

Jet Propulsion Laboratory
California Institute of Technology

Automated ISA Classifier

AmandaMarie Adams, George Fox University, Computer Science (BS) Information Systems (BS)
Mentor: Grant Faris

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the Space Grant and Summer Internship Program and the National Aeronautics and Space Administration.

Introduction to the PRS

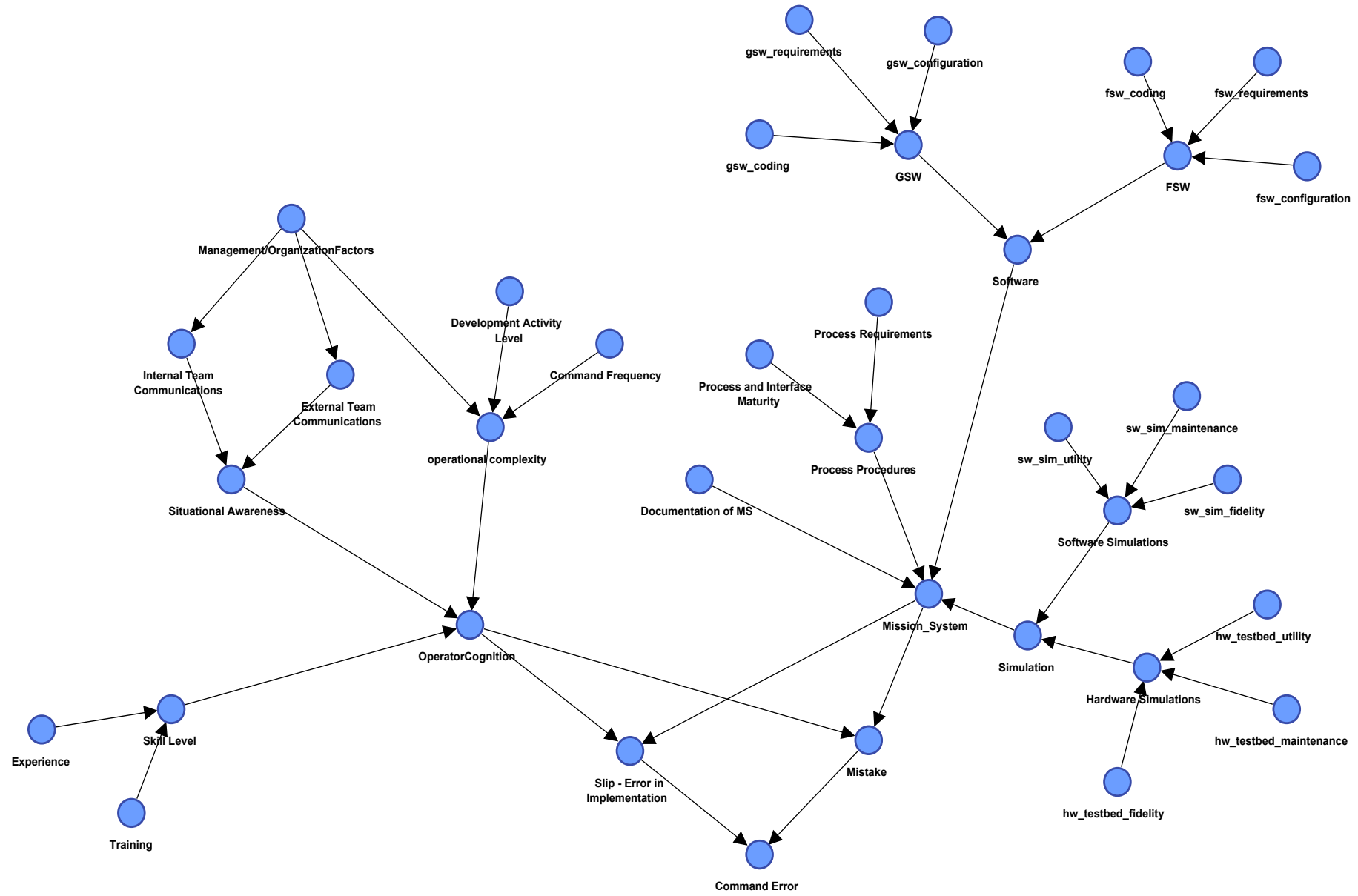
50,000+ records

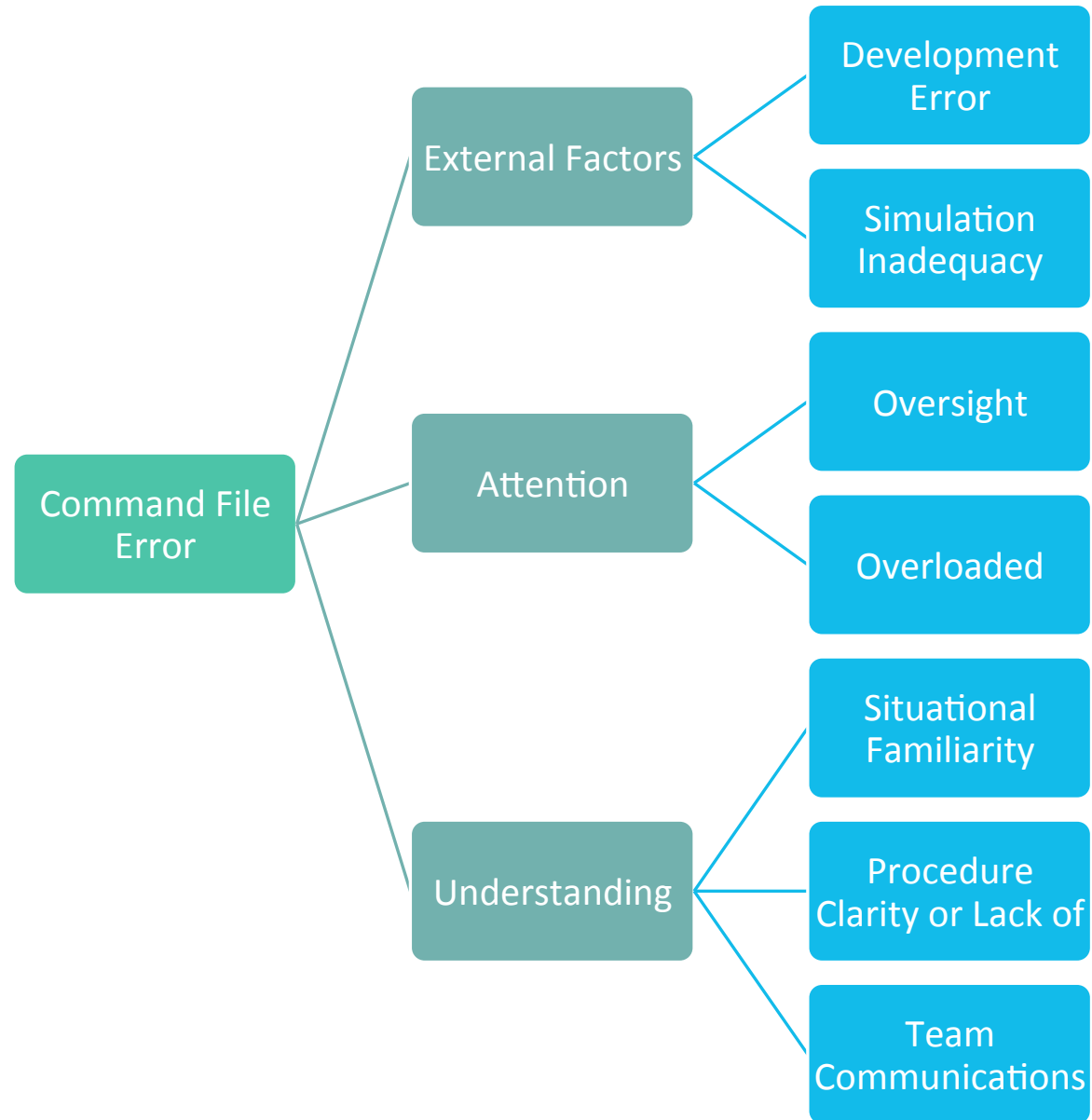
Benefits of PRS

Command File Error (CFE) according to DocID 35506, Rev. 7 Appendix B, Para 3.8.1:

- An error in a command file that was sent to the spacecraft.
- An error in the approval, processing, or uplink of a command file that was sent to the spacecraft.
- The omission of a command file that should have been sent to the spacecraft.

Root cause classification





Creating the Classifier

Approach 1: Word counts and frequencies

Approach 2: NLTK and Wordnet

- How does the brain classify?
- Subject-fault verb phrases

```
>>> wn.synsets('dog')
[Synset('dog.n.01'), Synset('frump.n.01'), Synset('dog.n.03'), Synset('cad.n.01'),
Synset('frank.n.02'), Synset('pawl.n.01'), Synset('andiron.n.01'), Synset('chase.v.01')]
>>> wn.synsets('dog', pos=wn.VERB)
[Synset('chase.v.01')]
```

```
>>> dog = wn.synset('dog.n.01')
>>> dog.hypernyms()
[Synset('canine.n.02'), Synset('domestic_animal.n.01')]
>>> dog.hyponyms()
[Synset('basenji.n.01'), Synset('corgi.n.01'), Synset('cur.n.01'), Synset('dalmatian.n.02'), ...]
>>> dog.member_holonyms()
[Synset('canis.n.01'), Synset('pack.n.06')]
>>> dog.root_hypernyms()
[Synset('entity.n.01')]
>>> wn.synset('dog.n.01').lowest_common_hypernyms(wn.synset('cat.n.01'))
[Synset('carnivore.n.01')]
```

Creating the Classifier

Approach 3: Weighted words

“updated procedure” vs “transposition”

Sureness level

High: 97% correct

Moderate: 93% correct

Low: 77% correct

NEI (Not Enough Information):

- Solely technical language: there is no verbiage that points to human or machine fault. The only way to categorize this kind of data is to have someone who is involved in the project classify it. Requires in-depth understanding of project and incident in report.
- Sparse text: the report is short and does not contain enough verbiage to classify it.

Classifier Evolution

	DAWN CFEs (1)	DAWN CFEs (2)	MSL CFEs (3)
Overall Accuracy	67%	84%	78%
Incorrect	14%	9%	10%
Not Classified	19%	7%	12%

Table 1 Classifier statistics for classifying: DAWN CFEs using text analysis from DAWN CFEs (1) and for DAWN CFEs and MSL CFEs (2); MSL CFEs using text analysis from DAWN CFEs and MSL CFEs (3).

19/22 are NEI

	DAWN CFEs	MSL CFEs
Overall Accuracy	97%	96%
Incorrect	3%	2%

Table 2 Classifier statistics for classifying: DAWN CFEs and MSEL CFEs using text analysis from DAWN CFEs and MSL CFEs. Percentages are calculated based on data that is not marked as NEI.

Improving Reports

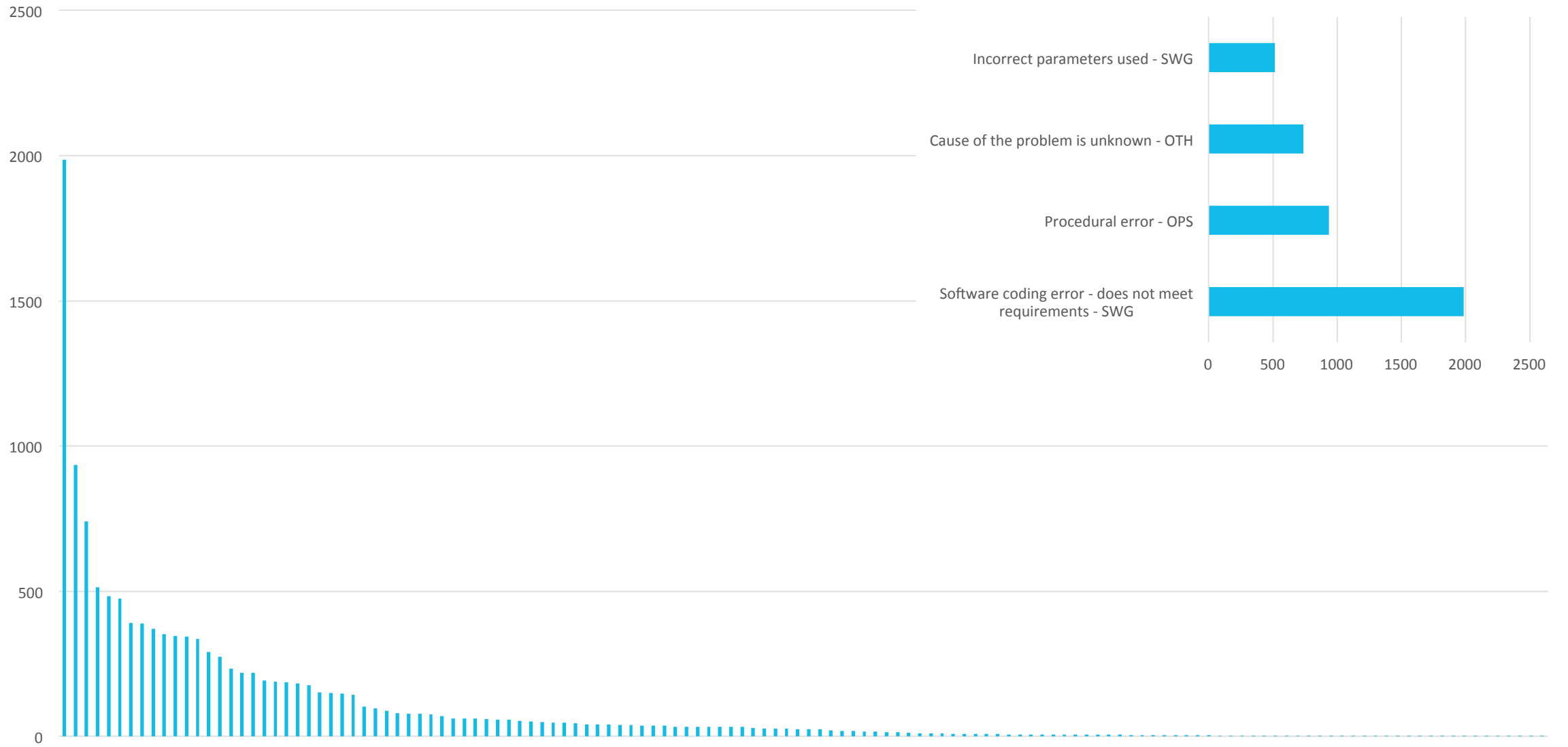
Training:

- Include subject-fault verb phrases (minimal needed)

Modifying PRS Reporting Form:

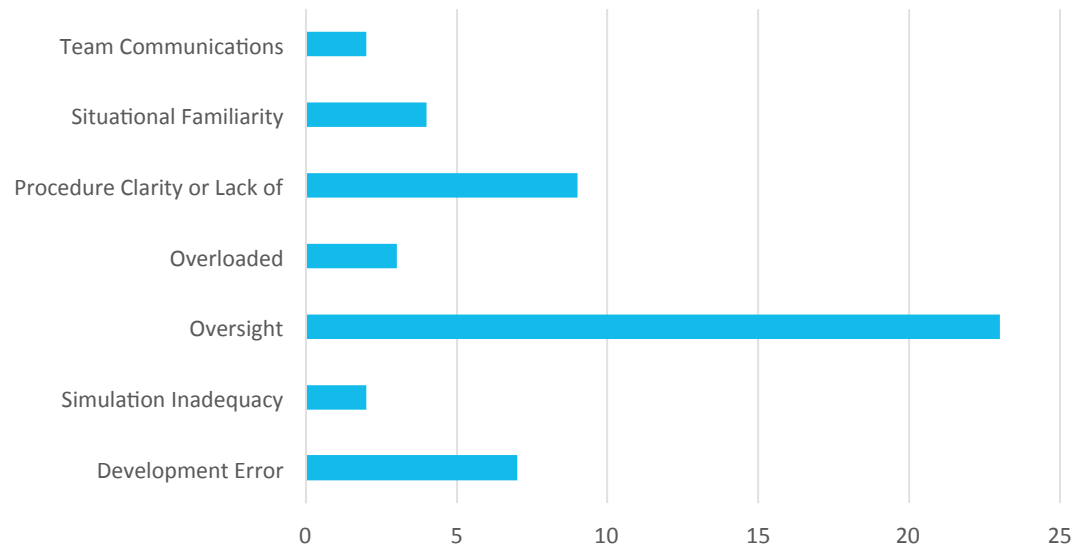
- Cause code categories
 - 130+ choices
 - 100+ are used <100 times
 - 13,000/50,000 use them
 - Dynamically created depending on lifecycle

Distribution of Cause Codes



Results

DAWN CFE Root Causes



MSL CFE Root Causes

