Source: NASA, https://www.nasa.gov/image-feature/scott-kelly-on-the-second-spacewalk-of-expedition-45

Presented by Timothy E. Wang and Prof. Pierluigi Nuzzo

Timothy E. Wang*
Chanwook Oh[x]
Matthew Low[x]
Isaac Amundson#
Zamira Daw%
Alessandro Pinto^
Massimiliano Chiodo*
Guoqiang Wang*
Saqib Hasan#
Ryan Melville*
Pierluigi Nuzzo[x]

*: Raytheon Technologies Research Center
x: University of Southern California
#: Collins Aerospace
^: NASA Jet Propulsion Laboratory
%: University of Stuttgart

Computer-Aided Generation of Assurance Cases

SafeComp 2023, SASSUR 2023

September 19th-22nd, 2023

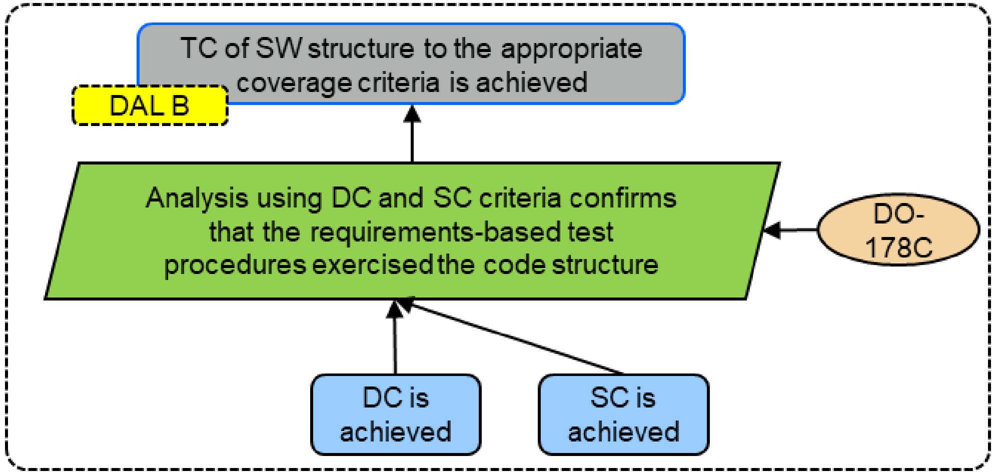
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Motivation

"Moving from current process to an argument-based process…"

What is an assurance case?

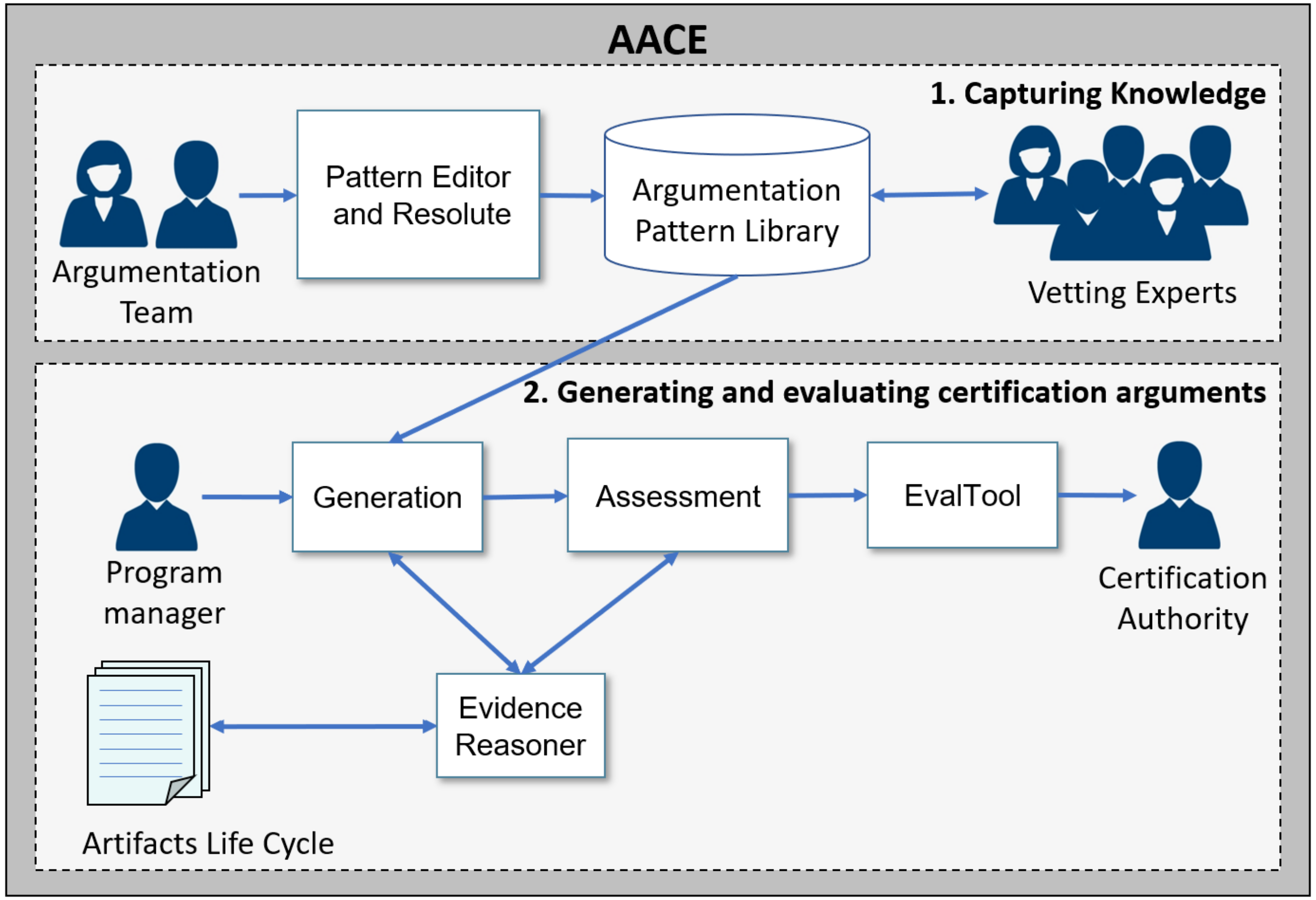
Assurance cases (ACs) are explicit arguments that desired properties (safety, security, standards compliance,..) have been adequately established for a given system.


TC of SW structure to the appropriate coverage criteria is achieved
DAL B
Analysis using DC and SC criteria confirms that the requirements-based test procedures exercised the code structure
DO-178C
DC is achieved
SC is achieved

Assurance case elements: Claim, argument, and evidence or other sub-claims

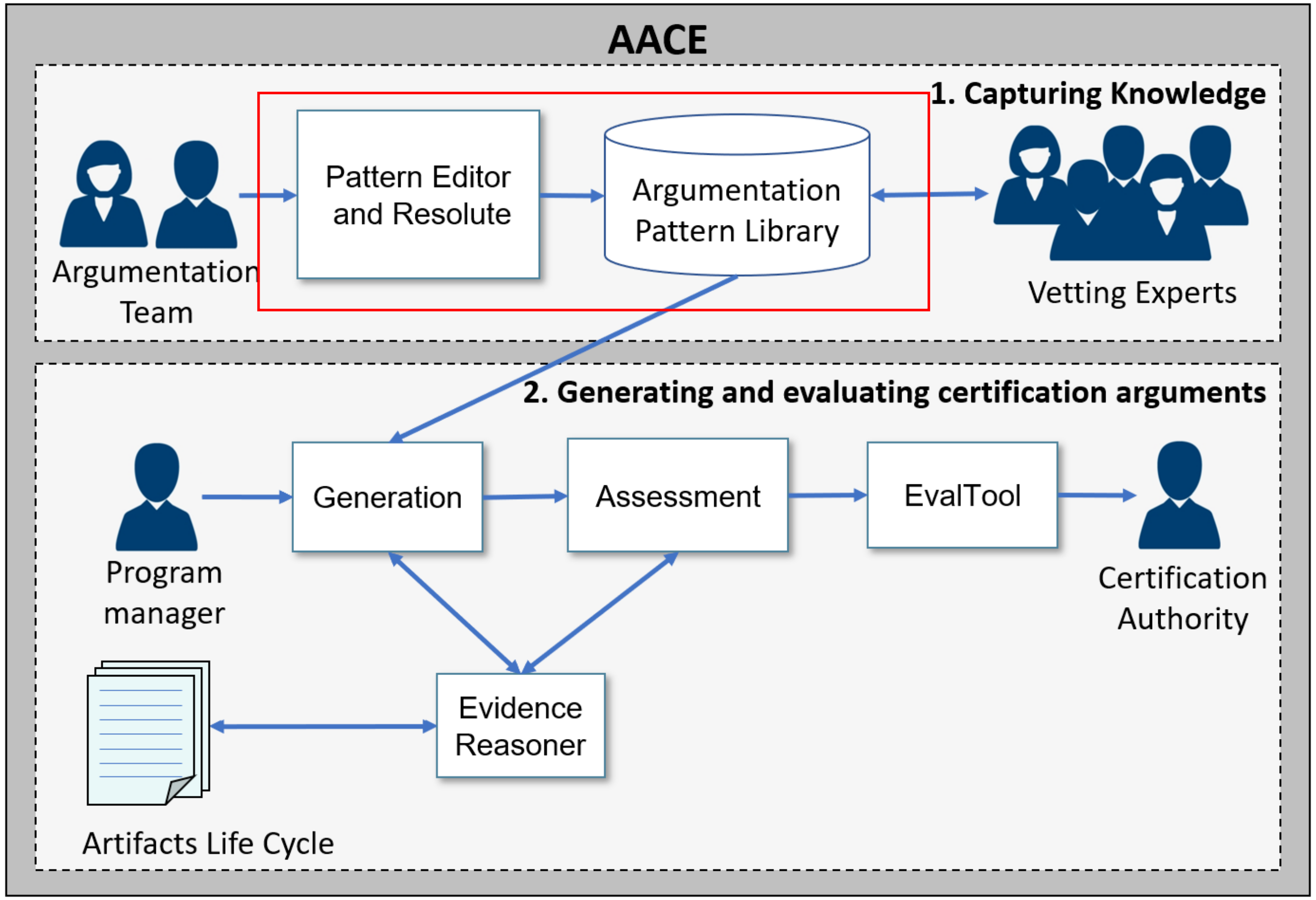
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

AACE: Automated Assurance Case Environment



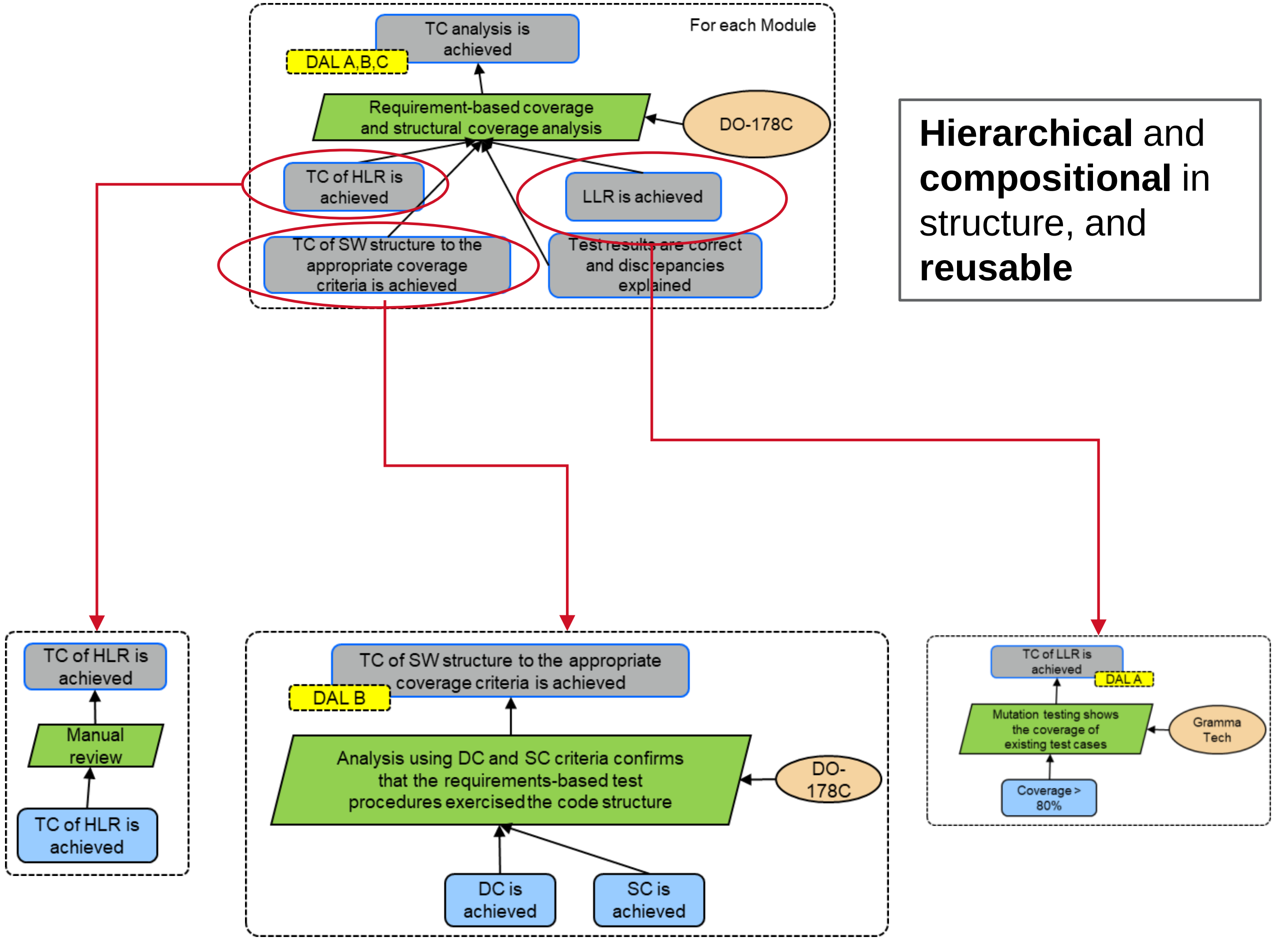
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

AACE: Automated Assurance Case Environment



Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Assurance Case Patterns


For each Module
TC analysis is achieved
DAL A,B,C
Requirement-based coverage and structural coverage analysis
DO-178C
TC of HLR is achieved
LLR is achieved
TC of SW structure to the appropriate coverage criteria is achieved
Test results are correct and discrepancies explained
TC of HLR is achieved
Manual review
TC of HLR is achieved
TC of SW structure to the appropriate coverage criteria is achieved
DAL B
Analysis using DC and SC criteria confirms that the requirements-based test procedures exercised the code structure
DO-178C
DC is achieved
SC is achieved
TC of LLR is achieved
DAL A
Mutation testing shows the coverage of existing test cases
Gramma Tech
Coverage > 80%

Hierarchical and **compositional** in structure, and **reusable**

Assurance case argumentation pattern:

- An argumentation pattern consists of a given **claim** (or conclusion), the associated **evidence** or sub-claims and the **argument** for why the claim could be concluded in a given context or given restrictions

- Context information can be categorized into User Domain, Rationale, and Restriction, and used for automatic selection of patterns

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Formalizing Assurance Case Patterns

Hierarchical Contract Nets (HCN)

Human-Understandable (FAN Notation and Justification)

Believing

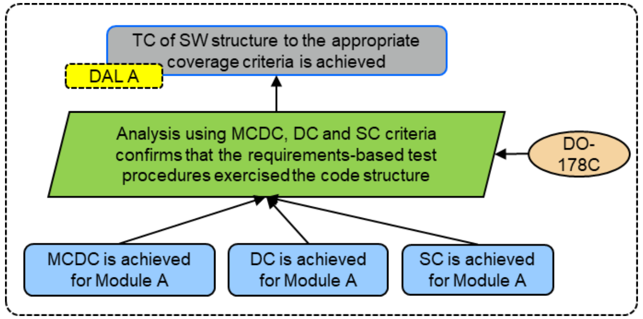
Structure TC of software Module A to the appropriate criteria is achieved.

Is justified by applying

DO-178C standard for DAL A software

To these premises

1. MCDC is achieved for Module A
2. DC is achieved for Module A
3. SC is achieved for Module A



TC of SW structure to the appropriate coverage criteria is achieved

DAL A

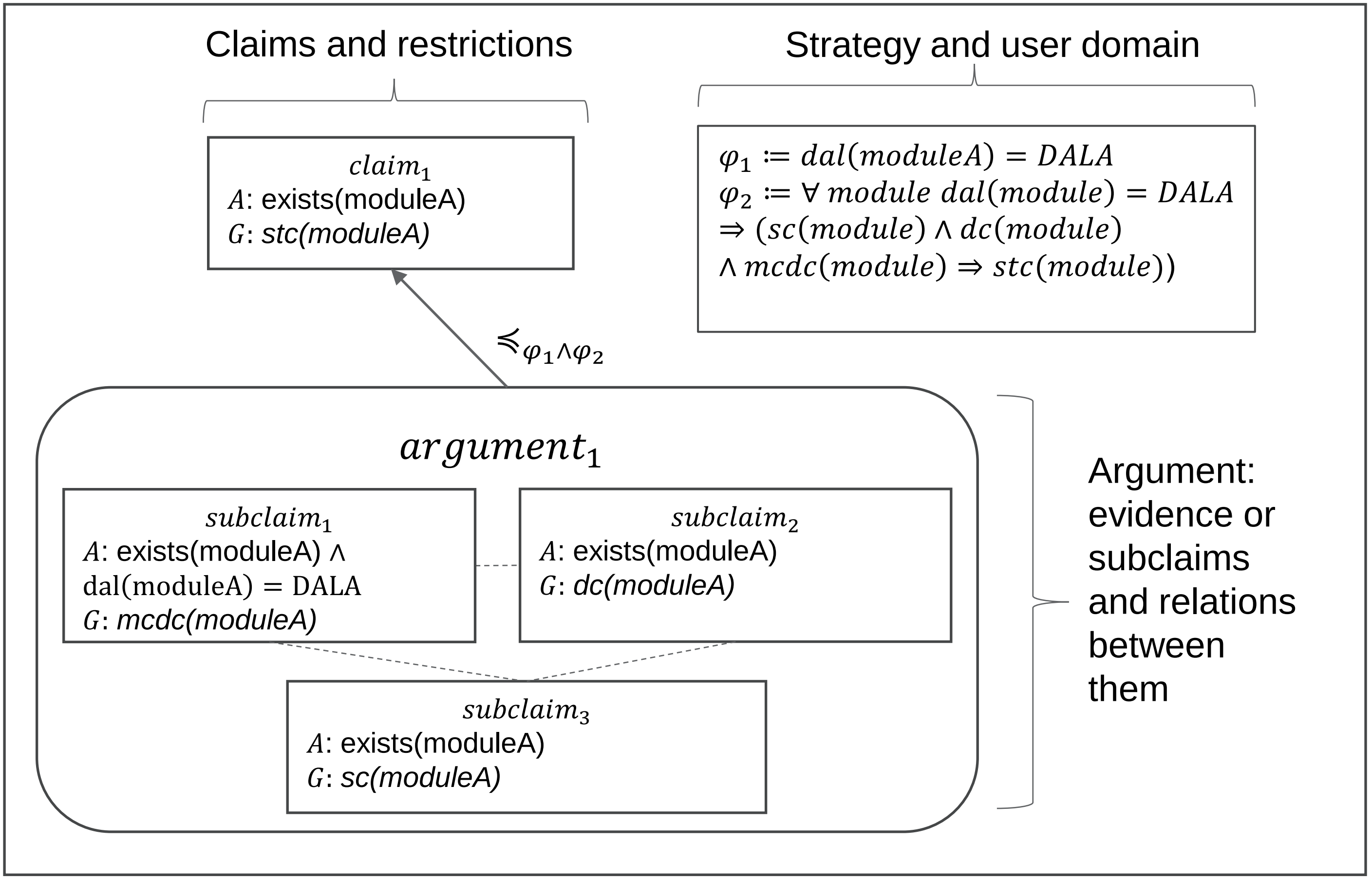
Analysis using MCDC, DC and SC criteria confirms that the requirements-based test procedures exercised the code structure

DO-178C

MCDC is achieved for Module A

DC is achieved for Module A

SC is achieved for Module A



Claims and restrictions

Strategy and user domain

$claim_1$
A: exists(moduleA)
G: *stc(moduleA)*

$\varphi_1 := dal(moduleA) = DALA$
$\varphi_2 := \forall\ module\ dal(module) = DALA \Rightarrow (sc(module) \land dc(module) \land mcdc(module) \Rightarrow stc(module))$

$\preccurlyeq_{\varphi_1 \land \varphi_2}$

$argument_1$

$subclaim_1$
A: exists(moduleA) $\land$ $\mathrm{dal(moduleA)} = \mathrm{DALA}$
G: *mcdc(moduleA)*

$subclaim_2$
A: exists(moduleA)
G: *dc(moduleA)*

$subclaim_3$
A: exists(moduleA)
G: *sc(moduleA)*

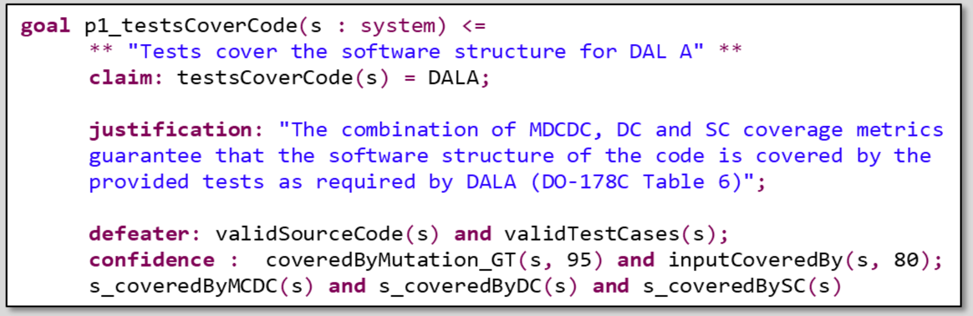
Argument: evidence or subclaims and relations between them

Timothy E. Wang, Zamira Daw, Pierluigi Nuzzo, and Alessandro Pinto. *Hierarchical Contract-Based Synthesis for Assurance Cases*. In NASA Formal Methods Symposium, 2022.

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Capturing Knowledge

Argumentation patterns and models for uncertainty quantification and decision making



```
goal p1_testsCoverCode(s : system) <=
     ** "Tests cover the software structure for DAL A" **
     claim: testsCoverCode(s) = DALA;

     justification: "The combination of MDCDC, DC and SC coverage metrics
     guarantee that the software structure of the code is covered by the
     provided tests as required by DALA (DO-178C Table 6)";

     defeater: validSourceCode(s) and validTestCases(s);
     confidence :  coveredByMutation_GT(s, 95) and inputCoveredBy(s, 80);
     s_coveredByMCDC(s) and s_coveredByDC(s) and s_coveredBySC(s)
```

Human-Understandable Argument

Believing

Tests cover source code

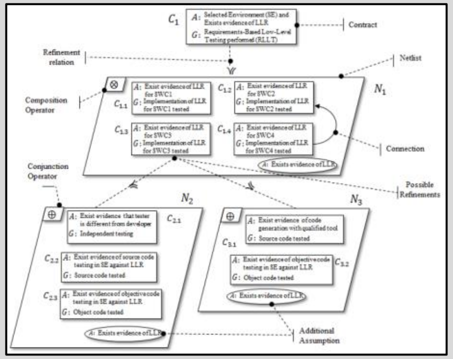
To these premises

1. MCDC is achieved
2. DC is achieved
3. Statement coverage is achieved

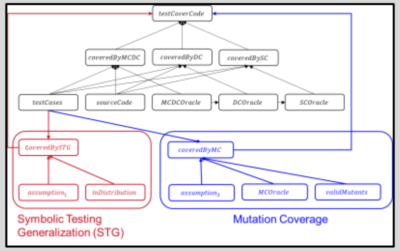
Confidence affected by

1. Input space is covered
2. Testing mutation coverage

Logic Argument (HCN)


C_1
A: Selected Environment (SE) and Exists evidence of LLR
G: Requirements-Based Low-Level Testing performed (RLLT)
Contract
Refinement relation
Netlist
Composition Operator
N_1
$C_{1.1}$ A: Exist evidence of LLR for SWC1 G: Implementation of LLR for SWC1 tested
$C_{1.2}$ A: Exist evidence of LLR for SWC2 G: Implementation of LLR for SWC2 tested
$C_{1.3}$ A: Exist evidence of LLR for SWC3 G: Implementation of LLR for SWC3 tested
$C_{1.4}$ A: Exist evidence of LLR for SWC4 G: Implementation of LLR for SWC4 tested
Connection
A: Exists evidence of LLR
Conjunction Operator
Possible Refinements
N_2
N_3
A: Exist evidence that tester is different from developer G: Independent testing $C_{2.1}$
$C_{2.2}$ A: Exist evidence of source code testing in SE against LLR G: Source code tested
$C_{2.3}$ A: Exist evidence of objective code testing in SE against LLR G: Object code tested
A: Exists evidence of LLR
A: Exist evidence of code generation with qualified tool G: Source code tested $C_{3.1}$
A: Exist evidence of objective code testing in SE against LLR G: Object code tested $C_{3.2}$
A: Exists evidence of LLR
Additional Assumption

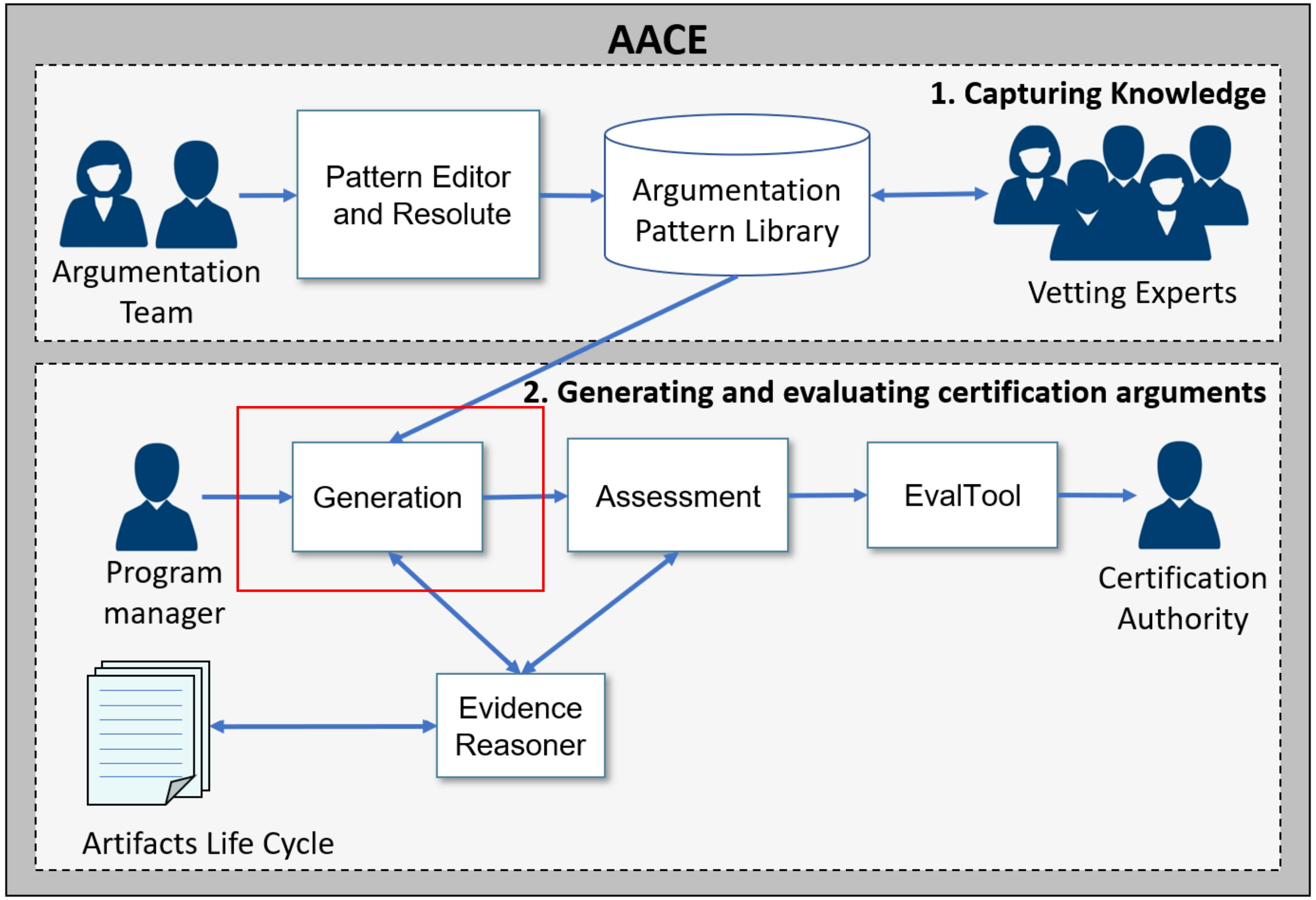
Confidence Argument (BN)



- Make knowledge **reusable**
- **Vetting process** involves consensus among different experts
  - Elicitation process
  - Validation process
  - Formal methods to verify validity of the patterns

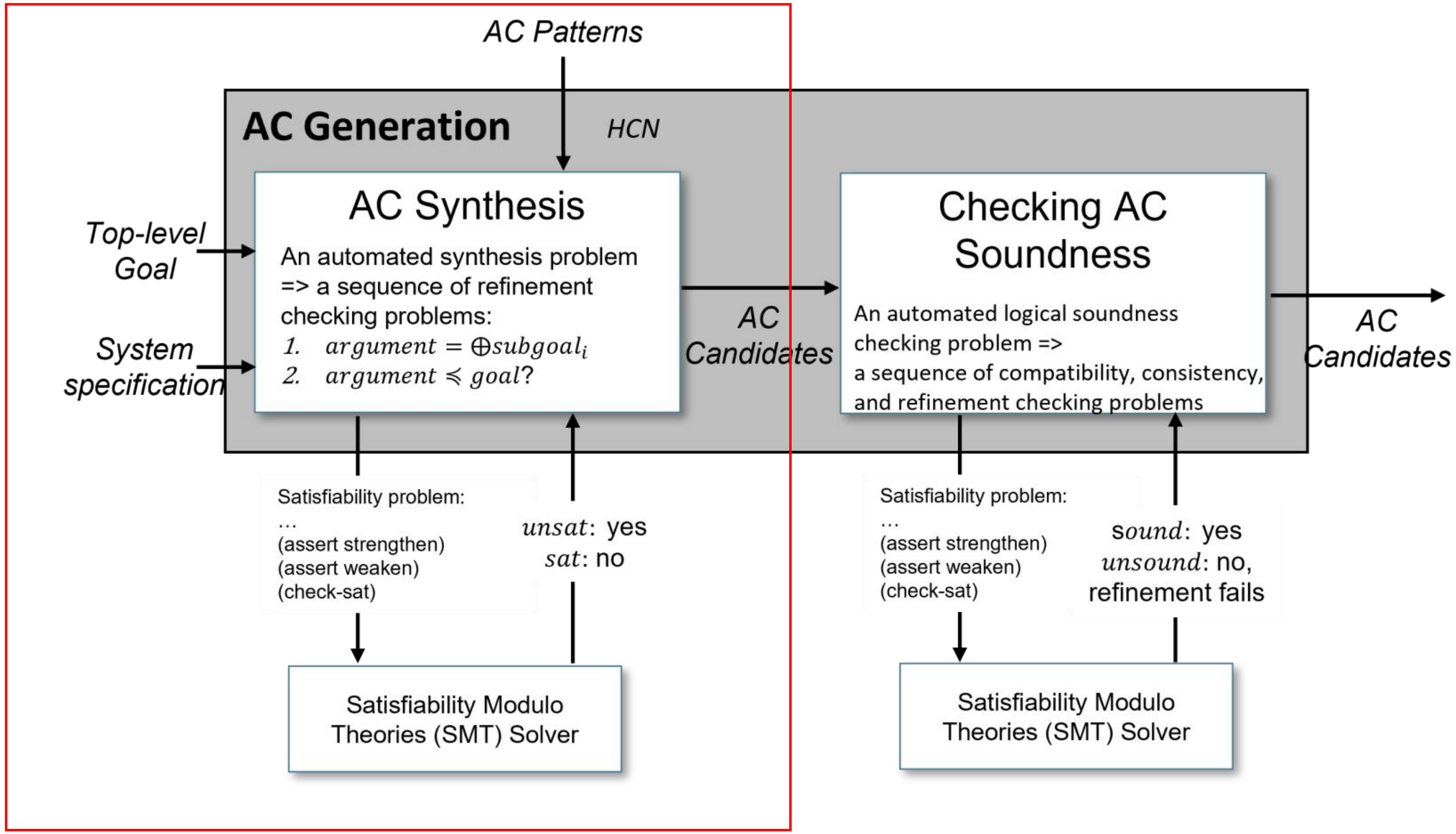
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

AACE: Automated Assurance Case Environment



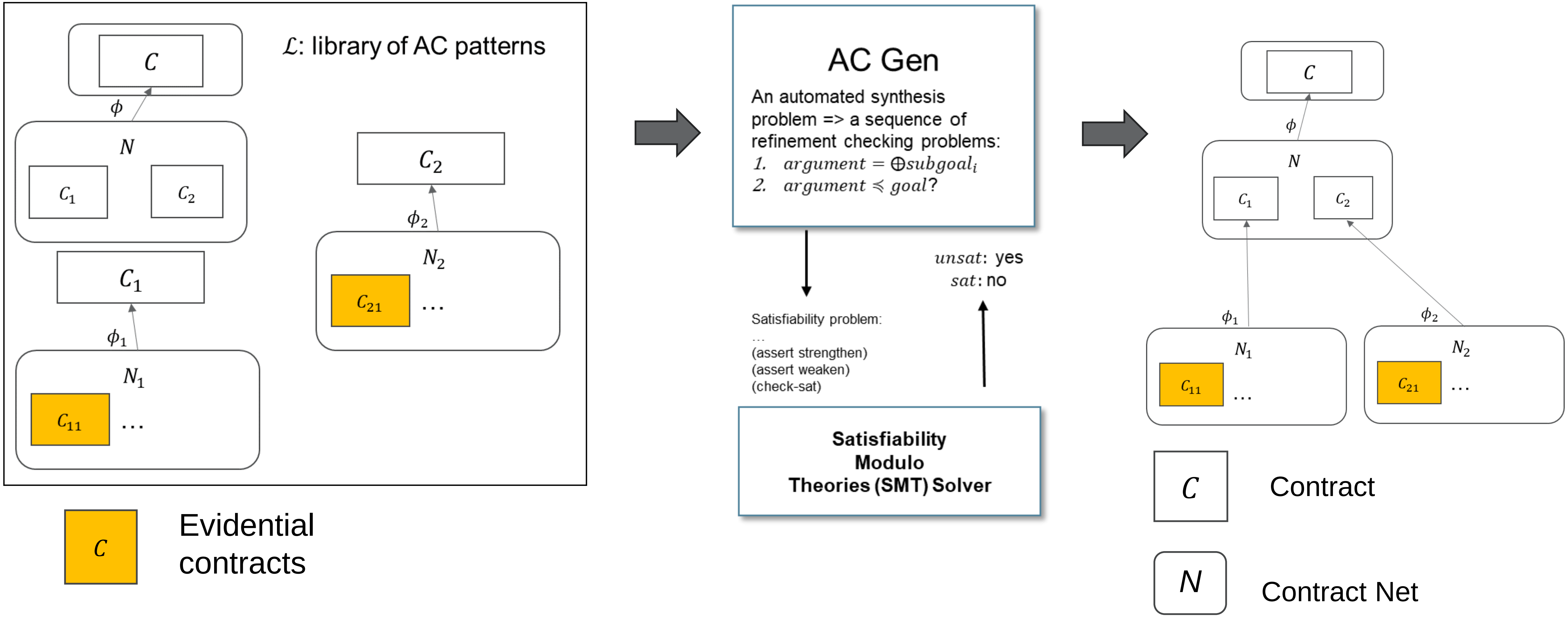
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Generating Certification Arguments


AC Patterns
AC Generation
HCN
Top-level Goal
System specification
AC Synthesis
An automated synthesis problem => a sequence of refinement checking problems:
1. $argument = \oplus subgoal_i$
2. $argument \preccurlyeq goal$?
AC Candidates
Checking AC Soundness
An automated logical soundness checking problem => a sequence of compatibility, consistency, and refinement checking problems
AC Candidates
Satisfiability problem:
…
(assert strengthen)
(assert weaken)
(check-sat)
$unsat$: yes
sat: no
Satisfiability Modulo Theories (SMT) Solver
Satisfiability problem:
…
(assert strengthen)
(assert weaken)
(check-sat)
$sound$: yes
$unsound$: no, refinement fails
Satisfiability Modulo Theories (SMT) Solver

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Automatic Synthesis



Evidential contracts

AC Gen

An automated synthesis problem => a sequence of refinement checking problems:

1. $argument = \bigoplus subgoal_i$
2. $argument \preccurlyeq goal$?

Satisfiability problem:
…
(assert strengthen)
(assert weaken)
(check-sat)

$unsat$: yes
sat: no

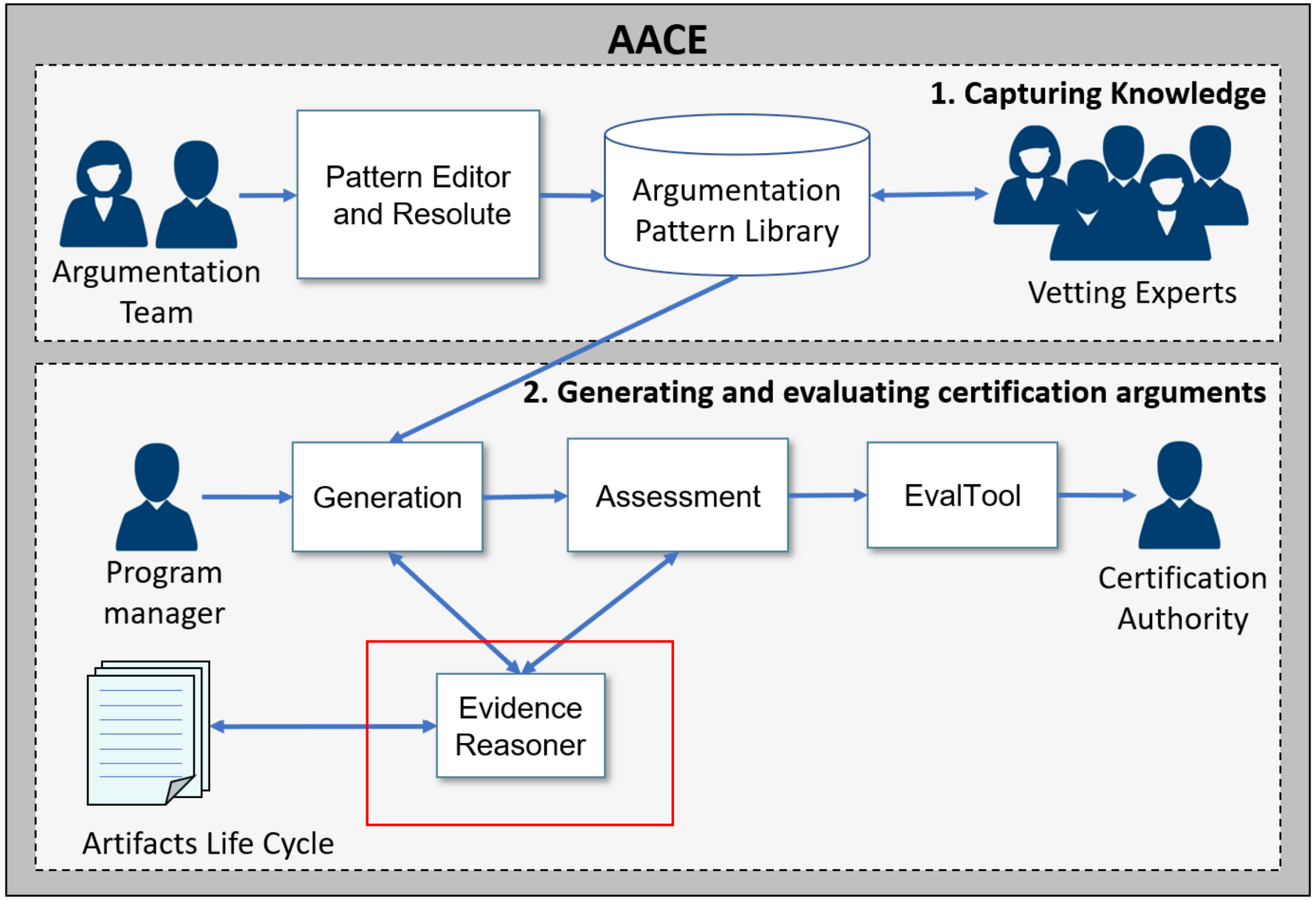
Satisfiability Modulo Theories (SMT) Solver

C Contract

N Contract Net

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

AACE: Automated Assurance Case Environment



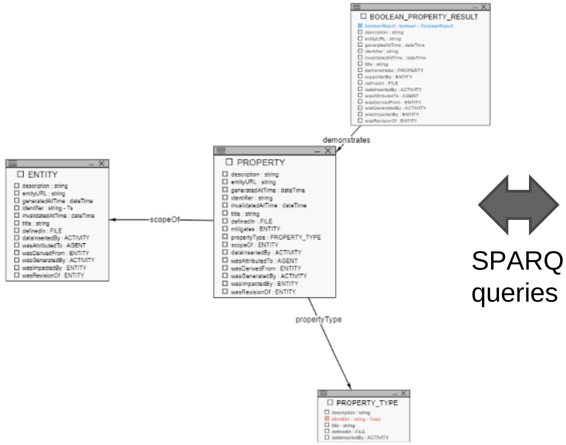
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Reasoning About Certification Artifacts

- Evidence produced during the entire lifecycle is captured in an ontology database
- The ontology defines key software certification concepts such as Component, Requirement, Tests,…
- An ontology reasoner between the assurance case generation and the evidence extraction allows a varied source of evidence and pre-analysis to identify inconsistencies and conflicts
- Currently, we are using as ontology-based database known as RACK

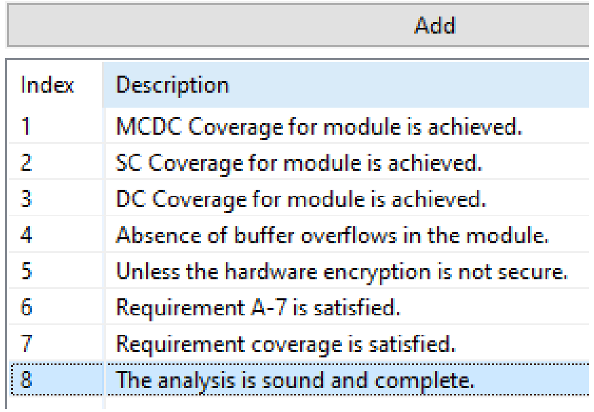
Heterogeneous evidence extraction

Ontological model of evidence (GE RACK)



SPARQL queries

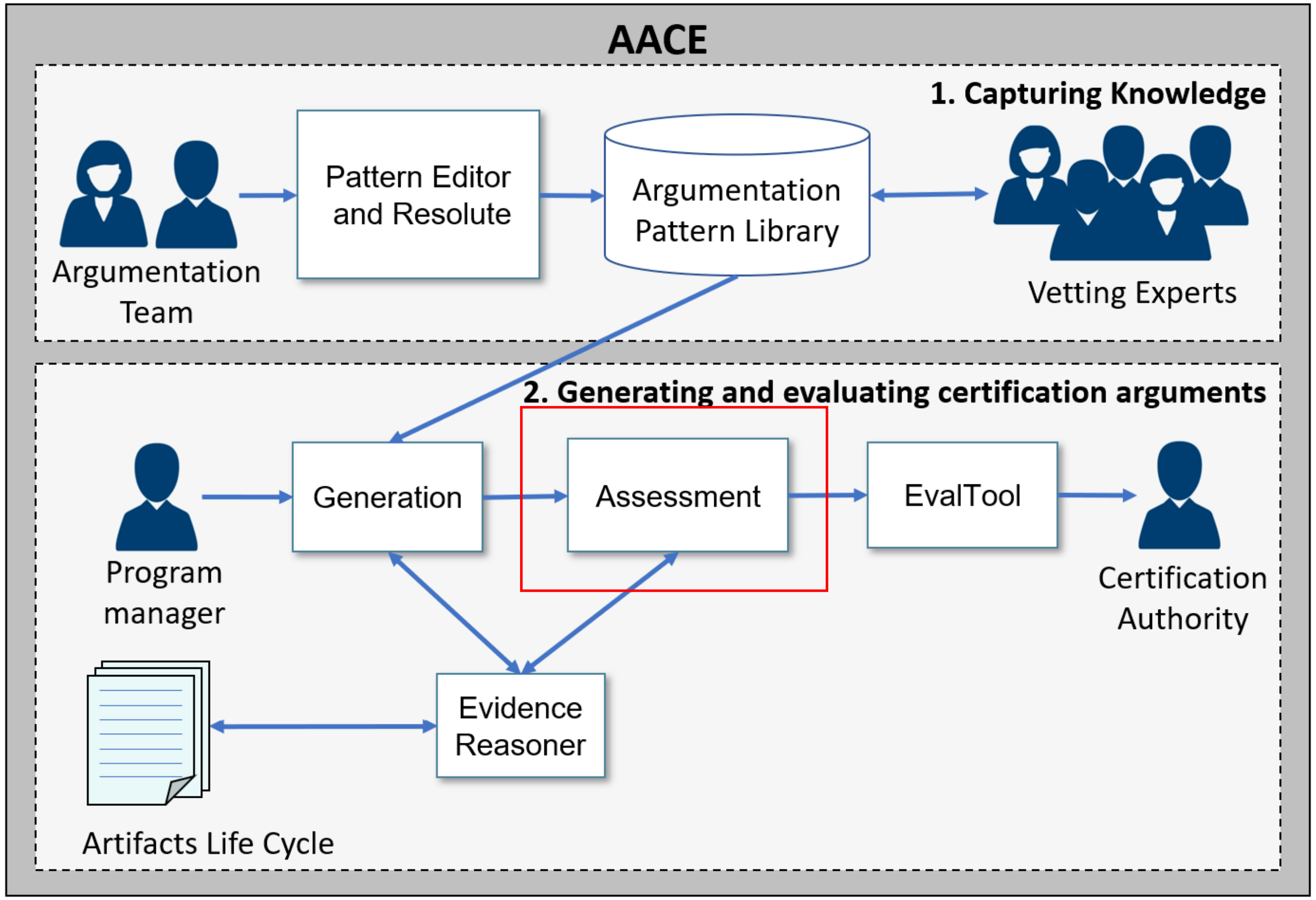
Evidence as premises, defeaters, and sources of doubt



Index	Description
1	MCDC Coverage for module is achieved.
2	SC Coverage for module is achieved.
3	DC Coverage for module is achieved.
4	Absence of buffer overflows in the module.
5	Unless the hardware encryption is not secure.
6	Requirement A-7 is satisfied.
7	Requirement coverage is satisfied.
8	The analysis is sound and complete.

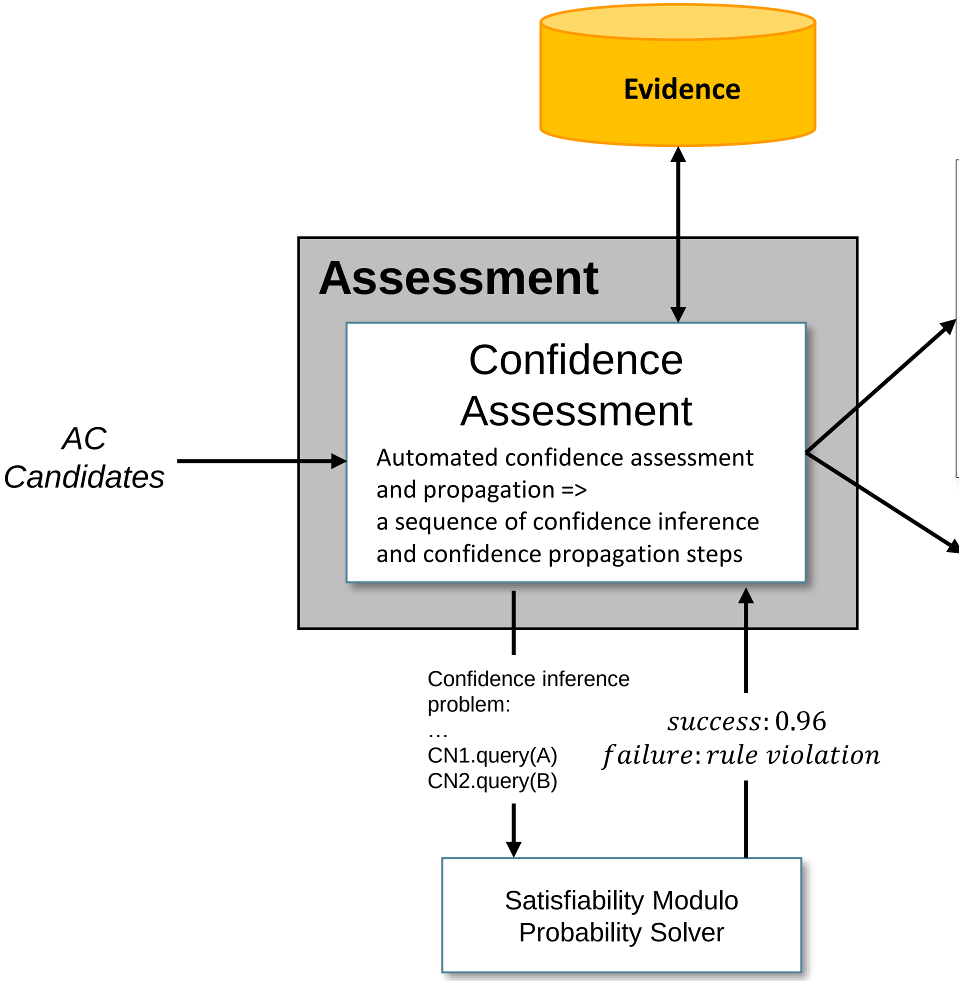
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

AACE: Automated Assurance Case Environment



Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Assessing Certification Arguments



Evidence

Assessment

Confidence Assessment

Automated confidence assessment and propagation =>
a sequence of confidence inference and confidence propagation steps

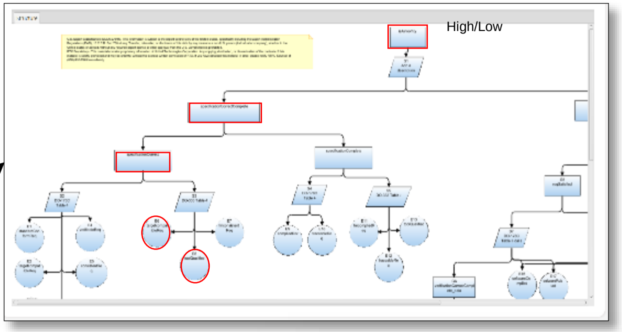
AC Candidates

Confidence inference problem:
…
CN1.query(A)
CN2.query(B)

$success: 0.96$
$failure: rule\ violation$

Satisfiability Modulo Probability Solver

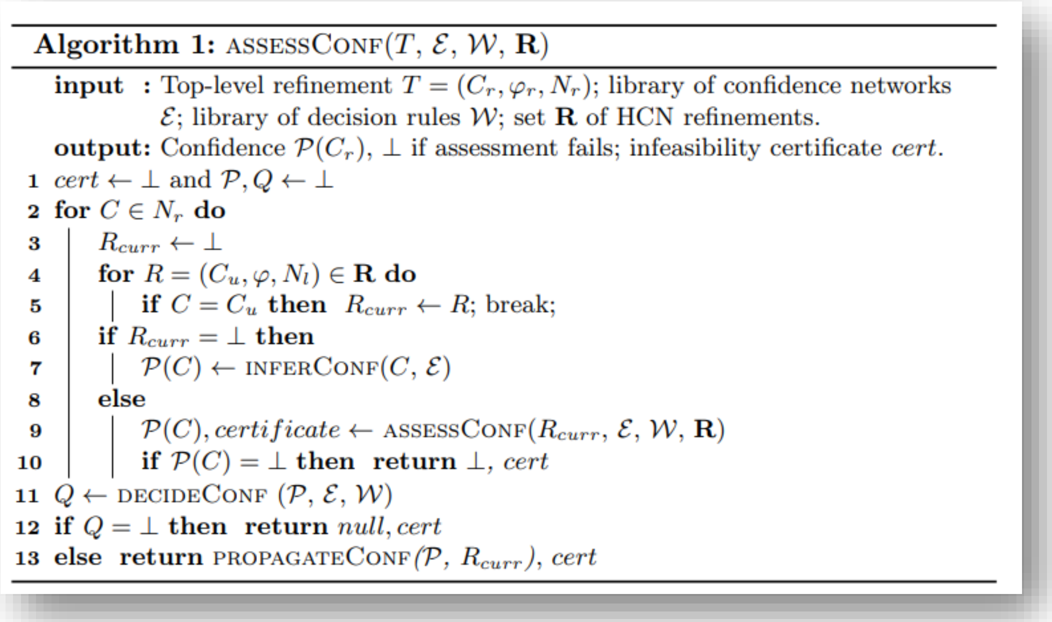
Valid AC(s)



Risk & Infeasibility Certificate

Completion Metric:

Select valid arguments with minimum number of ”missing” evidence items



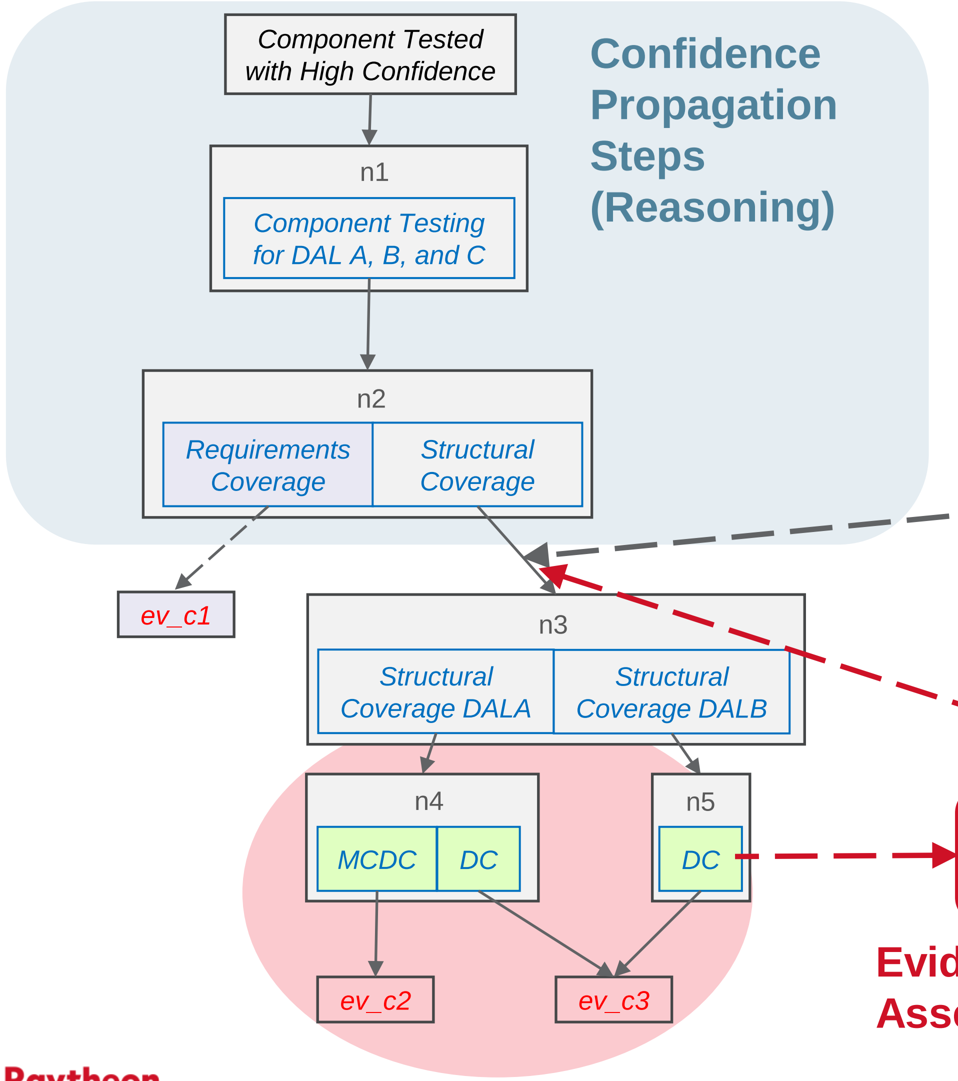
```
Algorithm 1: AssessConf($T$, $\mathcal{E}$, $\mathcal{W}$, $\mathbf{R}$)
input : Top-level refinement $T = (C_r, \varphi_r, N_r)$; library of confidence networks
        $\mathcal{E}$; library of decision rules $\mathcal{W}$; set $\mathbf{R}$ of HCN refinements.
output: Confidence $\mathcal{P}(C_r)$, $\bot$ if assessment fails; infeasibility certificate $cert$.
$cert \leftarrow \bot$ and $\mathcal{P}, Q \leftarrow \bot$
for $C \in N_r$ do
    $R_{curr} \leftarrow \bot$
    for $R = (C_u, \varphi, N_l) \in \mathbf{R}$ do
        if $C = C_u$ then $R_{curr} \leftarrow R$; break;
    if $R_{curr} = \bot$ then
        $\mathcal{P}(C) \leftarrow$ InferConf($C$, $\mathcal{E}$)
    else
        $\mathcal{P}(C), certificate \leftarrow$ AssessConf($R_{curr}$, $\mathcal{E}$, $\mathcal{W}$, $\mathbf{R}$)
        if $\mathcal{P}(C) = \bot$ then return $\bot$, $cert$
$Q \leftarrow$ DecideConf ($\mathcal{P}$, $\mathcal{E}$, $\mathcal{W}$)
if $Q = \bot$ then return $null, cert$
else return PropagateConf($\mathcal{P}$, $R_{curr}$), $cert$
```

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

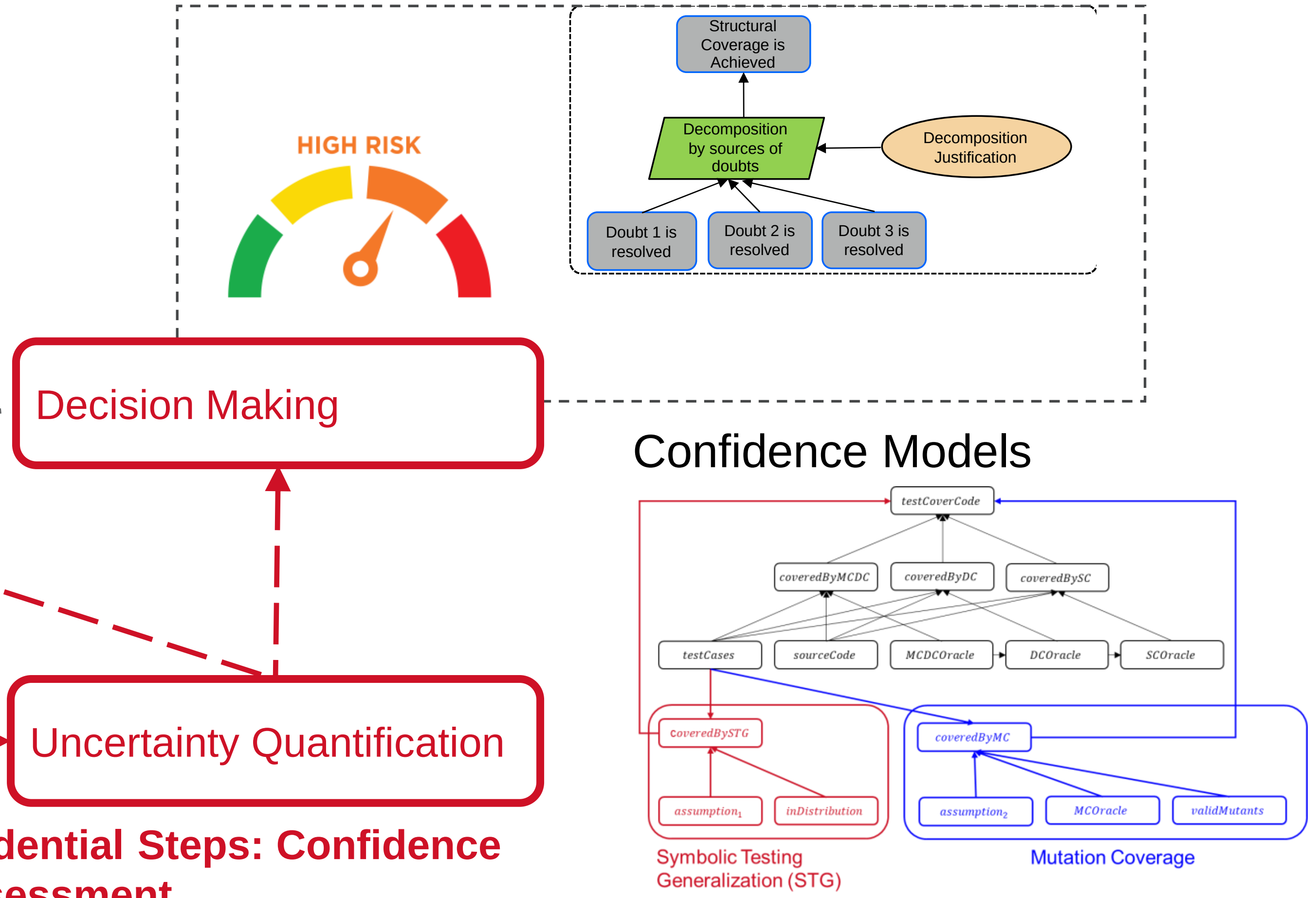
Assessing Certification Arguments

Hierarchical Stochastic Contract Nets (HSCN)

Chanwook Oh, Nikhil Vijay Naik, Zamira Daw, Timothy E. Wang and Pierluigi Nuzzo. *ARACHNE: Automated Validation of Assurance Cases with Stochastic Contract Networks*, SafeComp 2022.



Confidence Propagation Steps (Reasoning)



Evidential Steps: Confidence Assessment

Raytheon Technologies

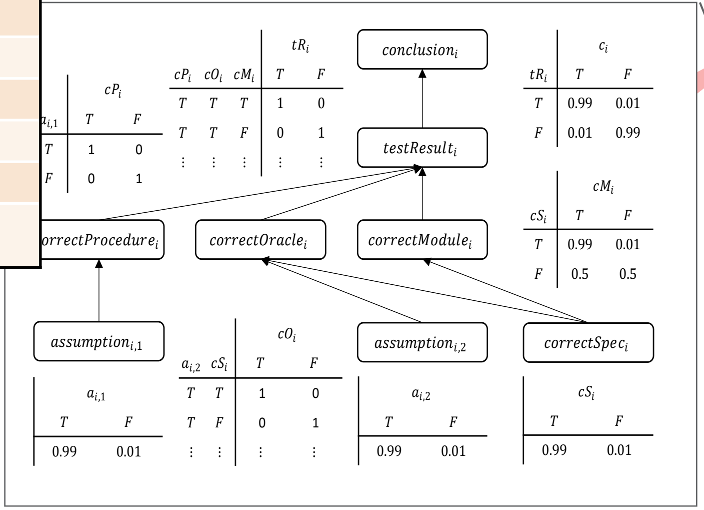
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Assessing Certification Arguments: Example

LLR	HLR	Analysis	Output
LLR-1	HLR-1	Manual Review	True
LLR-3	HLR-2	Model Checking Verification	False
LLR-4	HLR-2	Model Checking Verification	False
LLR-5	HLR-3	Simulink verification	True
LLR-6	HLR-3	Simulink verification	True

SW-Comp	Analysis	Output
Comp-1	MCDC	True
Comp-1	DC	True
Comp-1	SC	True
Comp-1	Mutation > 80%	True
Comp-1	Input Space Covered	True
Comp-1	Requirements address robustness	False

Confidence Propagation



$a_{i,1}$	cP_i T	cP_i F
T	1	0
F	0	1

cP_i	cO_i	cM_i	tR_i T	tR_i F
T	T	T	1	0
T	T	F	0	1
⋮	⋮	⋮	⋮	⋮

tR_i	c_i T	c_i F
T	0.99	0.01
F	0.01	0.99

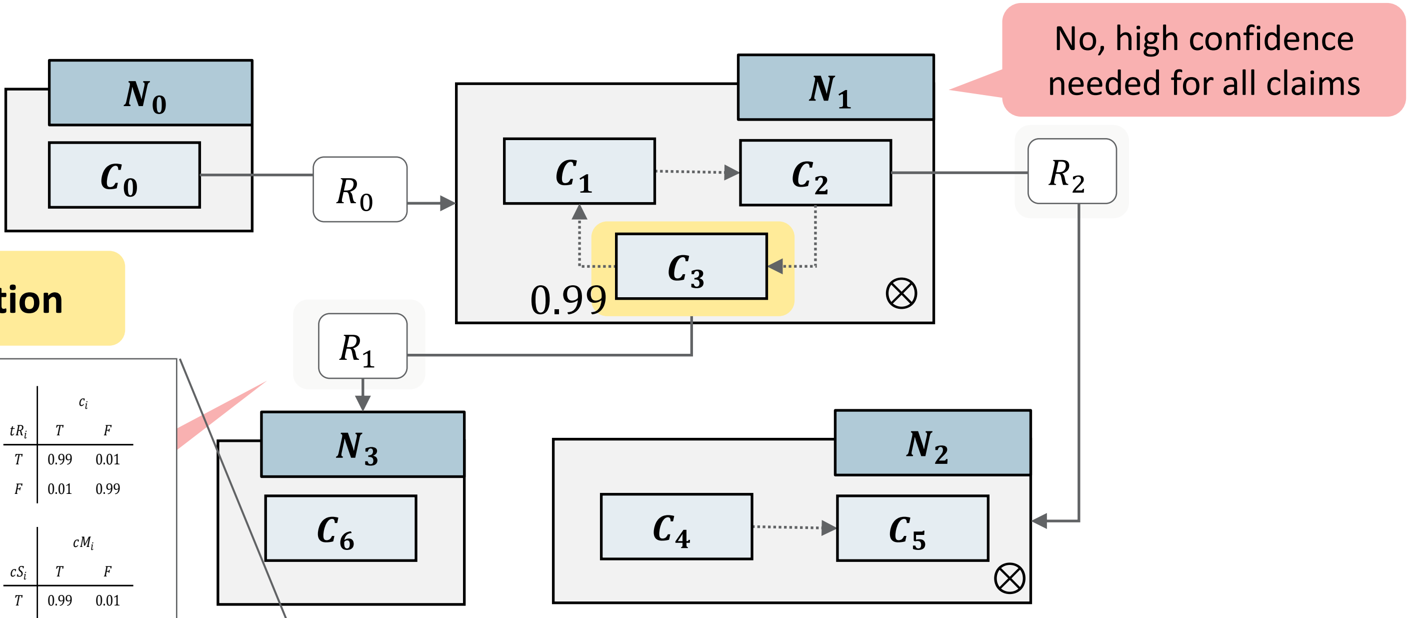
cS_i	cM_i T	cM_i F
T	0.99	0.01
F	0.5	0.5

$a_{i,1}$ T	$a_{i,1}$ F
0.99	0.01

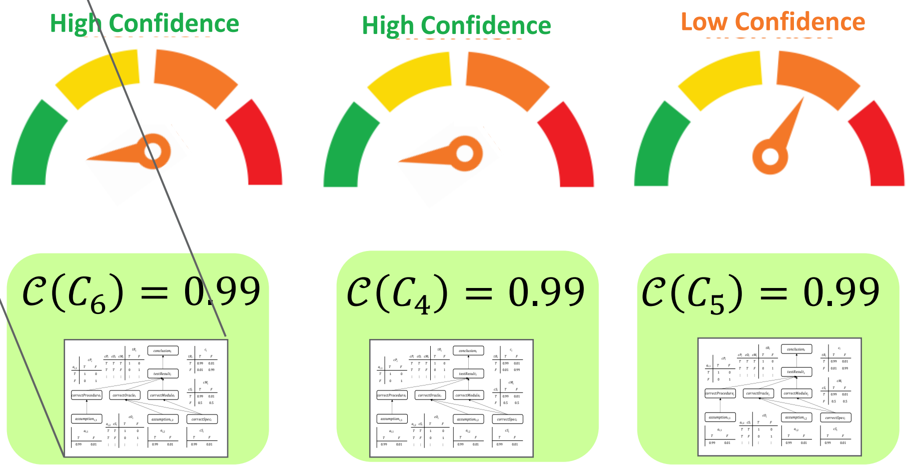
$a_{i,2}$	cS_i	cO_i T	cO_i F
T	T	1	0
T	F	0	1
⋮	⋮	⋮	⋮

$a_{i,2}$ T	$a_{i,2}$ F
0.99	0.01

cS_i T	cS_i F
0.99	0.01



No, high confidence needed for all claims



$\mathcal{C}(C_6) = 0.99$

$\mathcal{C}(C_4) = 0.99$

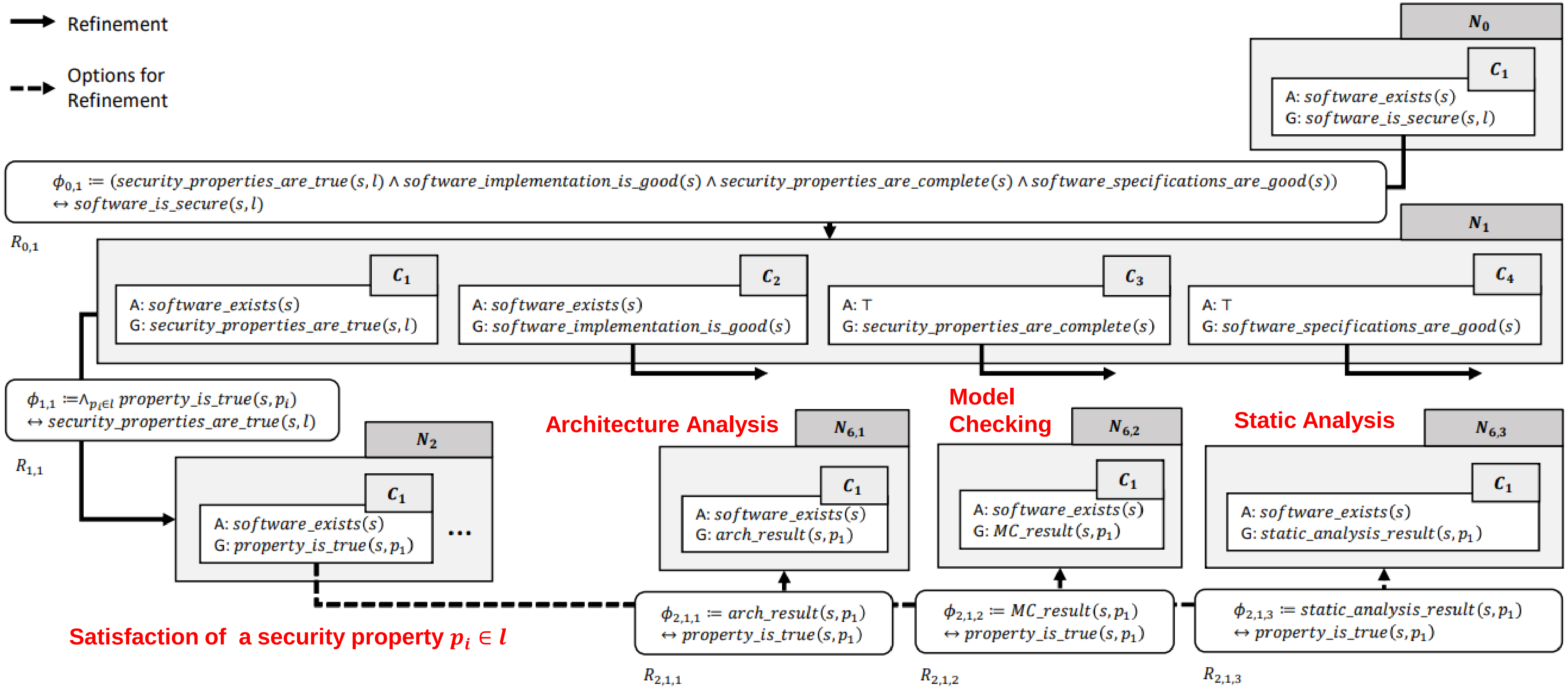
$\mathcal{C}(C_5) = 0.99$

Confidence Computation
Using **Bayesian Inference**
on Bayesian Network

Raytheon Technologies

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

ArduCopter Security Assurance



Refinement

Options for Refinement

N_0

C_1

A: $\mathit{software\_exists}(s)$
G: $\mathit{software\_is\_secure}(s,l)$

$\phi_{0,1} \coloneqq (\mathit{security\_properties\_are\_true}(s,l) \wedge \mathit{software\_implementation\_is\_good}(s) \wedge \mathit{security\_properties\_are\_complete}(s) \wedge \mathit{software\_specifications\_are\_good}(s))$
$\leftrightarrow \mathit{software\_is\_secure}(s,l)$

$R_{0,1}$

N_1

C_1

A: $\mathit{software\_exists}(s)$
G: $\mathit{security\_properties\_are\_true}(s,l)$

C_2

A: $\mathit{software\_exists}(s)$
G: $\mathit{software\_implementation\_is\_good}(s)$

C_3

A: $\top$
G: $\mathit{security\_properties\_are\_complete}(s)$

C_4

A: $\top$
G: $\mathit{software\_specifications\_are\_good}(s)$

$\phi_{1,1} \coloneqq \wedge_{p_i \in l}\, \mathit{property\_is\_true}(s,p_i)$
$\leftrightarrow \mathit{security\_properties\_are\_true}(s,l)$

$R_{1,1}$

N_2

C_1

A: $\mathit{software\_exists}(s)$
G: $\mathit{property\_is\_true}(s,p_1)$

…

Architecture Analysis

$N_{6,1}$

C_1

A: $\mathit{software\_exists}(s)$
G: $\mathit{arch\_result}(s,p_1)$

Model Checking

$N_{6,2}$

C_1

A: $\mathit{software\_exists}(s)$
G: $\mathit{MC\_result}(s,p_1)$

Static Analysis

$N_{6,3}$

C_1

A: $\mathit{software\_exists}(s)$
G: $\mathit{static\_analysis\_result}(s,p_1)$

Satisfaction of a security property $p_i \in l$

$\phi_{2,1,1} \coloneqq \mathit{arch\_result}(s,p_1)$
$\leftrightarrow \mathit{property\_is\_true}(s,p_1)$

$R_{2,1,1}$

$\phi_{2,1,2} \coloneqq \mathit{MC\_result}(s,p_1)$
$\leftrightarrow \mathit{property\_is\_true}(s,p_1)$

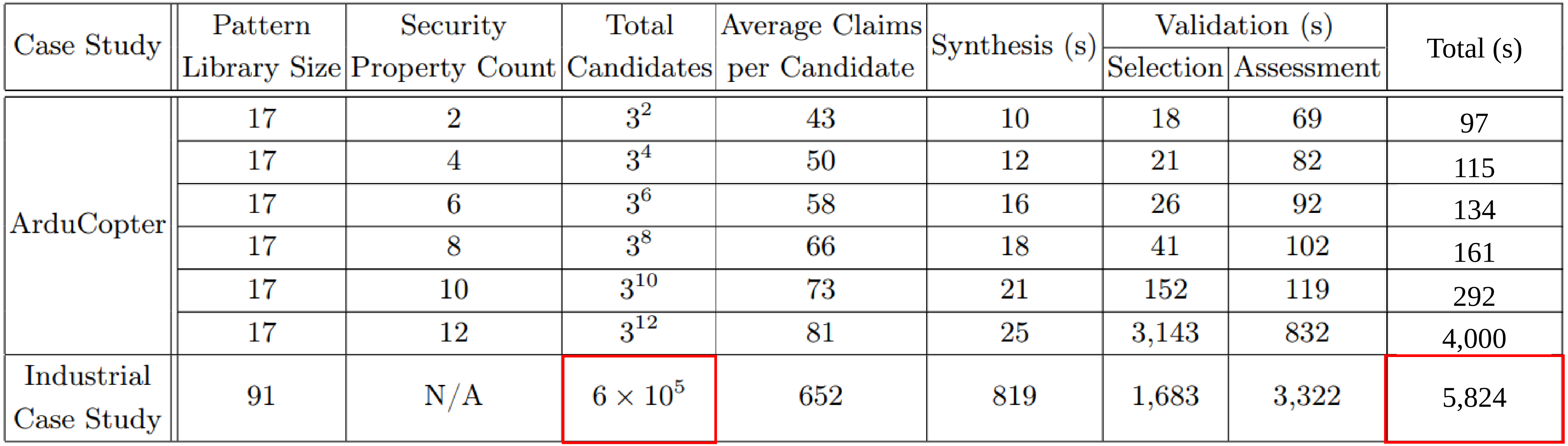
$R_{2,1,2}$

$\phi_{2,1,3} \coloneqq \mathit{static\_analysis\_result}(s,p_1)$
$\leftrightarrow \mathit{property\_is\_true}(s,p_1)$

$R_{2,1,3}$

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Synthesized and Evaluated over 10^5 Arguments in <100 min

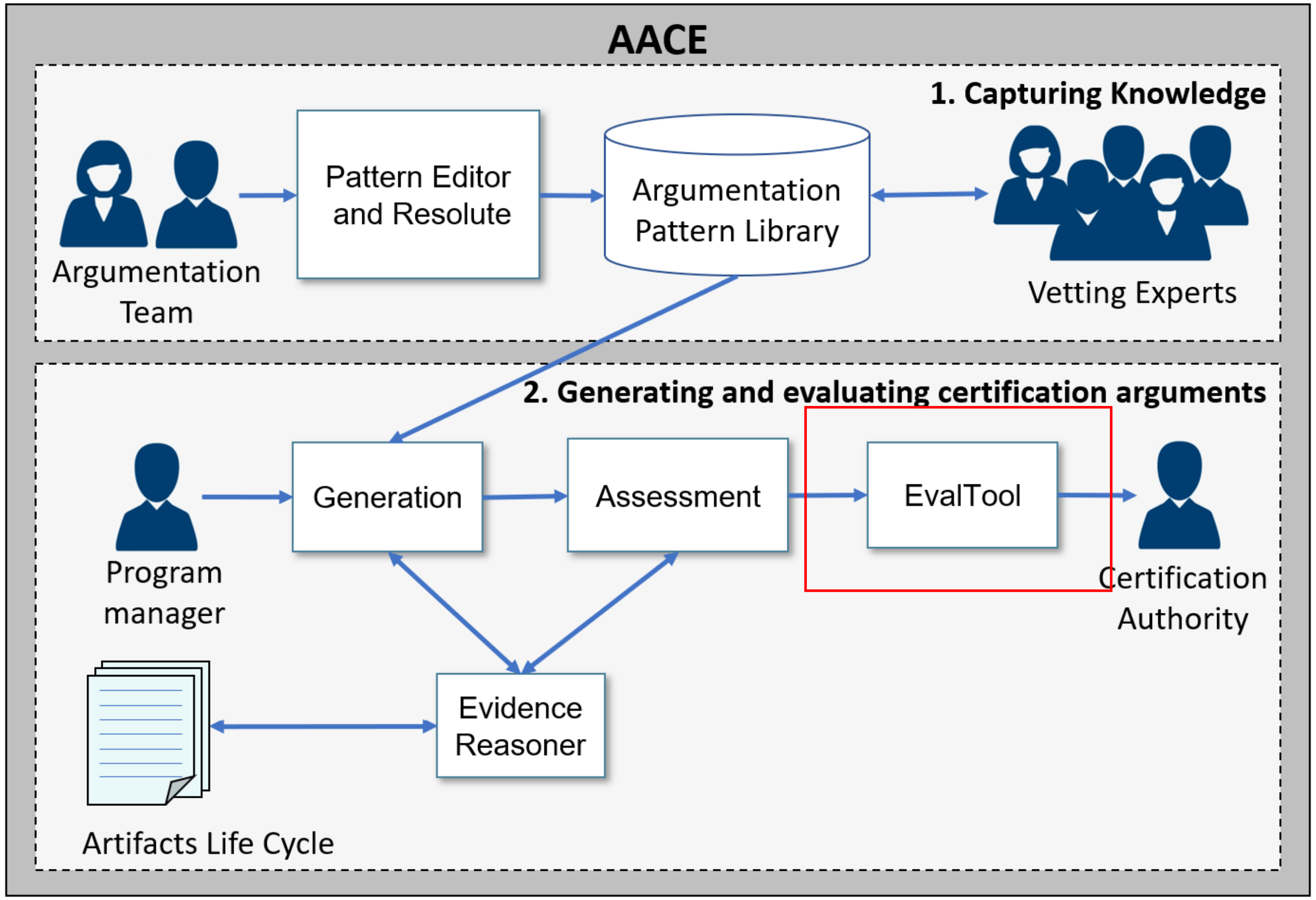


Case Study	Pattern Library Size	Security Property Count	Total Candidates	Average Claims per Candidate	Synthesis (s)	Validation (s)		Total (s)
						Selection	Assessment	
ArduCopter	17	2	3^2	43	10	18	69	97
	17	4	3^4	50	12	21	82	115
	17	6	3^6	58	16	26	92	134
	17	8	3^8	66	18	41	102	161
	17	10	3^{10}	73	21	152	119	292
	17	12	3^{12}	81	25	3,143	832	4,000
Industrial Case Study	91	N/A	6×10^5	652	819	1,683	3,322	5,824

Raytheon Technologies

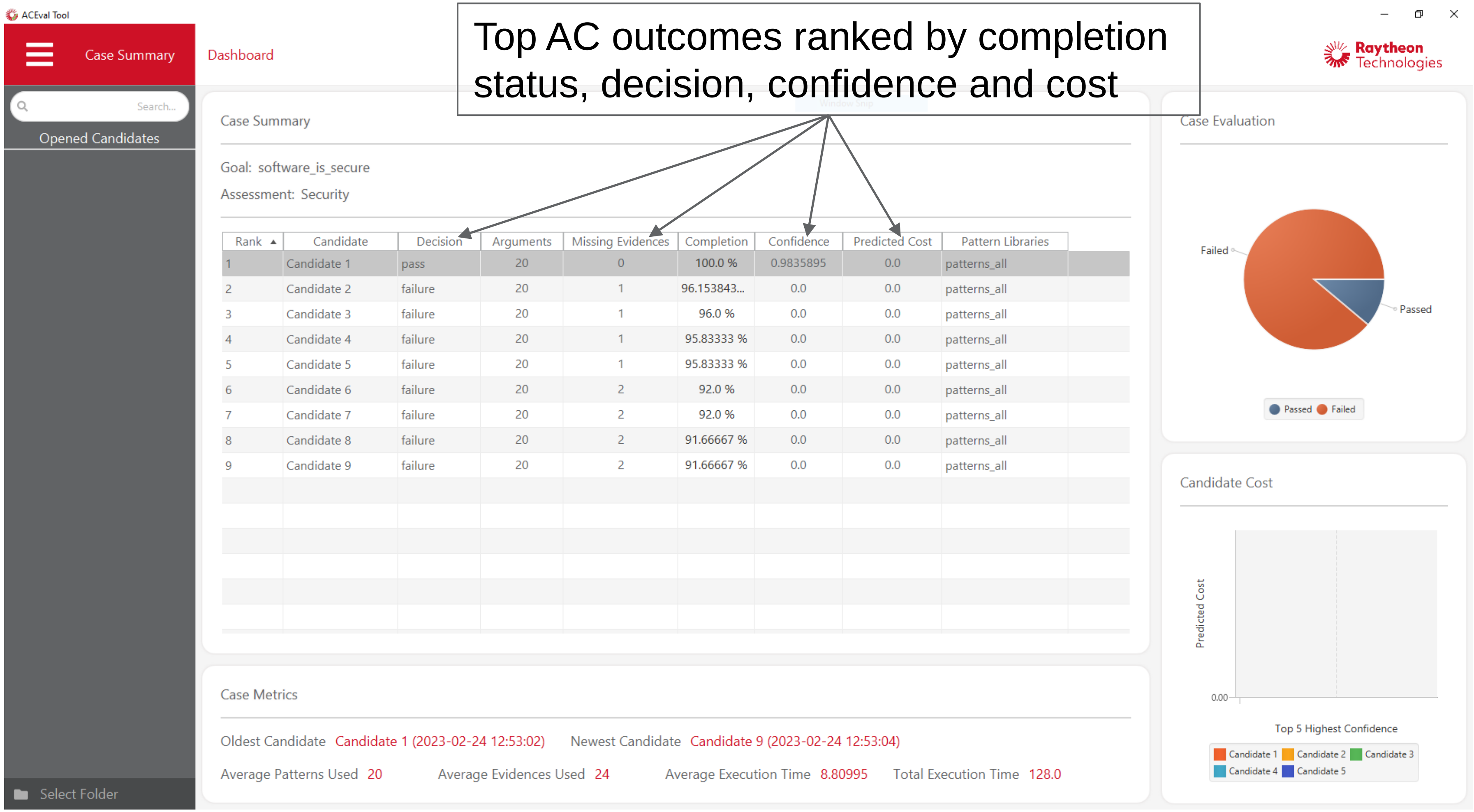
Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

AACE: Automated Assurance Case Environment



Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Eval Tool



Top AC outcomes ranked by completion status, decision, confidence and cost

Case Summary

Goal: software_is_secure

Assessment: Security

Rank ▲	Candidate	Decision	Arguments	Missing Evidences	Completion	Confidence	Predicted Cost	Pattern Libraries
1	Candidate 1	pass	20	0	100.0 %	0.9835895	0.0	patterns_all
2	Candidate 2	failure	20	1	96.153843...	0.0	0.0	patterns_all
3	Candidate 3	failure	20	1	96.0 %	0.0	0.0	patterns_all
4	Candidate 4	failure	20	1	95.83333 %	0.0	0.0	patterns_all
5	Candidate 5	failure	20	1	95.83333 %	0.0	0.0	patterns_all
6	Candidate 6	failure	20	2	92.0 %	0.0	0.0	patterns_all
7	Candidate 7	failure	20	2	92.0 %	0.0	0.0	patterns_all
8	Candidate 8	failure	20	2	91.66667 %	0.0	0.0	patterns_all
9	Candidate 9	failure	20	2	91.66667 %	0.0	0.0	patterns_all

Case Metrics

Oldest Candidate Candidate 1 (2023-02-24 12:53:02) Newest Candidate Candidate 9 (2023-02-24 12:53:04)

Average Patterns Used 20 Average Evidences Used 24 Average Execution Time 8.80995 Total Execution Time 128.0

Cost may consist of actual dollar cost, time, or any other limited or scarce resource

Z. Daw, T. Wang, C. Oh, M. Low, I. Amundson, G. Wang, R. Melville, and P. Nuzzo, “Computer-aided evaluation for argument-based certification,” in 42nd AIAA/IEEE Digital Avionics Systems Conference, 2023

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Thank You.

Raytheon Technologies

Distribution Statement A (Approved for Public Release, Distribution Unlimited). This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA) contract FA875020C0508. The views, opinions, or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.