achieving impressive scores of over 95 for all ROGUE metrics.

Other Relevant Experience

Cyber-Security

- Joined team “CS_Gang” which was 6th in Singapore at CTFTIME in July 2020 (2020)
- Completed cyber-security courses offered by Immersive Labs (2019)
- Participated in Young Defence Scientists Programme (YDSP) World of Science – Computer Security (2018)

Machine Learning and Artificial Intelligence

- Wrote a Jupyter Notebook which uses different machine learning techniques to determine the nature of a program based on the imported libraries in Windows (2020)
- Produced a 4D prediction program using neural network and genetic algorithm for school project (2020)
- Completed multiple DataCamp courses on machine learning and AI (2019-2020)
- Participated in Young Defence Scientists Programme (YDSP) World of Science – Artificial Intelligence (2018)

Full Stack Development

- Built a website for Hackathon to allow students to shorten their URL that used React.js as the front end and Express.js as the backend and MySQL as the database. (2020)
- Built a website for police database on criminals and offences for school project that used HTML, CSS, JS as well as backend PHP and MySQL (2019)

Internet of Things/Microcontroller

- Produced a GUI in Python to control LED lights from a Raspberry Pi (2016)

Competitive Programming

- Completed a school module on Algorithms and Data Structures (2019)

Computer Application

- Created a computer game using Java OOP and JavaFX for school project (2017)

Android App Development

- Made a To-Do Android app for school module project (2018)

Education

- Volunteering at Touch Young Arrow as a tutor for children from lower-income background (2023)
- Student Helper for Computer Science Booth for NUS High School Open House (2019)
- Peer tutoring at Pei Tong Primary School (2017)

Achievements

- 2023
 - Volunteered as an organiser at Hack&Roll 2023, Singapore largest student-run hackathon with over 600 participants
 - Worked with Indigitous Singapore to partner with Digital Wesley Methodist to organise a hackathon for the Singapore Wesley Methodist community
 - Finanlist in CodeExp, a mobile app Hackathon organised by DSTA
- 2022
 - Certification of Completion of the following courses on Coursera
 - Google Cloud Fundamentals: Core Infrastructure (Google Cloud)
 - Developing a Google SRE Culture (Google Cloud)
 - Foundations of Project Management (Google)
 - Bookkeeping Basics (Intuit)
 - Decentralized Finance (DeFi) Opportunities and Risks (Duke University)
 - Decentralized Finance (DeFi) Primitives (Duke University)
 - Decentralized Finance (DeFi) Deep Dive (Duke University)
 - Decentralized Finance (DeFi) Infrastructure (Duke University)
 - Certification of Specialisation on Coursera
 - Decentralized Finance (DeFi): The Future of Finance (Duke University)
 - Certification of Completion of Google Digital Garage The Fundamentals of Digital Marketing
 - Participation at Indigitous #HACK hackathon 2022
- 2021
 - Participation in CodeFiesta, a Hackathon organised by GovTech
- 2020
 - DSO Star Award for achieving A- in NUS CS2100 module
 - 11th in STACK the Flags 2020 (CTF)
 - 17th in Cyber Defenders Discovery Camp (CTF)
 - 9th in WhiteHacks2020 (CTF)
 - Participation in NUSH Hacks'20 (Hackathon)
 - Participation in NUS Department of Statistics & Applied Probability 2020 Statistics Competition
 - Participation in Singapore Science and Engineering Fair 2020
 - NUS High School Leadership Commendation Award for Mentor Representative
- 2019
 - Finanlist in Singapore Science and Engineering Fair 2019
 - DSO Star Award for achieving A in NUS CS1010S module
 - Participation in Cyberthon (CTF)
 - Bronze in Cyber Defenders Discovery Camp (CTF)
 - NUS High School CCA Commendation Award for Soccer CCA
- 2018
 - Merit in Science Mentorship Program
 - Presented at Raffles Institution Research Education Congress 2018 at the Interactive Session Project Presentation
 - Participation in Singapore Science and Engineering Festival 2018
- 2017
 - NUS High School CCA Commendation Award for Soccer CCA
 - NUS High School Service Commendation Award
 - Edusave Scholarships for Integrated Programme Schools (ESIP)
- 2016
 - Clinched best algorithm award in building an Arduino robot to navigate through the maze

Educational Qualification

2015-2020 *National University of Singapore High School of Maths and Science Diploma with Major in Chemistry and Honours in Mathematics and Computer Science*

- Python: Basics, OOP, GUI, Raspberry Pi
- Java: Basics, OOP, GUI, Android
- Web development: HTML, CSS, JS, PHP, MySQL
- Algorithms and Data Structures
- AI: Genetic Algorithms, Machine Learning, Neural Network, Fuzzy Logic
- Computer Networking

2018-2020 *National University of Singapore, Non-Graduating Programme*
Completion of the modules

- CS2106 Introduction to Operating Systems (2020)
- CS2100 Computer Organisation (2019)
- CS1010S Programming Methodology (2018)

Referees

Dr Chieu Hai Leong

Distinguished Member of Technical Staff

Information Division

DSO National Laboratories

(Research mentor for *Information Extraction from Air Traffic Control Communications*, *Decompilation of Assembly Language using Machine Learning* and *Exploring Decompilation of Assembly Language using Pointer-Generator Networks*)

Dr Khoo Wei Ming

Senior Member of Technical Staff

Information Division

DSO National Laboratories

(Research Mentor for *Decompilation of Assembly Language using Machine Learning* and *Exploring Decompilation of Assembly Language using Pointer-Generator Networks*)

Mr David Lim

Chief Executive Officer

Gracet.ai

Mr Ronald Luc

Chief Technological Officer

Fabrica.ai

Abstract of research projects

Information Extraction from Air Traffic Control Communications

Air Traffic Control (ATC) communication refers to the relay of crucial information between pilot and ground-based ATC personnel. Conveying this information is paramount in avoiding collisions with other planes, allowing a swift, efficient and safe transit.

Speech recognition in ATC will help mitigate the issue of increasing air traffic by allowing ATC personnel to focus on more critical situations. As air traffic becomes heavier, this increases the workload on ATC personnel and hinders them from focusing on more critical situations.

Speech-to-Text (STT) is a fundamental step in achieving speech recognition in ATC communications. However, this is a difficult problem due to the amalgamation of different issues such as high speech rate and noisy communication channels.

In this project, we propose using a State of The Art (SOTA) language model with different heuristics to reduce word error rate in lattice searching. In addition, we used a SOTA entity extraction tool along with Hidden Markov Model in identifying callsigns to improve STT accuracy.

Our results reveal that using a different language model provides a better transcription as compared to language model provided by STT toolkit. In addition, we have shown using Hidden Markov Models can improve Conditional Random Field sequence models in correcting extracted callsigns.

To advance the current work, we purpose using different language models and sophisticated Named Entity Recogniser in tandem with more training data to improve the accuracy of STT in ATC.

Decompilation of Assembly Programs using Machine Learning

In this project, we present a novel technique of using Pointer-Generator Network (PGN), a modified Sequence-to-Sequence attentional model, to decompile code into its equivalent C program. The viability and efficacy of PGN for decompilation were examined. Due to the complex and inefficient nature of traditional decompilation techniques, we present PGN as a possible alternative to traditional decompilation. This technique has the potential to eliminate problems that current compilers face such as compiler optimisations or the production of source code that is difficult for humans to read. As the PGN is trained on human-written code, the generated code from the PGN would hence be similar to human-written code. This allows the researchers to understand the logic of the program better. In this project, we trained the PGN on simple code as well as real-world code. To simplify the dataset for the neural network, various pre-processing techniques were examined. We concluded that PGN was able to decompile simple cases of assembly which it had not encountered before. When tested on real-world code, preliminary experimentation has shown that PGN was not able to decompile well. Hence, more work is needed to determine the full extent of decompilation using PGN.

Exploring Decompilation of Assembly Language using Pointer-Generator Networks

Currently traditional decompilation is done through rule-based procedures. Due to the exceptions that can arise in code, traditional decompilers must be programmed to catch these exceptions. As a result, decompilation is very complicated and the generated code is prone to error. In this report, we used a novel approach of using neural networks to decompile. The neural network was trained in assembly language to generate C code. The generated C code was evaluated using ROUGE to determine the accuracy of the neural network. Our results have shown that the generated C code shows a high level of accuracy on simple C programs containing assignment and basic operators. Hence, this experiment concluded that neural networks can do decompilation. However, in the future, more research needs to be done to determine how feasible neural networks are in decompilation on more complex C programs.