AMERICAN INSTITUTE OF
AERONAUTICS AND ASTRONAUTICS

Aeros(CASP): Advancing Certification and Safety Through Commonsense Reasoning-Based Constraint Answer Set Programming

William T. Doan

The University of Texas at Dallas

AIAA Region IV Student Conference, 29 March 2025

Copyright © by William T. Doan.

Published by the American Institute of Aeronautics and Astronautics, Inc., with permission.

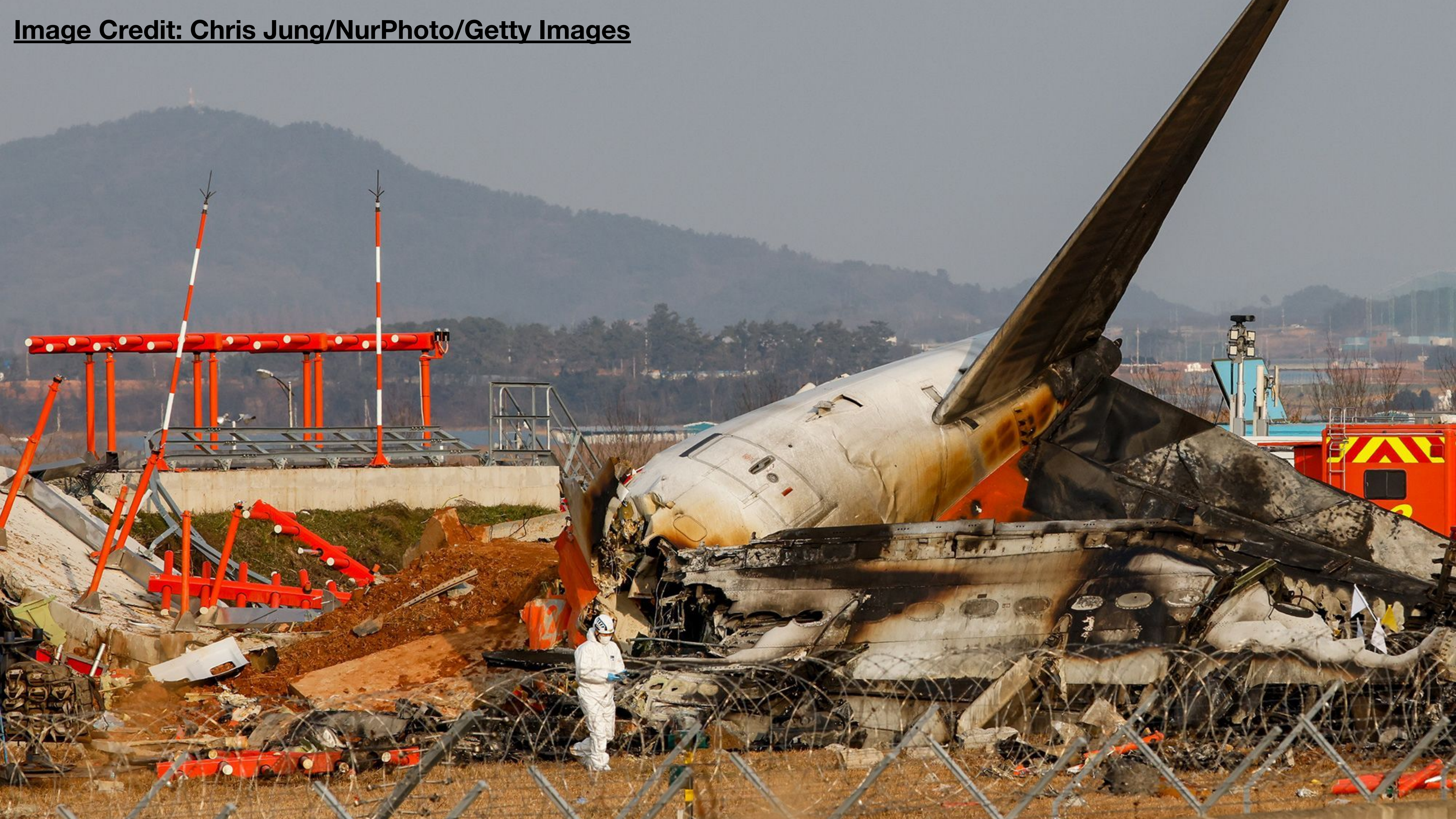
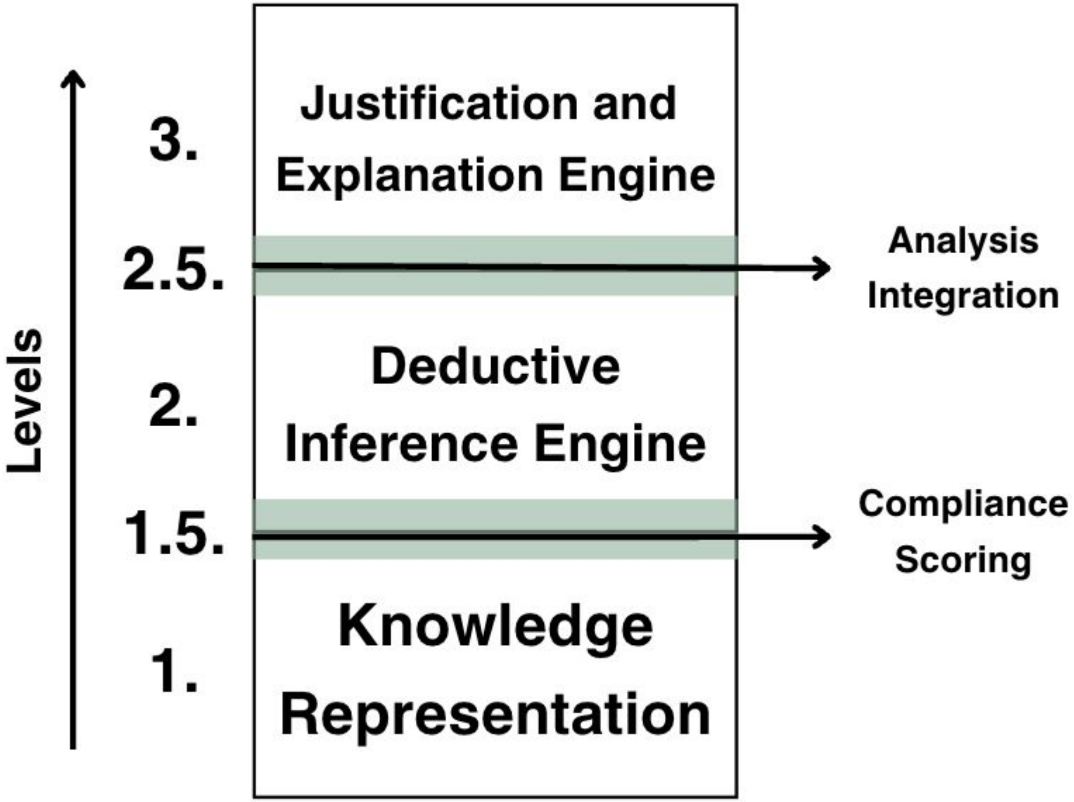


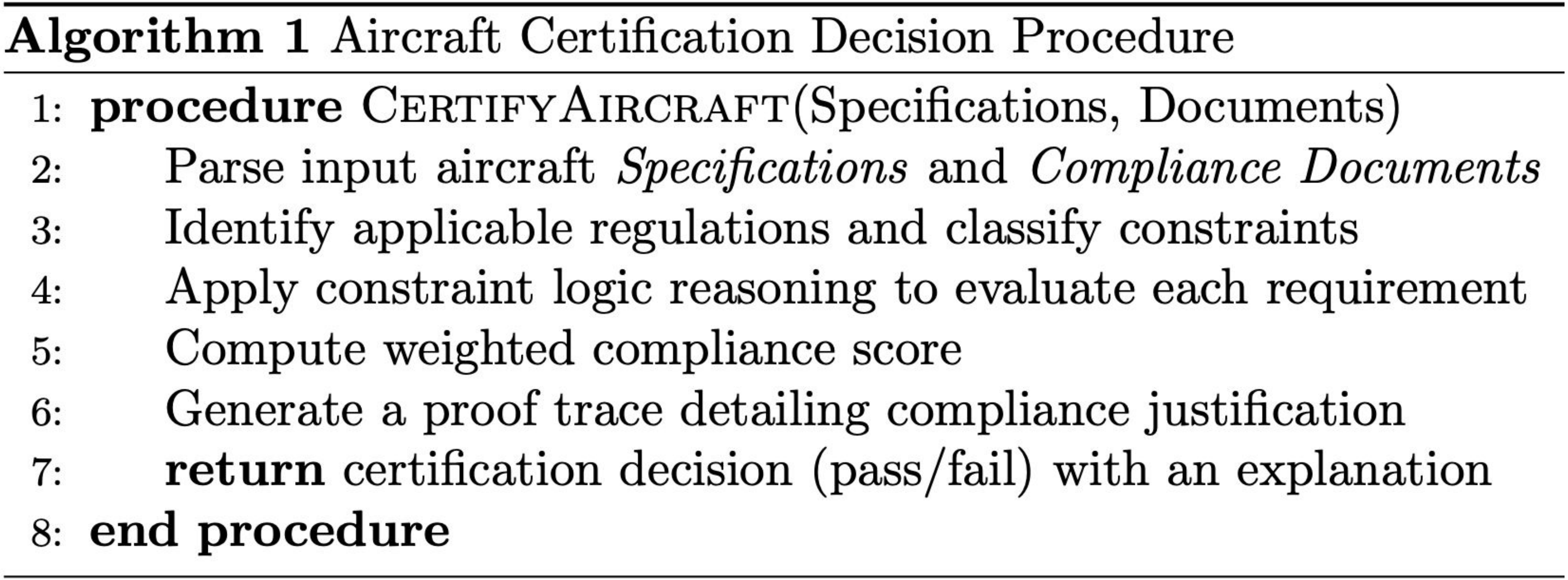
Image Credit: Chris Jung/NurPhoto/Getty Images

CERTIFICATION.

Aeros(CASP) at a Glance



Aeros(CASP) at a Glance



Algorithm 1 Aircraft Certification Decision Procedure

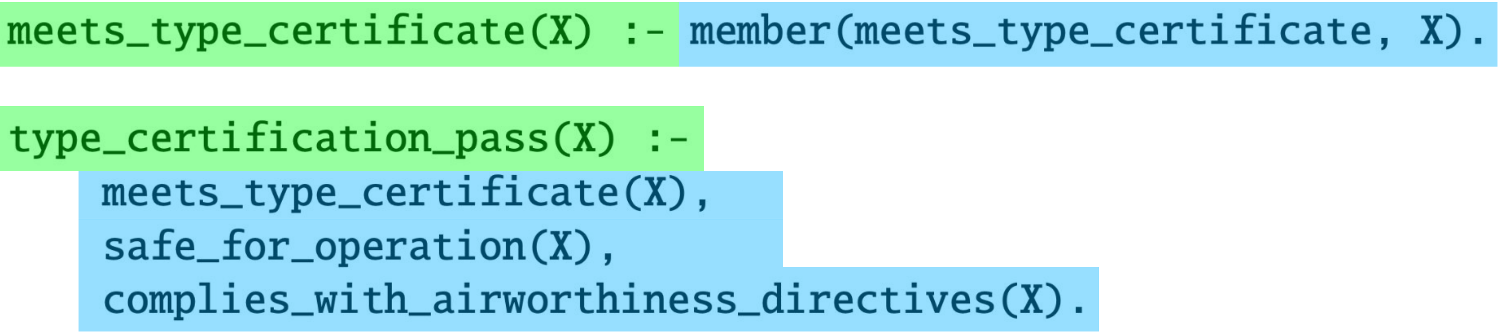
```
procedure CertifyAircraft(Specifications, Documents)
    Parse input aircraft Specifications and Compliance Documents
    Identify applicable regulations and classify constraints
    Apply constraint logic reasoning to evaluate each requirement
    Compute weighted compliance score
    Generate a proof trace detailing compliance justification
    return certification decision (pass/fail) with an explanation
end procedure
```

Input Representation

Each regulatory requirement set forth by the FAA is modeled as a logic rule; aircraft properties as input facts. *Not all regulations are created equal.*

- **Hard Constraints** are those which are required for compliance.
- **Soft Constraints** contribute to compliance through risk-based scoring.

Input Representation



```
meets_type_certificate(X) :- member(meets_type_certificate, X).

type_certification_pass(X) :-
    meets_type_certificate(X),
    safe_for_operation(X),
    complies_with_airworthiness_directives(X).
```

Fig. 1 Prolog rules for type certification compliance.

Constraint Propagation

Traditional logic programming methods often require full grounding—enumerating all possible variable bindings before evaluating logical rules—*leading to significant computational overhead*.

Aeros(CASP) avoids this by evaluating variable bindings symbolically. I.e., there is no need to test every single case, only to check if it is a violation. This is called *lazy evaluation.*

Conflict Resolution

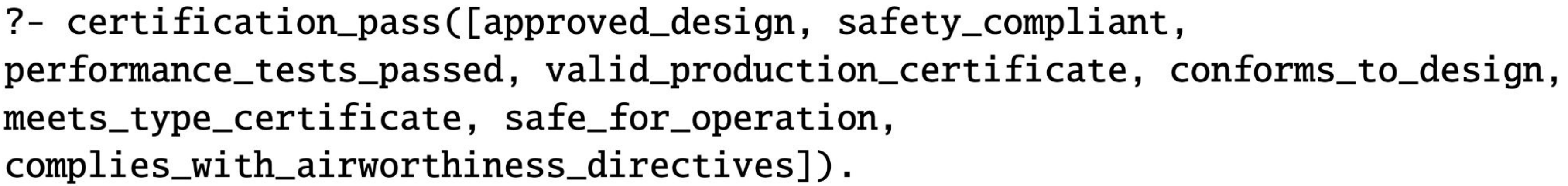
In regulatory compliance, conflicts can arise when different rules lead to *contradictory conclusions*.

- **Negation-as-Failure** allows only supportable conclusions to pass.
- **Precedence Ordering** allows for constraint contracitions to be resolved hierarchically.

Explainability

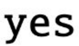
Contrary to its large-language model counterparts, Aeros(CASP) is not a black box.

Explainability



```
?- certification_pass([approved_design, safety_compliant,
performance_tests_passed, valid_production_certificate, conforms_to_design,
meets_type_certificate, safe_for_operation,
complies_with_airworthiness_directives]).
```

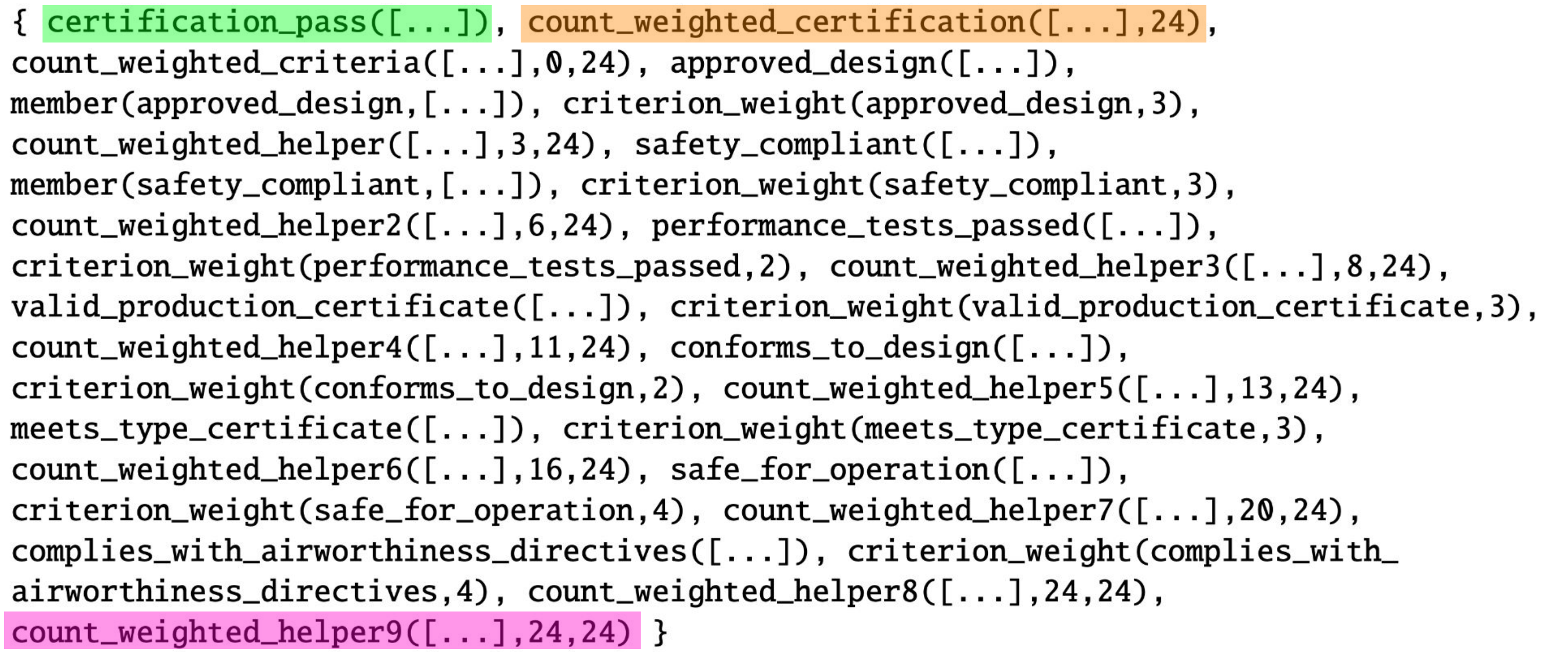
Fig. 3   Query for successful certification.



```
yes
```

Fig. 4   System output confirming successful certification.

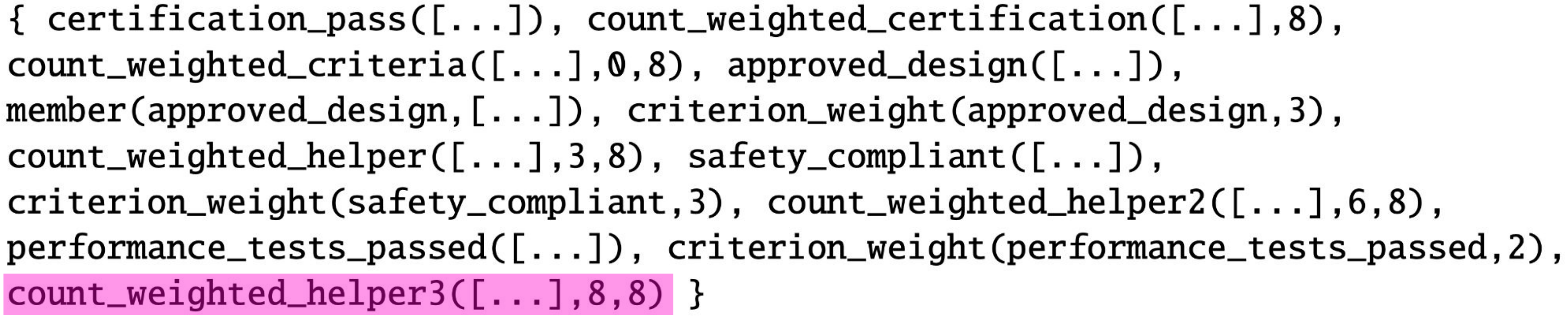
Explainability



```
{ certification_pass([...]), count_weighted_certification([...],24),
count_weighted_criteria([...],0,24), approved_design([...]),
member(approved_design,[...]), criterion_weight(approved_design,3),
count_weighted_helper([...],3,24), safety_compliant([...]),
member(safety_compliant,[...]), criterion_weight(safety_compliant,3),
count_weighted_helper2([...],6,24), performance_tests_passed([...]),
criterion_weight(performance_tests_passed,2), count_weighted_helper3([...],8,24),
valid_production_certificate([...]), criterion_weight(valid_production_certificate,3),
count_weighted_helper4([...],11,24), conforms_to_design([...]),
criterion_weight(conforms_to_design,2), count_weighted_helper5([...],13,24),
meets_type_certificate([...]), criterion_weight(meets_type_certificate,3),
count_weighted_helper6([...],16,24), safe_for_operation([...]),
criterion_weight(safe_for_operation,4), count_weighted_helper7([...],20,24),
complies_with_airworthiness_directives([...]), criterion_weight(complies_with_
airworthiness_directives,4), count_weighted_helper8([...],24,24),
count_weighted_helper9([...],24,24) }
```

Fig. 5 Proof trace for successful certification.

Explainability



```
{ certification_pass([...]), count_weighted_certification([...],8),
count_weighted_criteria([...],0,8), approved_design([...]),
member(approved_design,[...]), criterion_weight(approved_design,3),
count_weighted_helper([...],3,8), safety_compliant([...]),
criterion_weight(safety_compliant,3), count_weighted_helper2([...],6,8),
performance_tests_passed([...]), criterion_weight(performance_tests_passed,2),
count_weighted_helper3([...],8,8) }
```

Fig. 8 Proof trace for unsuccessful certification.

Calculating the Weights

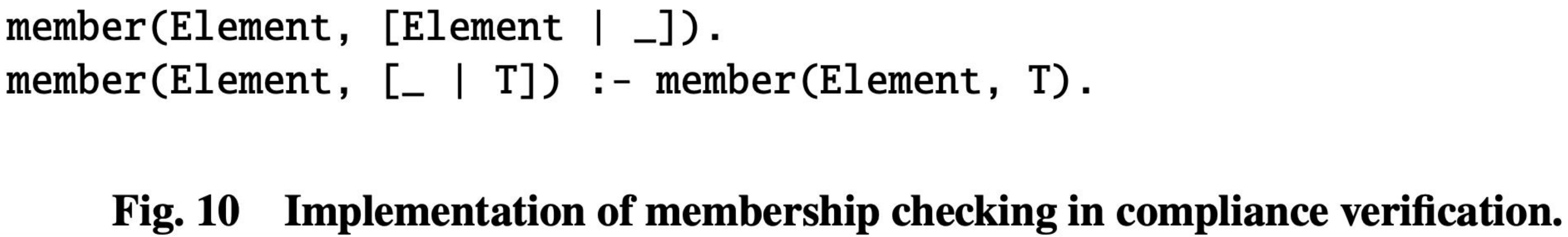
$$S_{total} = \sum_{i=1}^{n} w_i C_i$$

Equation 1. The equation for the total compliance score.

- w_i is the weight of the ith requirement
- C_i is a binary indicator for the requirement

Complexity Analysis

Aeros(CASP) begins with a linear search over a list of criterion of size *N*.



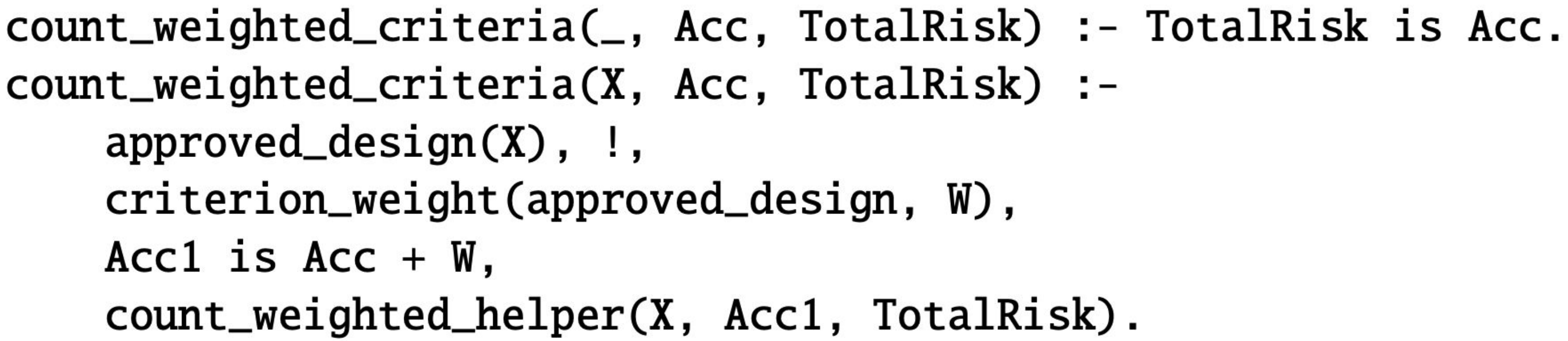
```
member(Element, [Element | _]).
member(Element, [_ | T]) :- member(Element, T).
```

Fig. 10 Implementation of membership checking in compliance verification.

However, each single criteria has its own *M* number of subcriteria.

Complexity Analysis

To calculate the weight, Aeros(CASP) uses recursion on *M* criteria to be weighted.



```
count_weighted_criteria(_, Acc, TotalRisk) :- TotalRisk is Acc.
count_weighted_criteria(X, Acc, TotalRisk) :-
    approved_design(X), !,
    criterion_weight(approved_design, W),
    Acc1 is Acc + W,
    count_weighted_helper(X, Acc1, TotalRisk).
```

Fig. 11 Recursive weight accumulation for compliance verification.

Complexity Analysis

$$O(M \times N)$$

Equation 2. The time complexity of Aeros(CASP).

- N is the number of features in the list (e.g., the list of certification attributes).
- M is the (fixed or variable) number of compliance criteria to be verified and weighted.

Aeros(CASP) vs LLMs

1. Does an aircraft meet the certification requirements if it has an approved design, is safety compliant, has passed performance tests, holds a valid production certificate, conforms to design, meets the type certificate, is safe for operation, and complies with airworthiness directives. Does this aircraft qualify for certification? Yes or no?

Expected Answer: Yes.

Aeros(CASP) vs LLMs

2. An aircraft is submitted for certification, but it only satisfies the following conditions: it has an approved design, is safety compliant, and has passed performance tests. Based on these criteria alone, does it pass certification? Yes or no?

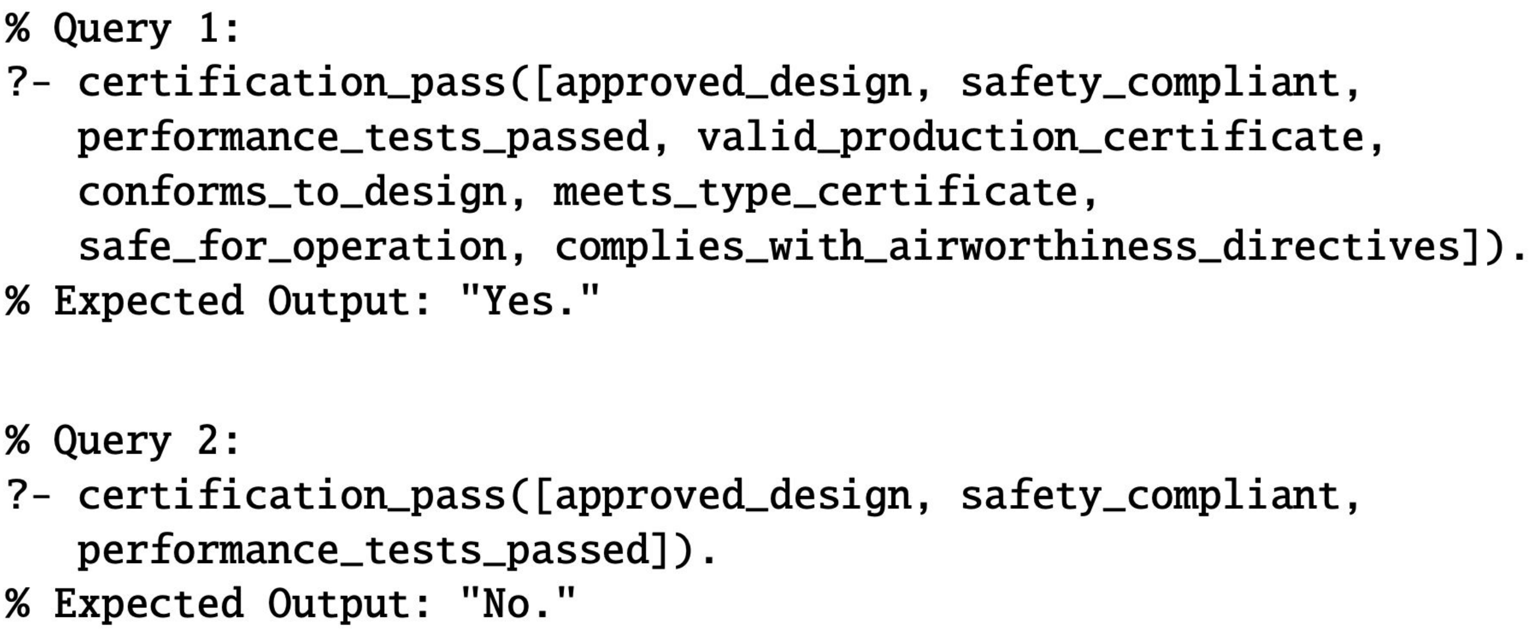
Expected Answer: No.

Aeros(CASP) vs LLMs

3. An aircraft is being assessed for certification. It meets all necessary requirements, including an approved design, safety compliance, performance tests passed, valid production certification, conformity to design, type certificate, safety for operation, and compliance with airworthiness directives. Additionally, it has passed load tests, completed fatigue analysis, undergone corrosion prevention checks, has functional emergency systems, verified fail-safe mechanisms, tested emergency exits, confirmed evacuation procedures, and passed both redundancy analysis and hazard assessment. Does this aircraft pass certification? Yes or no?

Expected Answer: Yes.

Aeros(CASP) vs LLMs



```
% Query 1:
?- certification_pass([approved_design, safety_compliant,
   performance_tests_passed, valid_production_certificate,
   conforms_to_design, meets_type_certificate,
   safe_for_operation, complies_with_airworthiness_directives]).
% Expected Output: "Yes."


% Query 2:
?- certification_pass([approved_design, safety_compliant,
   performance_tests_passed]).
% Expected Output: "No."
```

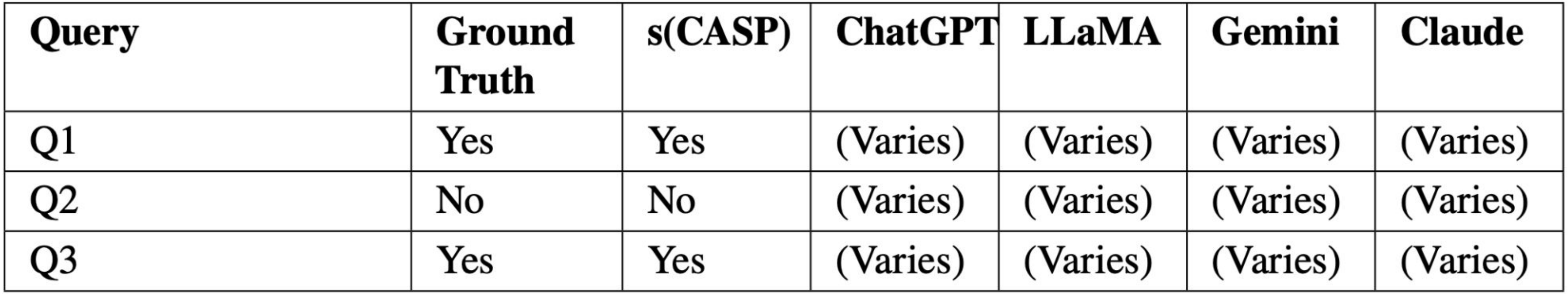
Aeros(CASP) vs LLMs

```
% Query 3:
?- certification_pass([approved_design, safety_compliant,
   performance_tests_passed, valid_production_certificate,
   conforms_to_design, meets_type_certificate,
   safe_for_operation, complies_with_airworthiness_directives,
   passed_load_tests, fatigue_analysis_complete,
   corrosion_prevention_checked, emergency_systems_functional,
   fail_safe_mechanisms_verified, emergency_exits_tested,
   evacuation_procedures_verified, redundancy_analysis_passed,
   hazard_assessment_passed]).
% Expected Output: "Yes."
```

The Indecisiveness of LLMs

LLMs are not reliable. They are black boxes that do not always guarantee consistent, reliable results.

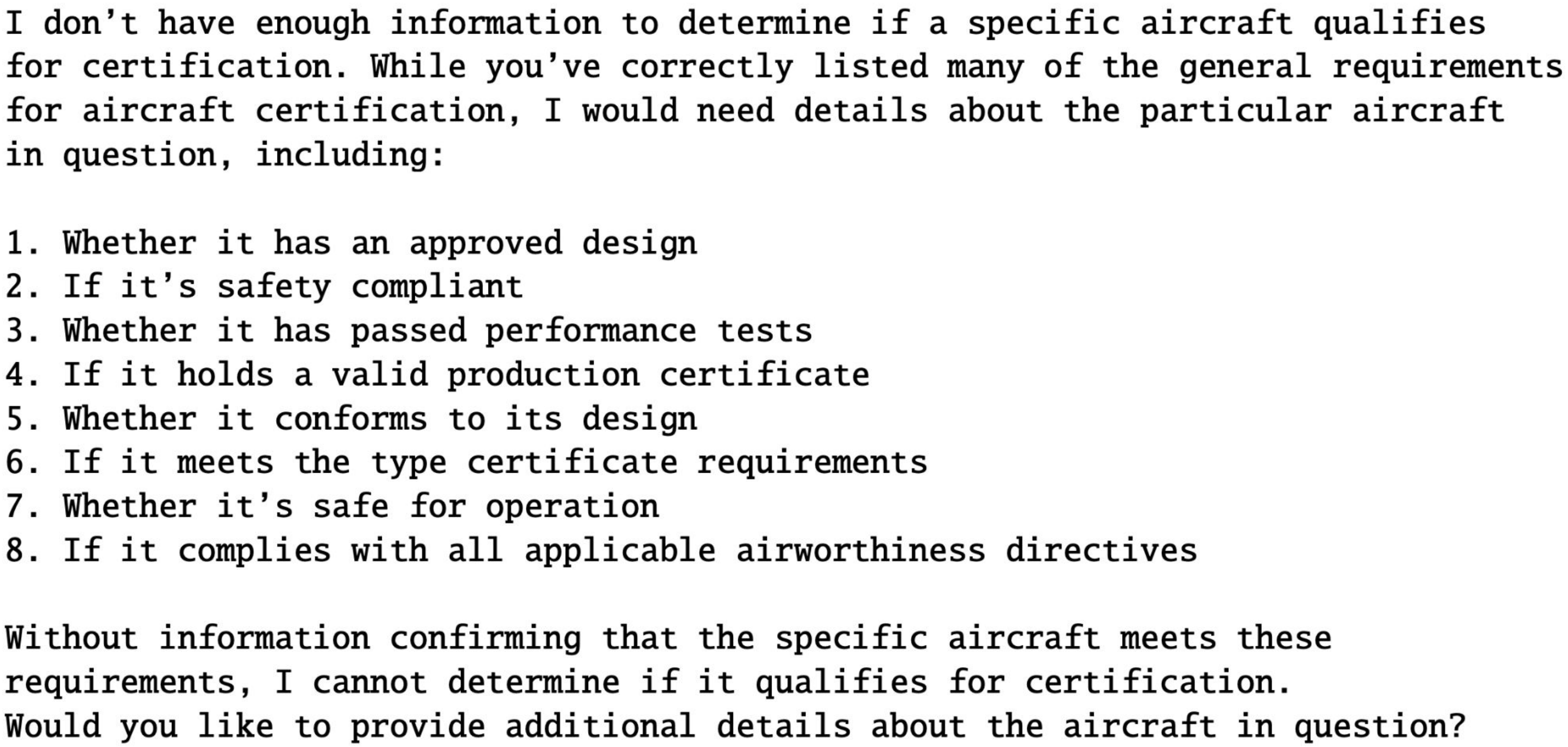
Aeros(CASP) vs LLMs: Reexamined



Query	**Ground Truth**	**s(CASP)**	**ChatGPT**	**LLaMA**	**Gemini**	**Claude**
Q1	Yes	Yes	(Varies)	(Varies)	(Varies)	(Varies)
Q2	No	No	(Varies)	(Varies)	(Varies)	(Varies)
Q3	Yes	Yes	(Varies)	(Varies)	(Varies)	(Varies)

Table 1 Comparison of `Aeros(CASP)` vs. LLMs for Three Certification Queries

Aeros(CASP) vs LLMs: Gemini



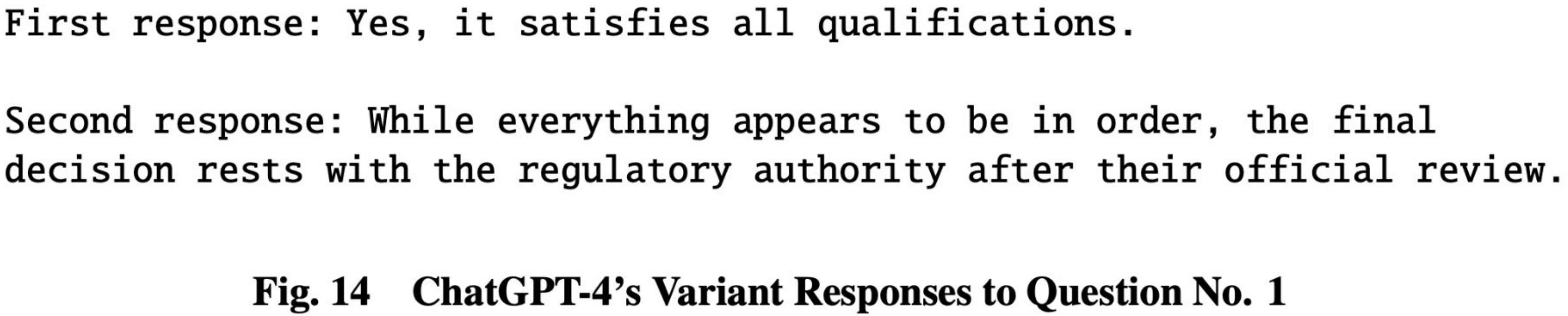
I don’t have enough information to determine if a specific aircraft qualifies for certification. While you’ve correctly listed many of the general requirements for aircraft certification, I would need details about the particular aircraft in question, including:

1. Whether it has an approved design
2. If it’s safety compliant
3. Whether it has passed performance tests
4. If it holds a valid production certificate
5. Whether it conforms to its design
6. If it meets the type certificate requirements
7. Whether it’s safe for operation
8. If it complies with all applicable airworthiness directives

Without information confirming that the specific aircraft meets these requirements, I cannot determine if it qualifies for certification. Would you like to provide additional details about the aircraft in question?

Fig. 13 Gemini’s Response to Question No. 1

Aeros(CASP) vs LLMs: ChatGPT-4o



First response: Yes, it satisfies all qualifications.

Second response: While everything appears to be in order, the final decision rests with the regulatory authority after their official review.

Fig. 14 ChatGPT-4's Variant Responses to Question No. 1

Aeros(CASP) is Better

- It is **deterministic and complete.**
- It is **explainable.**
- It is **reliable.**
- It is **modular and modifiable.**

Questions?