37th SafeTRANS Industrial Day — Submission

Speaker: Tom Bienmüller, BTC Embedded Systems AG

Title: Keeping the Chain of Trust Closed: Formal Methods for AI-Generated Safety Evidence in Continuously Evolving Systems

Abstract:

Generative AI is becoming a productivity engine in embedded software development: it drafts specifications, writes code, and derives tests far faster than engineers can. For safety-critical systems, however, speed without evidence is worthless — a safety argument cannot rest on artifacts that are merely plausible. Frequent updates sharpen this problem: the faster AI produces and the more often systems change, the wider the gap between what has been generated and what has been assured.

We argue that trustworthiness must be established as an unbroken chain: at every point where AI generates a safety-relevant artifact, formal methods complement it and make it a trusted artifact, turning "the AI probably got it right" into machine-checkable evidence that feeds directly into the safety case and standards such as ISO 26262. With high automation, this guarantee keeps pace with AI's speed. We illustrate this with two use cases.

Use Case 1 — Trusted AI Code Generation: AI-generated code is only as trustworthy as its specification. We separate two trust concerns, both discharged by automated formal analysis. First, the specification is made a trusted "golden" reference — formal methods detect inconsistency, incompleteness, and ambiguity. Second, the generated code is formally verified against this golden spec.

Use Case 2 — Requirement-Based System-Level Testing: AI generates test cases from requirements, and formal specification plus formal checking assures that the tests faithfully and completely reflect the specified behavior.

We present the current state of both use cases and close with an outlook on the challenges that remain when this chain of trust has to be kept closed across continuous updates.