

Scraping online documentation of Ardupilot UAV system to generate a knowledge-based causal graph

Master's/Bachelor's Thesis

Supervisor: Prof. Dr. Alexander Pretschner

Advisor: Ehsan Zibaei

Email: {alexander.pretschner, ehsan.zibaei}@tum.de

Phone: +49 (89) 289 - 17885

Starting date: immediately



Fakultät für Informatik
Lehrstuhl 4
Software & Systems Engineering
Prof. Dr. Alexander Pretschner

Boltzmannstraße 3
85748 Garching bei München

Tel: +49 (89) 289 - 17885
<https://www4.in.tum.de>

Context

When a UAV crashes, it is difficult with the current technologies to identify the root cause of the failure. One approach to automate and potentially improve this procedure is to build a causal graph and trace back the events that led to the final failure. Examples of such graphs are proposed in [1]. However, the generation of such causal graphs for UAVs is still an open problem.

One method to build a causal graph is to collect cause and effect pairs mentioned by domain experts in online documentations. For example, in a forum post an experienced user argues that switching off the engines caused a UAV to lose its control and crash [2]. Similar pieces of information can be encoded as pairs of graph nodes that have a direct edge between them. By integrating those pairs, one can build a causal graph representing users' and developers' domain knowledge about the system. Fortunately, there exist open-source autopilots such as Ardupilot that own comprehensive documentation [3, 4]. Its users and developers have mentioned several cause and effect pairs regarding the UAV system in these resources.

In the first step of this thesis, standard web scraping tools [5] are utilized to extract the cause and effect pairs in the online documentation of Ardupilot UAV system and integrate them into a single causal graph. In the second step, the obtained causal graph will be validated against the already-available sanitized data-set of UAV flight logs. By looking into several UAV crashes, we will determine whether those mentioned events are observable in the data and how frequently the causation is seen in the data.

Goal

This thesis aims to automatically extract domain knowledge about Ardupilot UAVs from online natural language documents and encode it in a causal graph. In particular, we aim to answer these research questions:

1. How diverse are the mentioned cause and effect events in the online resources?
2. How consistent are the causal relationships in the online resources?
3. How detectable are such cause and effect events in the data set?
4. How accurate are the mentioned causal relationships considering the historical data?

Working Plan

1. Familiarize yourself with Ardupilot autopilot system, causal graphs, and different categories of causation in robotic systems [6]
2. Implement a workflow that scrapes the online resources and generates a causal graph
3. Measure the detectability of the graph nodes in the data
4. Measure the accuracy of the causal edges in the data
5. Write the thesis

Deliverables

- Source code of the implementation preferably deployed in Docker
- Technical report with comprehensive documentation of the implementation, i.e. design decision, architecture description, API description and usage instructions.
- Final thesis report written in conformance with TUM guidelines.

References

- [1] Ibrahim, Amjad, et al. "Practical causal models for cyber-physical systems." NASA Formal Methods Symposium. Springer, Cham, 2019.
- [2] <https://bit.ly/2XSZbRv>
- [3] <https://ardupilot.org/>

- [4] <https://discuss.ardupilot.org/>
- [5] Mitchell, Ryan. Web scraping with Python: Collecting more data from the modern web. "O'Reilly Media, Inc.", 2018.
- [6] Hellström, Thomas. "The relevance of causation in robotics: A review, categorization, and analysis." Paladyn, Journal of Behavioral Robotics 12.1 (2021): 238-255.



Fakultät für Informatik
Lehrstuhl 4
Software & Systems Engineering
Prof. Dr. Alexander Pretschner

Boltzmannstraße 3
85748 Garching bei München

Tel: +49 (89) 289 - 17885
<https://www4.in.tum.de>