Requirements and Risk Analysis

AI Engineering - Recitation 4

The World & the Machine

- Requirement gathering is the most important step in building software systems
- Ground all requirements around the World & the Machine

- Concepts:
    - **World / Environment** - The place where the system lives and manipulates
    - **Requirement** - A desired state of the world, a goal for the system
    - **Machine / Software** - Interprets and manipulates the environment as per requirement
    - **Shared Phenomenon** - Interface used by machine to manipulate the world
    - **Assumptions** - Assumed properties of the world
    - **Specification** - Actions taken by the machine to achieve a requirement

What Could Go Wrong?

- Missing / incorrect environmental assumptions
- Wrong / violated specification
- Inconsistency in assumptions and specifications / requirements
- Feedback loop: Behavior of the machine affects the world, which in turn affects input to the machine, and so on.
- Data drift: Behavior of the world changes over time, causing assumptions to become invalid
- Adversaries: Bad actors deliberately manipulate inputs / violate assumptions

Amazon Product Recommendations

- Requirements (in the world)
  - Recommend products that the user would like (and is more likely to buy)
- Specifications (for the machine)
  - Recommend highly rated products up front or higher in the list
  - Return a list of products with the same category as items in purchase history higher in the list
- Assumptions (about the world and shared phenomena)
  - ??
- Problems
  - ??

Amazon Product Recommendations

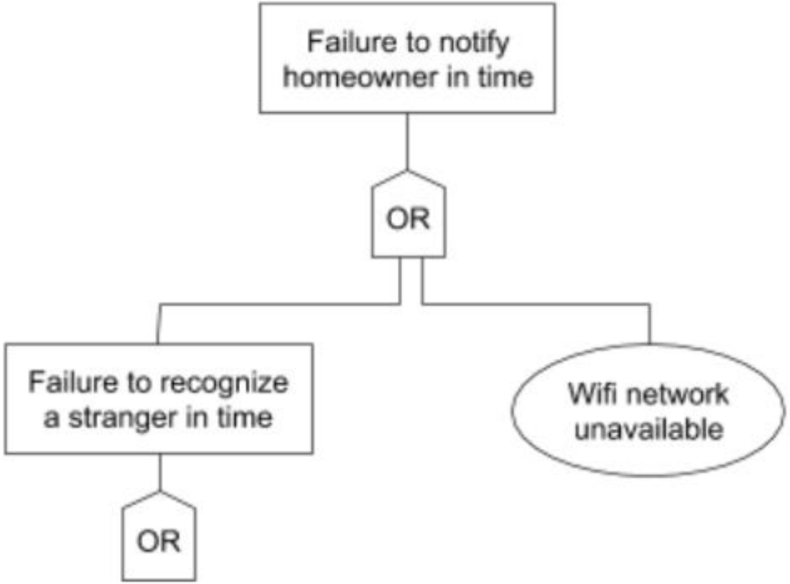
- Assumptions
  - Information about products from vendors are accurate
  - Product ratings are authentic and represent the real quality of that product
  - Products are tagged with the appropriate category by vendors
- Problems
  - What if the ratings are tampered with?
  - What if products are labeled incorrectly?
  - We recommend based on product type -> User purchases those products -> ...
  - New product / product types based on the latest trend
  - Should we recommend just based on product type?

Fault Tree Analysis

Consider a home assistant robot, that is capable of moving around obstacles in a home on its own. It has many capabilities to help around with household chores, but also has the capability to alert the head of the household (by sending a notification to their phone) if it detects that a stranger has entered the house.

Requirement:
The homeowner must be contacted in time if a stranger is present in the house

Complete the fault tree


Failure to notify homeowner in time
OR
Failure to recognize a stranger in time
Wifi network unavailable
OR

Utility for FTA:

https://github.com/troeger/fuzzed