A Framework for Evidence-Based Licensure of Autonomous Systems

David Tate

Christopher Martin, Frank Moses, David Sparrow

Institute for Defense Analyses

Work sponsored by AFRL and OASD(R&E)

5 July 2016

We know TEVV* for autonomous systems will be hard

Can't test exhaustively

Can't statistically sample

System learning can invalidate past results

*Test, Evaluation, Verification and Validation

Evidence Based Licensure (EBL) might be a solution

Define **dependability cases**

Accumulate **evidence**

Construct explicit **dependability arguments**

Establish **third-party confidence**

License for use within **defined limits**

Form 25

6390

Department of Public Safety - Bureau of Building Inspection

CITY OF PHILADELPHIA

Elevator Operator's License

Olga Villari

having been duly examined and having qualified in accordance with the Ordinance of Council, Ordinance Revised, June 11, 1940, is hereby licensed to operate a Passenger Elevator for one year from date of issue.

Geo. U. Siegrist

CHIEF

DATE OF ISSUE Oct. 23, 1944

“Dependability” means everything we care about

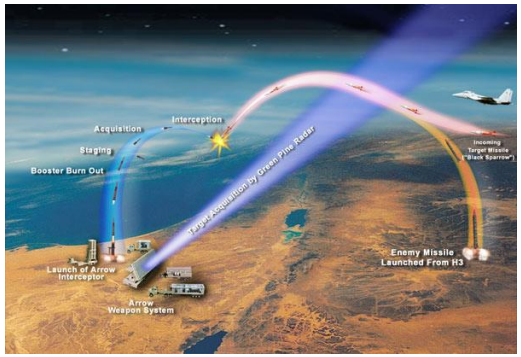
Mission performance

Safety

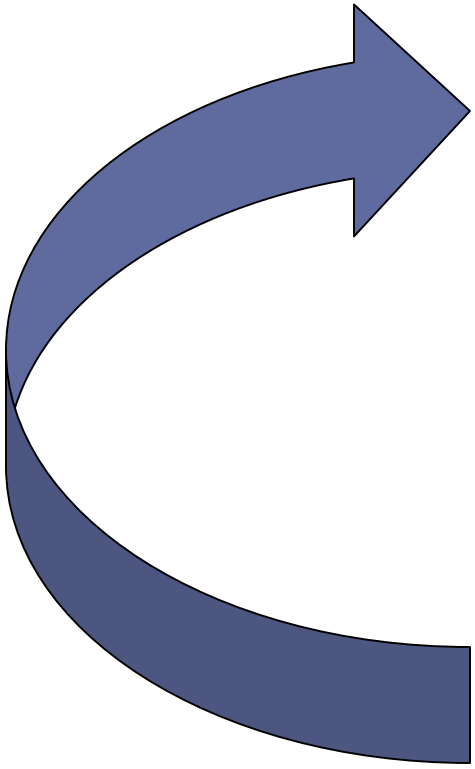
Security

Reliability

⋮



Proposed TEVV birth-to-retirement framework



Formal Methods

- Architecture & Design phase

Normative Oracles

- Full life cycle

Scalable Virtual Testbed

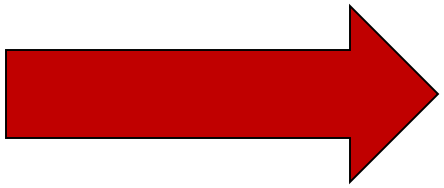
Run-time Monitors

- Development phase

Fielding with Limits

Ongoing Recertification

Today we're focusing on this part



- Formal Methods
  - Architecture & Design phase
- **Normative Oracles**
  - Full life cycle
- Scalable Virtual Testbed
- Run-time Monitors
  - Development phase
- Fielding with Limits
- Ongoing Recertification

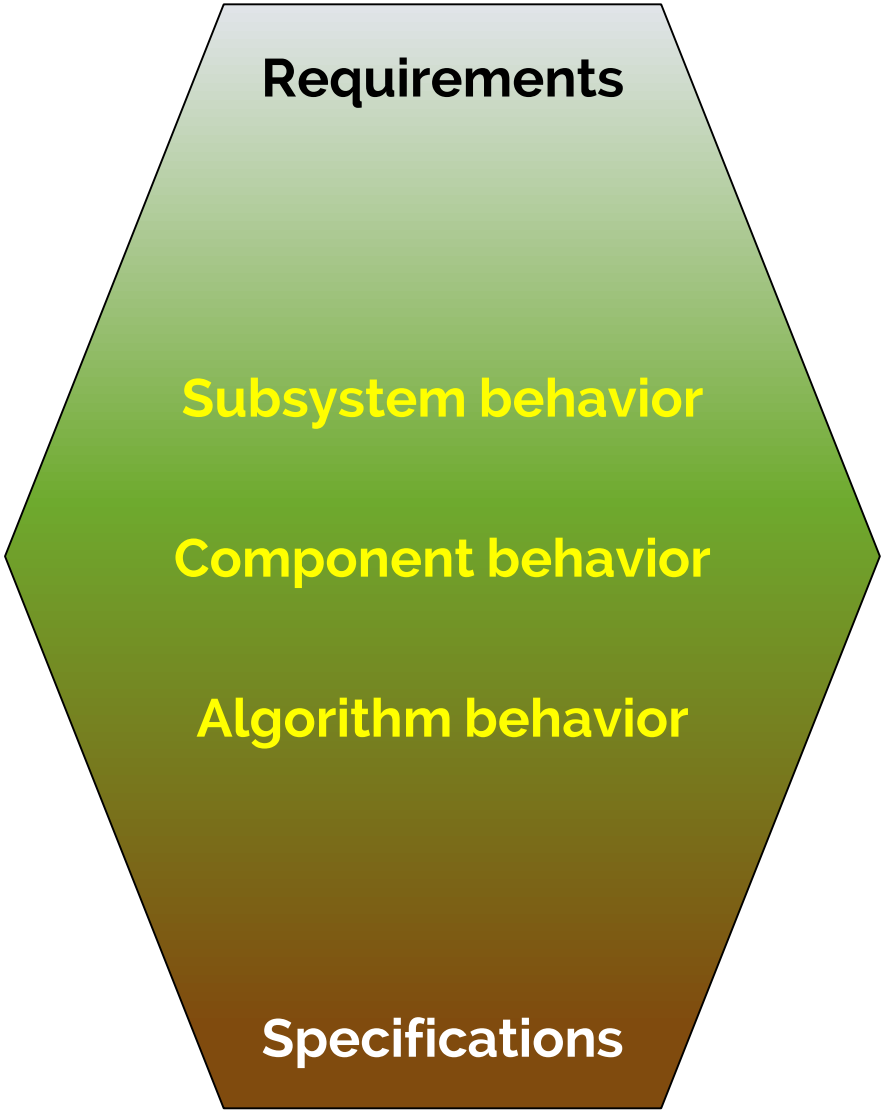
What is a “normative oracle”?

Oracles answer questions.

Normative oracles answer questions about whether behavior is correct / appropriate / desirable

Is the system behaving **as it ought**?

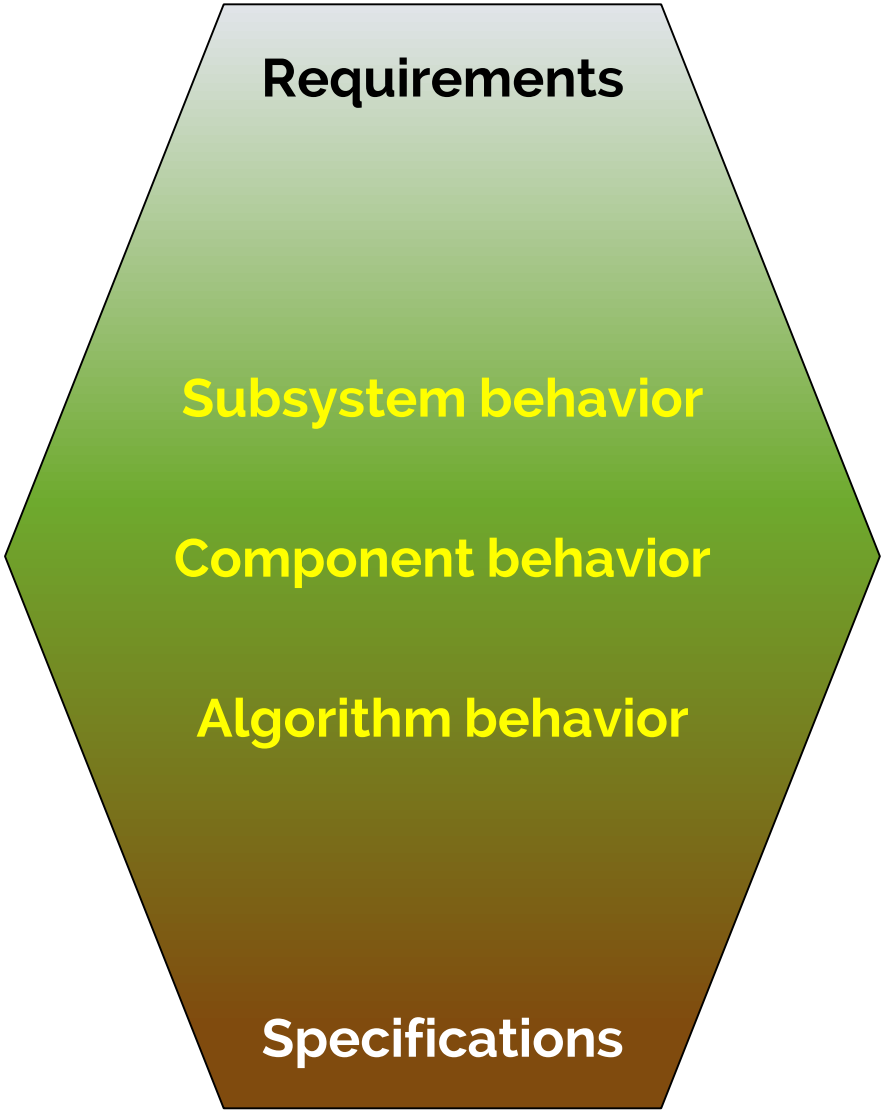
Different oracles are appropriate for different levels



High-level oracles

Design-independent
Derived from requirements
(including safety etc.)

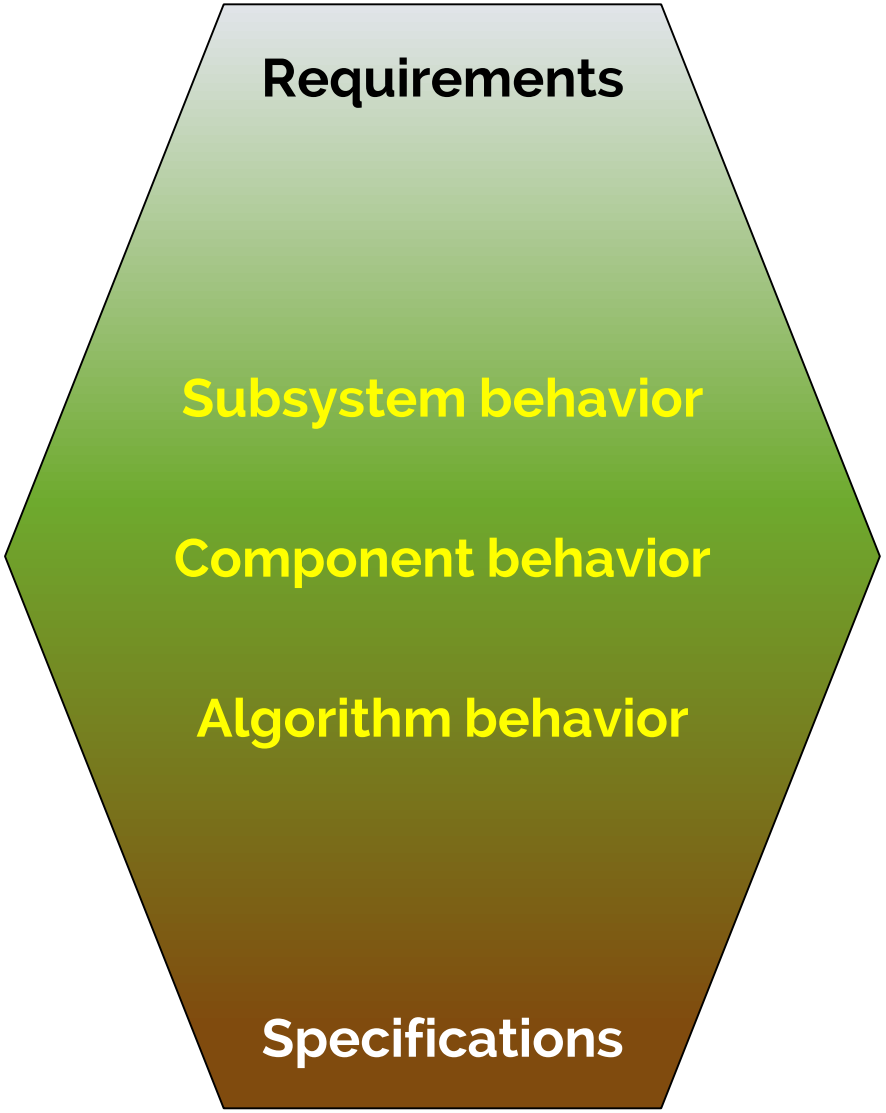
Different oracles are appropriate for different levels



Low-level oracles

- Test instrumentation
- Design-dependent

Different oracles are appropriate for different levels



Mid-level oracles

SME notions of what success looks like

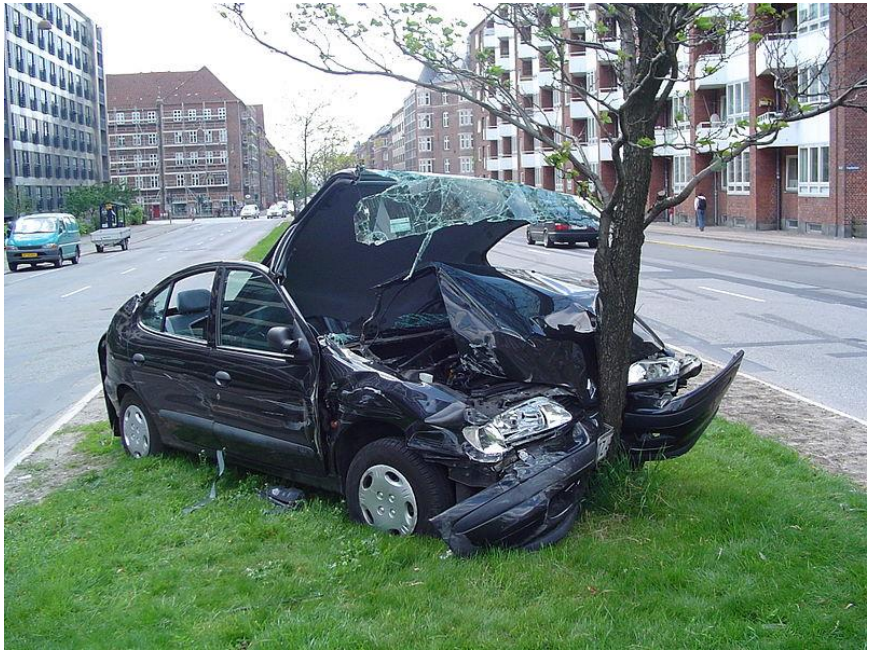
Some design-dependent

Many required

Example: Self-Driving Car

High level oracle:

Don't crash into trees



Example: Self-Driving Car

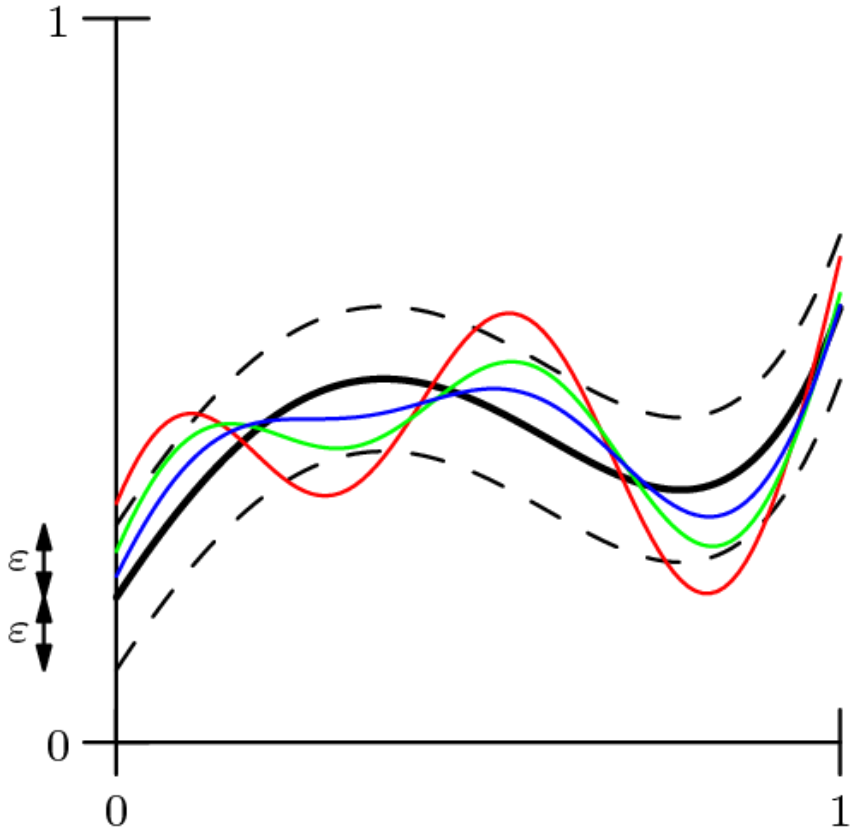
Subsystem oracle:

Motion should not make passengers uncomfortable

Example: Self-Driving Car

Algorithm oracle:

Weights in NN prediction of lead car behavior should converge, not oscillate



Example: Self-Driving Car

Low level oracle:

self-assessed
speed ±0.3 kph
vs ground truth

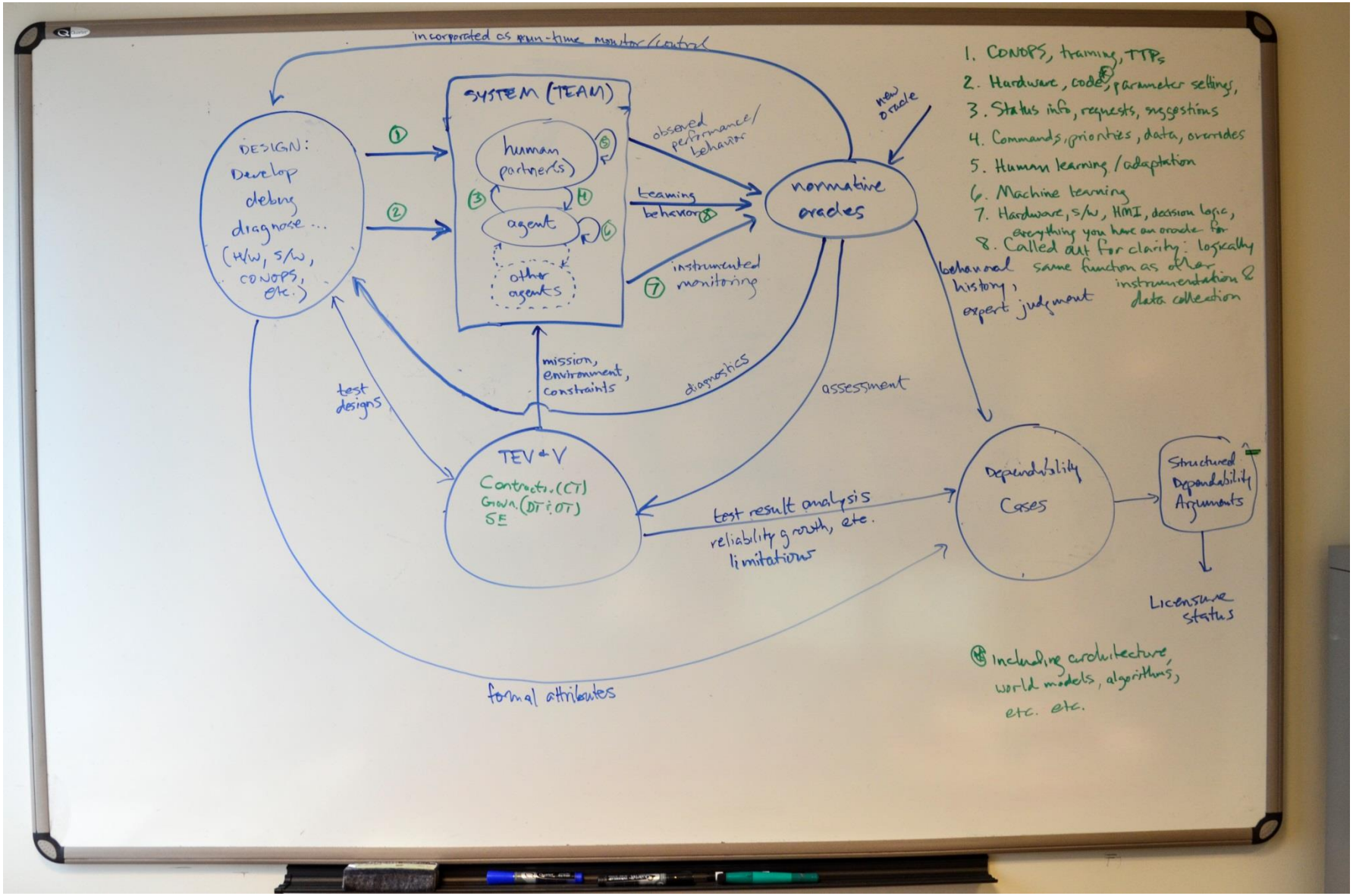
Humans are part of the system, too

If operations require human-machine teaming, you will need explicit oracles for

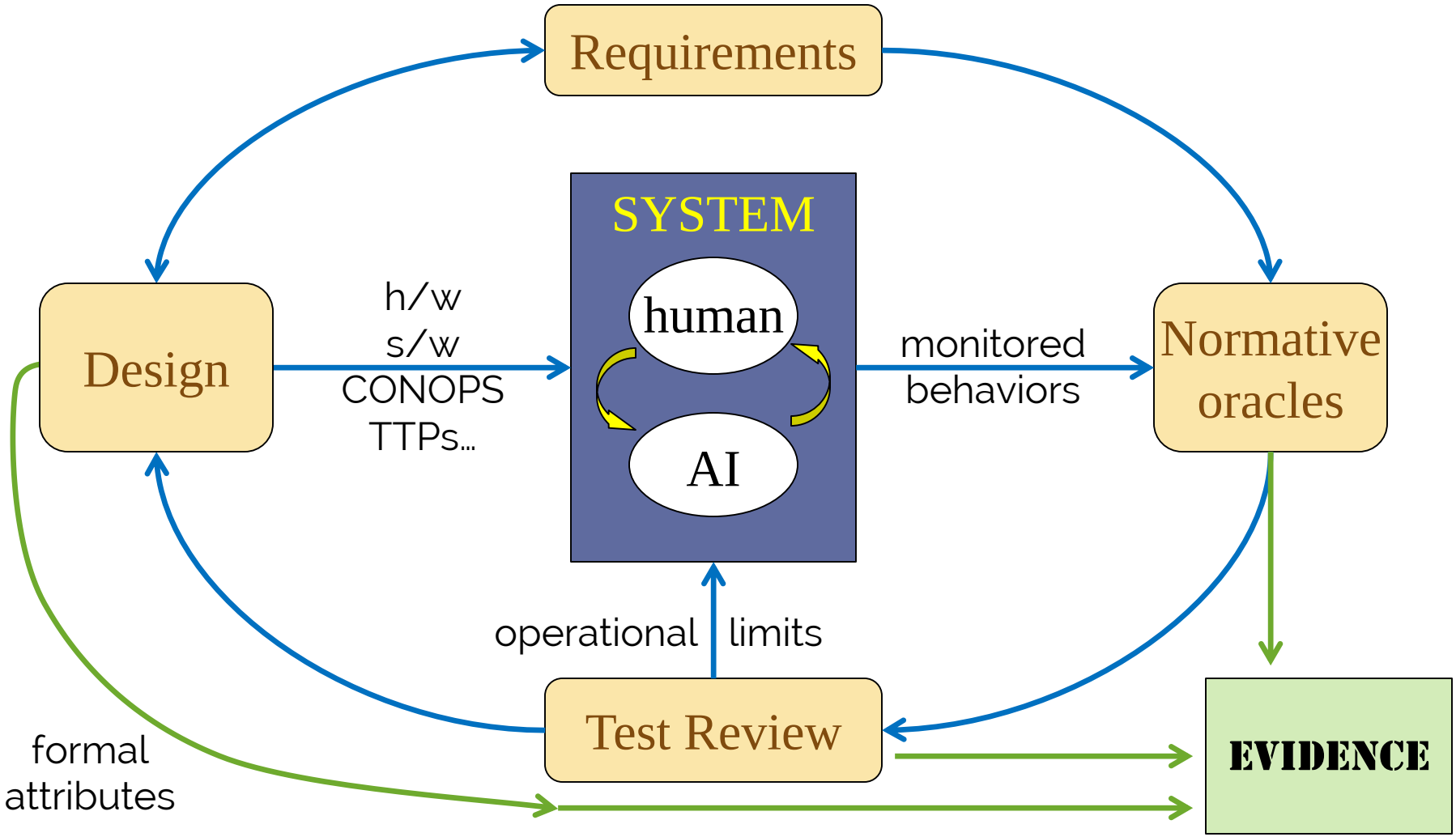
1. Machine behavior (internal and external)
2. Human behavior (likewise?)
3. Performance of the human-machine interaction

Coactive design seems like a good approach, and a source of oracle definitions

It's really quite simple…



Here's the bare bones



What does this approach change?

Not much at highest- and lowest-levels
(oracles are part of current SE and DT practice)

Mid-level oracles will be SME labor-intensive:
definition
implementation
interpretation
argument generation

What does this effort buy us?

The **time series of performance** against the oracles provides a **richer body of evidence** toward potential licensure than simple pass/fail testing

Quantifies **robustness**, based on history of behavior in novel situations

Supports **partial licensure** by identifying **operational bounds** within which performance is most dependable, **evolving over time** pre- and post-fielding

Summary

For EBL to be successful, third-party licensing bodies will have to be confident of system dependability

Explicit dependability arguments for autonomous systems will need more compelling evidence than pass/fail testing can provide

Normative oracles will be key to developing the time series of evidence that supports confident fielding

For a much more detailed description…

A Framework for Evidence-Based Licensure of Adaptive Autonomous Systems

IDA Paper P-5325

March 2016

https://www.ida.org/idamedia/Corporate/Files/Publications/IDA_Documents/STD/2016/P-5325.ashx

Questions?

Gory details – evidence accumulates iteratively