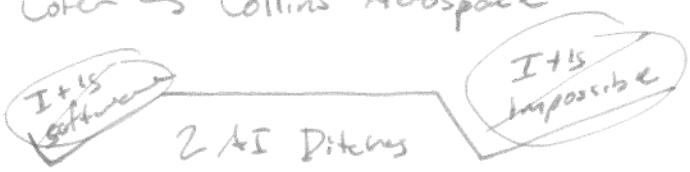
NASA Formal Methods Conference

June 11th 2025

- William and Mary is 2nd oldest ~~college~~ university? in US

Keynote - "AI in the Sky"

- Darren Cofer → Collins Aerospace



- Pilot shortages → Reduced staffing → more AI assistance?
- AI can be used from embedded systems to pilot monitoring to cabin monitoring
- Coding co pilots - documentation, test generation
- Regulators have determined current processes cannot prove compliance for AI/ML systems
- Extremely improbable failure → 10^{-9} probability per Flight hour.
- Neural networks are deterministic → but complexity makes them viewed like probabalistic...
  - Loss of 10^{-9} does not indicate 10^{-9} failure rate.
  - No way to test!
- Some approaches
  1. Embrace probabalism → huge verification set
  2. Improve redundancy
  3. Reject Probabalism → embrace formal methods
  4. Bound behavior of probabalistic requirements.

"Verify Generalization of a NN w/ FM"

- Run Time Assurance
  - ↳ AADL → Assume Guarantee model
    - ↳ Assurance argument graph
  - ↳ Counter examples found → can change training!

- Cautionary Tales podcast

THEOREM PROVING

Process Algebraic Semantics fo. Verifying Intelligent Robotic Control Software

didn't pay attention

Reuseable For Verif of DAG-Based Consensus Protocols

- TLA+
- what the fuck do any of these graphs mean??
  - update rthunk they're states.
- I really should've learned what a DAG is.
- Basically main contribution is providing building blocks for TLA+ for DAG based protocols

Verification of Anti-Unification Algorithm in PVS

Category: Methods
Context: Unification replaces variables with other expressions to make to computations the same

Clarity: Reasonably Clear. A lot of info in 33 slides.
Correctness: Have no clue.
Contributions: They made some functions to prove anti-unification of something.

VERIFICATION

Missed 1st One

Form. Verf. of Composite Field Multipliers

Category: New Method
Context: Basically Japan is trying to make antennas in space w/ a bunch of small satellites
Correctness: IDK
Contributions: They came a way to prove correctness of signals.
Clarity:

Refused

2/8

Formal Verification As A Service of PLCs

- From CERN
- PLCVerif
  - Didn't get into any specifics

Vellvm: Formalizing the Informal LLVM

- Compiler Correctness
- Compiled code $\stackrel{?}{\leq}$ regular code.
- Compiler maybe kills verification of correctness
- C/Rust → LLVM → Binary
- LLVM is a huge language that complicates verification
- Formalization has lead to insights about the language
- "axiomitized?"

SAT/SMT Solvers

MAXSat

A maxsat problem has soft constraints and hard constraints
- Soft constraints get optimized
- Hard constraints must always be satisfied

An SMT Foundation of Mixed Precision Multiplication

- Basically Hardware suppliers don't say how they implement precision
- They used SMT solvers to discover how certain hardware is actually performing operations.

TRACE: Toolkit for Requirements, Capture, and Elicitation

- Guy from CE Aerospace basically presenting a lecture on TRACE, a platform to write requirements, specifications and equations to do requirement analysis.

Thursday June 12

Keynote - Cedar

Category: New Prog Language
Context: Access control is really hard to get right. Cedar provides a new language that does authorization
Contributions: Cedar → a high assurance language for authorization
Legible, expressive, fast, and analyzable
Clarity: Emina Torlak is a great speaker
Correctness: Formal ♥ Written in Rust

- ☐ What is a DAG?
- ☐ Differential Random Testing ↔ Check language congruence

8 days to learn LEAN and 20k LOC for proofs... We're fucked.

Hybrid Automata

Size Reduction Thru Procedure Finding

Category: Methods
Context: Large automata are difficult to understand and implement
Contributions: Combine states thru marriage → this reduces automaton size. Save a lot of FPGA gates
Clarity: Eh I really don't know...
Correctness: Unable to identify

Mode Based Reactive Synthesis ☐ FOLLOW UP

☐ What is an atomic property?
Category: Method
Context: Reactive systems are those that take an input from the environment and produce an output
Synthesis is very challenging.
Contribution: Fracture controllers into modes. Modes are state predicates. Breaking into modes can make the problem easier to solve.
Clarity: I stopped understanding after a bit. No mention of plant or control objective.
Correctness: Unable to identify

MECHANIZED RS274 SEMANTICS FOR ADDITIVE MANUFACTURING

Category: ???
Context: RS274 is line oriented execution (G code...)
Codes often can be viewed as nodes - G90/G91 → local/absolute coordinates
Contributions: ~~Checks to make sure that~~ Freeze G code (RS274) to be analyzed. Pipes between written G and M codes and formal methods tools
Clarity: Literally no clue what's going on
Correctness: Unable to Identify

So much text. Literally no idea what the "Who cares?" is.
- ☐ What is mechanizing?

HYBID SYSTEMS

HyTWIN ☐ Purdue MODEL — So Loud!

Category:
Context: ICS systems are getting attacked all the time. Digital twins are use to determine if an adversary has attacked a system.
Contribution: ~~Some definitions talking about how a hybrid system w/ Digital twin monitor could work~~ Describes Attacks and mitigations as hybrid programs in Differential Dynamic Logic
Correctness: Honestly very suspect.
Clarity: Very obtuse. Very loud.

Bad. What happens if i spoof same value → no detection horizon?

~~Explainable Dynamic Logics w/ 1st Class Reasoning~~
Category: New Tool
Context: LTL for simple systems, ~~dL~~ dL for ΔF
Contributions: Expand dL with a relational extension to make things easier to verify
Correctness: This guy knows his shit
Clarity: Very well spoken, great explanation

5/8

RARE EVENT SIMULATION

Category: Simulation method
Context: Rare events are hard to capture in simulations
Contributions: Importance splitting assigns states an importance based on proximity to decision points. Makes hybrid simulation easier because you get more data near transition points.
Correctness: Looks good 2 me!
Clarity: Well spoken, a little cluttered but not too bad.

DISCRETE SYSTEMS

ELIMINATING FLAKINESS

Category: New testing Method
Context: Flakiness is when a test is not deterministic
Contribution: The "Rich" statement. Basically an implement in software testing that allows code to randomly select a value for a parameter, and simplify the following code to look for errors.
Correctness: Looks good to me!
Clarity: Pretty clear and easy to understand

QUERYING LABELED TIME SERIES DATA

□ What is a formal scenario?

Category: New software tool
Context: Simulation is a standard practice to test cyber-physical systems, but sensor realism yields sim failures that can't be realized IRL.
Contribution: Scenic querying time data. Basically a system where they can search for similar video data from real life that meets similar properties
Correctness: Compared to mixed model LLMs - Nice results
Clarity: Very well spoken!

ALGORITHMIC ANALYSIS OF EVENT B IN REWRITING LOGIC

Category:
Context: Event B is a FM based on set theory and first order logic, refinement based processes. Can be probabalistic
Contribution:
Correctness:
Clarity: I have no clue. This was straight math. 6/8

JUNE 13

~~KEYNOTE~~ ~~NASA & FORMAL METHODS~~

- ☐ Check out PVS
- ☐ Check out FRET
  - ↳ link to Simulink
- Autonomy does not require AI. We've done autonomy since the 60's
- AI/ML really lack safety assurances

☐ What is an F1 score?

~~RUNTIME MONITORING~~

~~TWO SAFE VERIF. RUNTIME MONITR - R2U2~~

Category: Towards Paper

Context: On-line runtime monitors constantly validate the integrity of a controller. R2U2 gets used as an online runtime monitor

Contribution: R2U2 in embedded Rust from C. Formall defn ~~past~~ past-time ~~Mission~~ Mission Time Temporal Logic ($p^{+\text{MTTL}}$ ~~MTTL~~)

Made optimizations

- ☐ Verus → Rust encoder verifier
  - ↳ Doesn't do floats

Correctness: Partially Formal

Clarity: Awesome, and they had a gif!

~~ENFORCING MAVLINK SAFETY & SECURITY PROPERTIES~~ (ARTAUD!)

Category: Experimental result

Context: MAVLink is a critical protocol, ~~for~~ drones, but ~~Contribution~~ has a lot of limitations, ~~of~~ bugs.

Valid AND unsure

Contribution: Datum, runtime verification for communications

Datum is a guard for parameter manipulation

Correctness: Very cool

Clarity: Assertion Evidence, great visuals 10/10

VISUALIZING TEMPORAL INTERVAL HEIARCHIES

Category:

Context: Logs are everywhere. Nfer helps process logs

~~Contribution~~

Contribution: Interval heirarchies can be hard to interpret. This tool makes logs easier to visualize.

Correctness:

Clarity

Temporal Logics

A Streamlined, Formal Approach to Requirement Based Testing

Category: New Method
Context: Formal methods for requirements testing is too hard.
Contributions: A Restricted English framework for writing requirements based tests. Can send reqs to CoCoSim.
Correctness: Formal?
Clarity: Lots of words. Recently spoken. NASA Ames

Language Partitioning for MLTL

Category:
Context: Requirements come from a long list and a big team. Elicitation is not easy! Now how do we process traces?
Contribution: Lacey-Rart, a function that takes in a MLTL formula and returns coformulas that are partitions of the language.
Correctness: Formal
Clarity: Really well spoken, and covers past and future operators with Lacey-Rart well

Closing Remarks

37% Acceptance rate
NFM 2026 at JPL in Ca