UNIVERSITÀ DEGLI STUDI
DI GENOVA

Formal Assessment of Complex Engineered Systems

Armando Tacchella

Dipartimento di Informatica, Bioingegneria,
Robotica e Ingegneria dei Sistemi (DIBRIS)

CPS Summer School 2019, September 23-27

May 24th , 2018

UNIVERSITÀ DEGLI STUDI
DI GENOVA

A (Nearly Epic) Tale of Courage, Passion, Amazing Victories and Major Defeats

Armando Tacchella

Dipartimento di Informatica, Bioingegneria,
Robotica e Ingegneria dei Sistemi (DIBRIS)

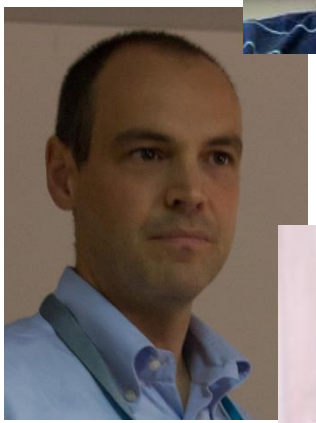
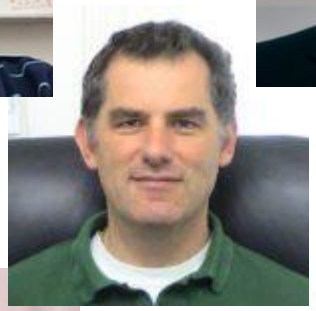
CPS Summer School 2019, September 23-27

May 24th , 2018

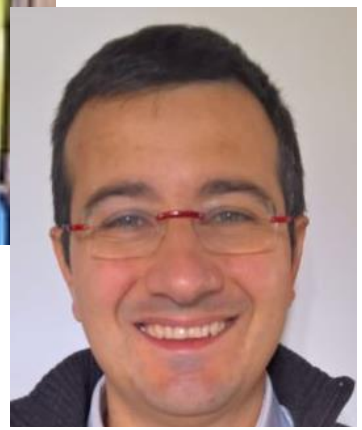
Starring...





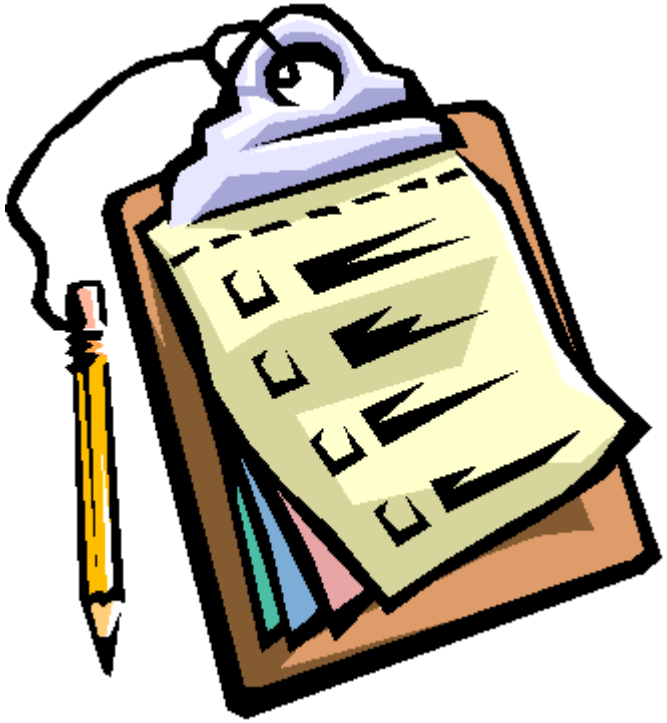






… and others who
did not fit the slide!

Agenda



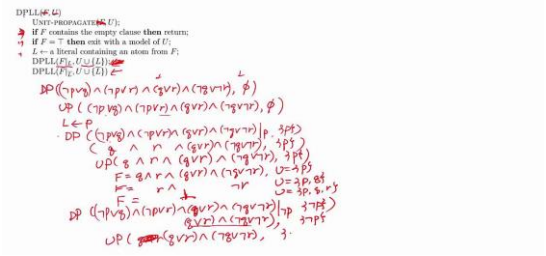
- **Beauty and the Beast**
  Formal methods vs. engineered systems
- **The Fantastic Four**
  Diverse application domains where we tried to apply formal methods
- **Edge of Tomorrow**
  Go formal, fail, repeat

Beauty...

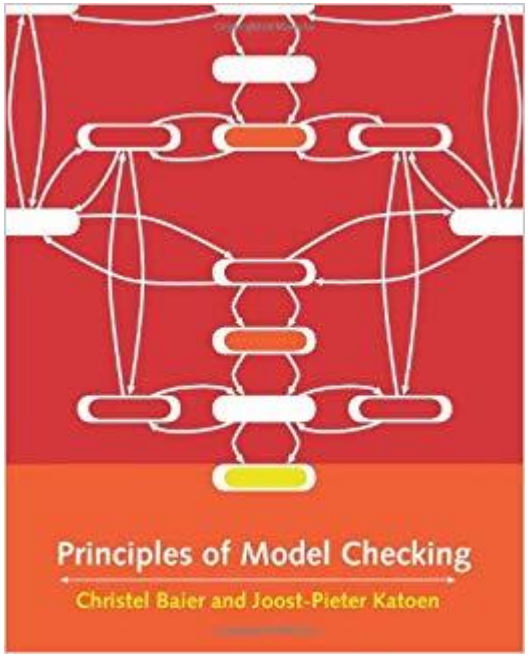
Rules of Inference

Modus Ponens	Modus Tollens	Hypothetical Syllogism
p $p \rightarrow q$ q	$\neg q$ $p \rightarrow q$ $\neg p$	$p \rightarrow q$ $q \rightarrow r$ $p \rightarrow r$
Addition	**Resolution**	**Disjunctive Syllogism**
p $p \vee q$	$p \vee q$ $\neg p \vee r$ $q \vee r$	$p \vee q$ $\neg p$ q
Simplification	**Conjunction**	
$p \wedge q$ p	p q $p \wedge q$	

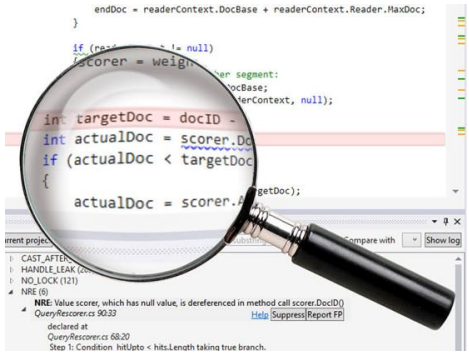
Theorem Proving



Decision Procedures

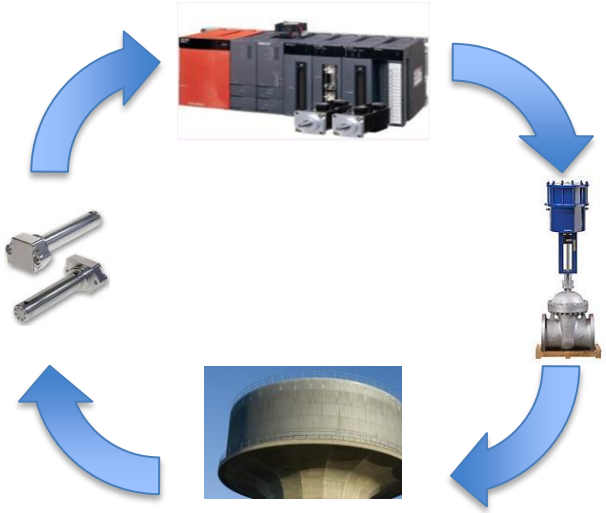


Model Checking



Static Analysis

... and the Beast!



Cyber-physical

Integrated

Connected

Socio-technical

EMERGENT BEHAVIORS

Complex

A Happy Ending?

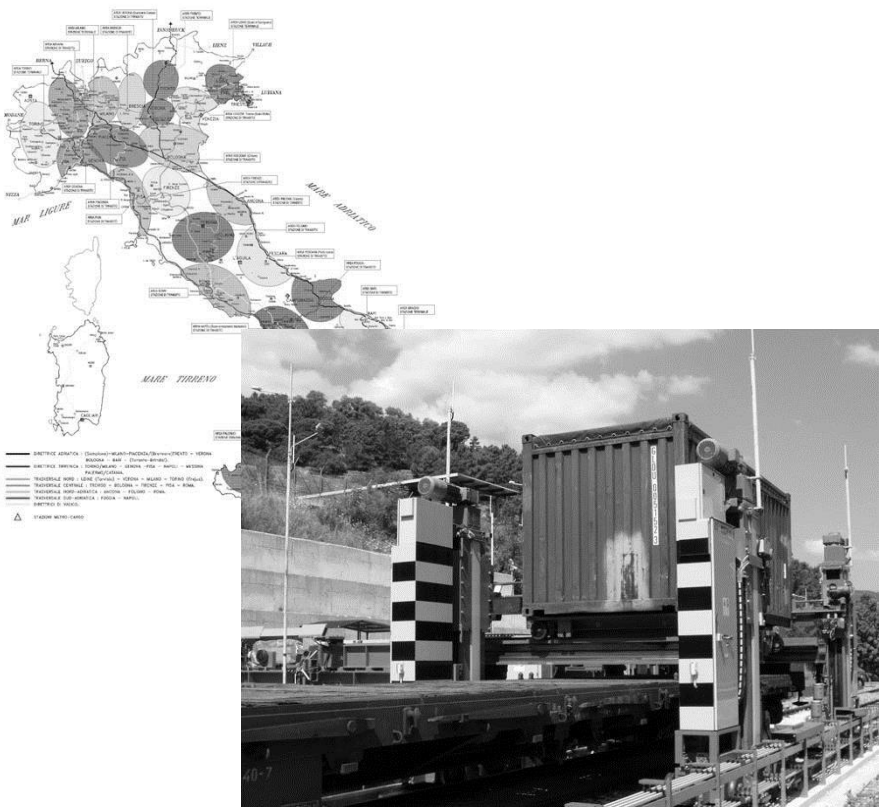
Do Formal Methods tame the beast...

... or maybe not?

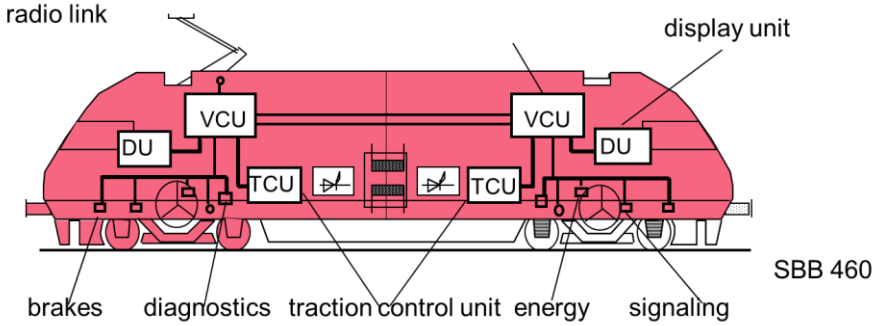
The Fantastic Four

1. System Monitoring and Maintenance
2. Predictability of Autonomous Systems
   - Analysis of Neural Nets (and similar stuff)
   - Safe Reinforcement Learning
   - Optimal Planning
3. Security of Critical Infrastructure
4. Computer-Automated Design

System monitoring and maintenance



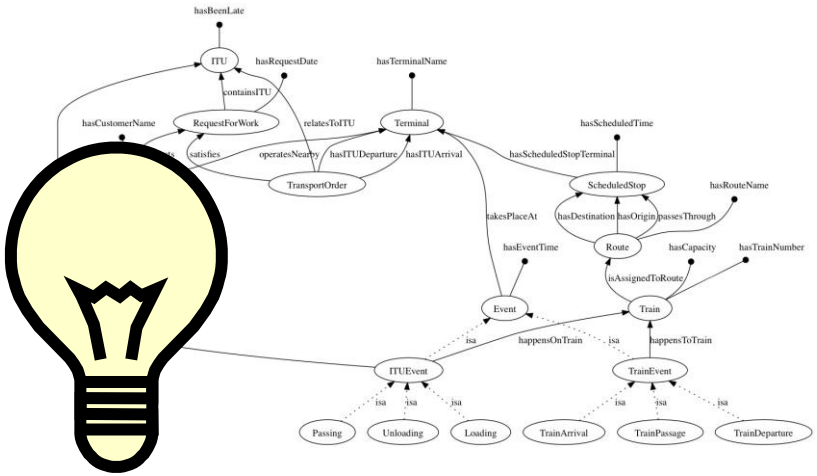
Networks of multimodal terminals



Locomotives

HVAC units

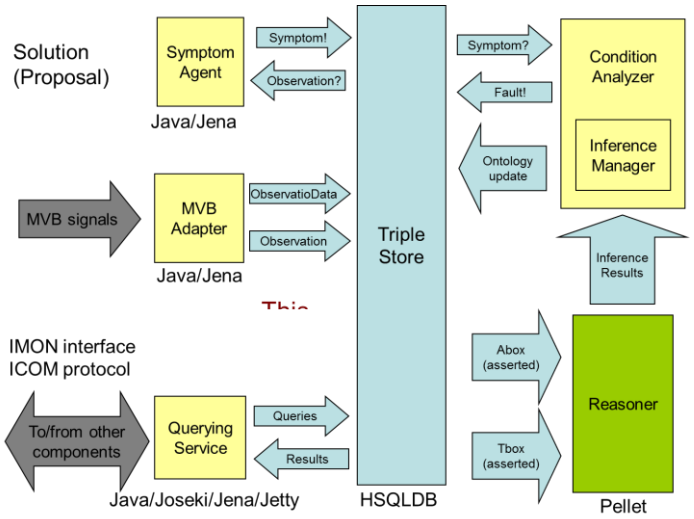
System monitoring and maintenance



Ontologies to represent components and relationships among them

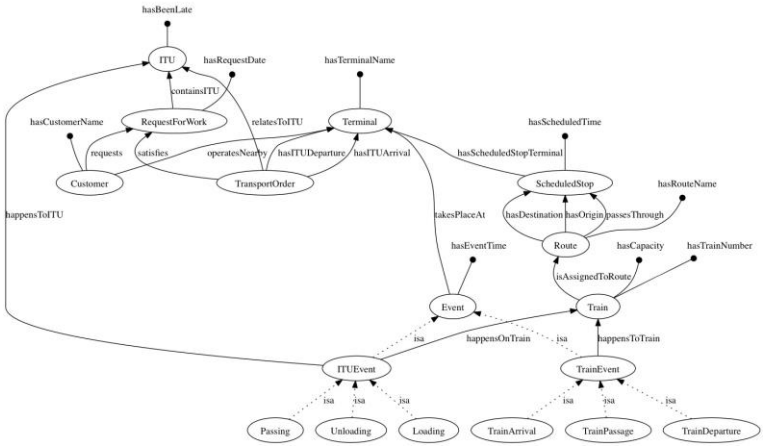
Not suitable to «reason about the system»

Cristina De Ambrosi, Cristiano Ghersi, Armando Tacchella:
An Ontology-Based Condition Analyzer for Fault Classification on Railway Vehicles. IEA/AIE 2009: 449-458

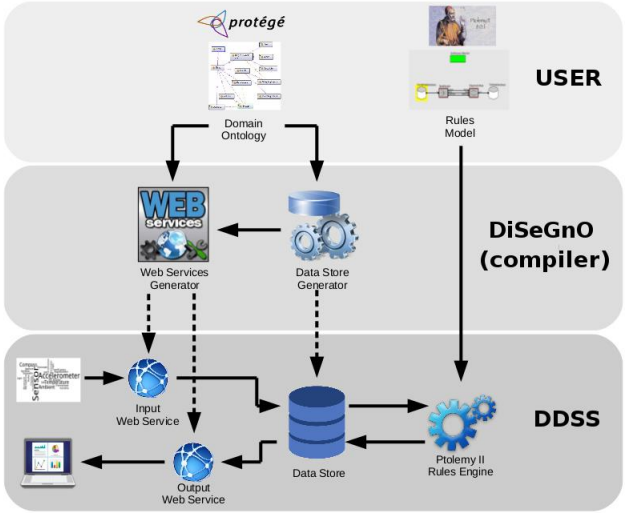


System monitoring and maintenance

Matteo Casu, Giuseppe Cicala, Armando Tacchella:
Ontology-based data access: An application to intermodal logistics. Information Systems Frontiers 15(5): 849-871 (2013)



Giuseppe Cicala, Marco De Luca, Marco Oreggia, Armando Tacchella:
A Multi-Formalism Framework To Generate Diagnostic Decision Support Systems. ECMS 2016: 628-634



Predictability of Autonomous Systems

DRAFT INTERNATIONAL STANDARD ISO/DIS 13482

ISO/TC **184**/SC **2** Secretariat: **SIS**

Voting begins on **2011-09-08** Voting terminates on **2012-02-08**

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

Robots and robotic devices — Safety requirements for non-industrial robots — Non-medical personal care robot

Robots et composants robotiques — Exigences de sécurité — Robots non médicaux pour les soins personnels

ICS 25.040.30

ISO/CEN PARALLEL PROCESSING

This draft has been developed within the International Organization for Standardization (ISO), and processed under the **ISO-lead** mode of collaboration as defined in the Vienna Agreement.

This draft is hereby submitted to the ISO member bodies and to the CEN member bodies for a parallel five-month enquiry.

Should this draft be accepted, a final draft, established on the basis of comments received, will be submitted to a parallel two-month approval vote in ISO and formal vote in CEN.

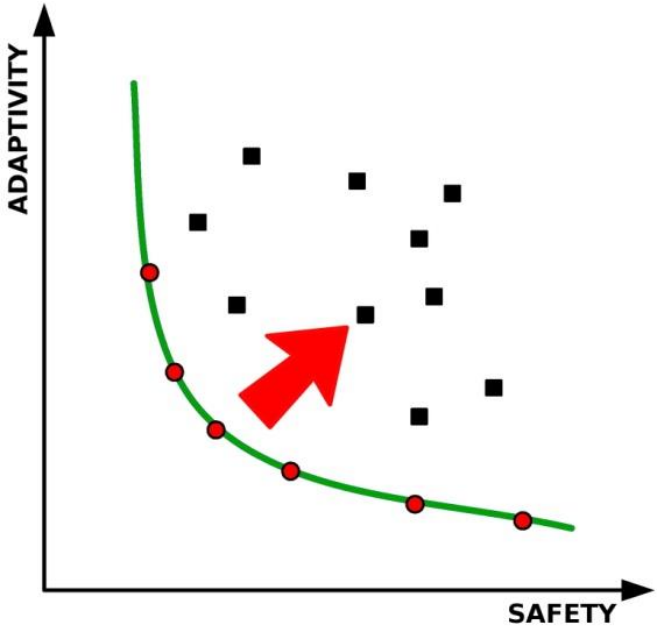
To expedite distribution, this document is circulated as received from the committee secretariat. ISO Central Secretariat work of editing and text composition will be undertaken at publication stage.

Pour accélérer la distribution, le présent document est distribué tel qu'il est parvenu du secrétariat du comité. Le travail de rédaction et de composition de texte sera effectué au Secrétariat central de l'ISO au stade de publication.

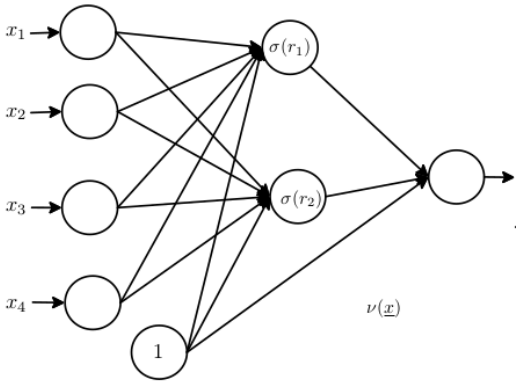
Adaptability, Safety, Performance...

ADAPTABILITY:
Capability to adapt internal parameters in the case of mutating environment mantaining a given performance

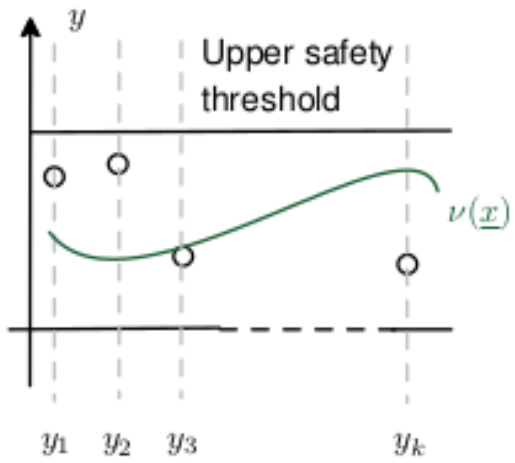
SAFETY:
Assurance that a dangerous behaviour is never reached by the system



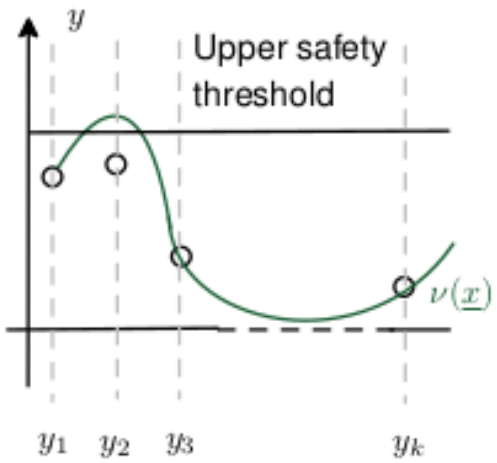
Analysis of Neural Nets



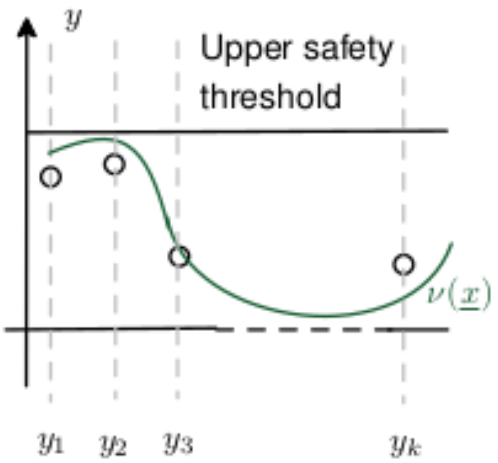
Neural Networks and other computational learning models are (or should become) commonplace in robotics and other autonomous systems



Safe but not accurate



Accurate but not safe

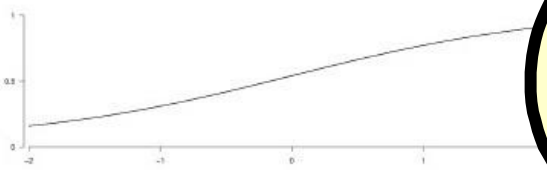


Accurate **and** safe

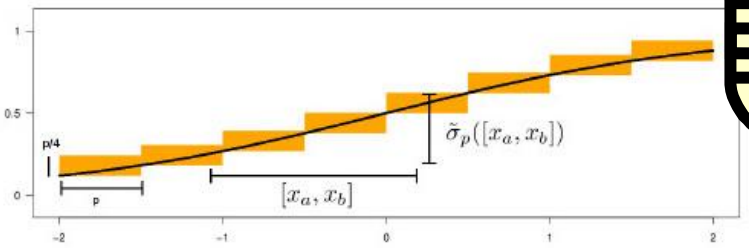
Analysis of Neural Nets

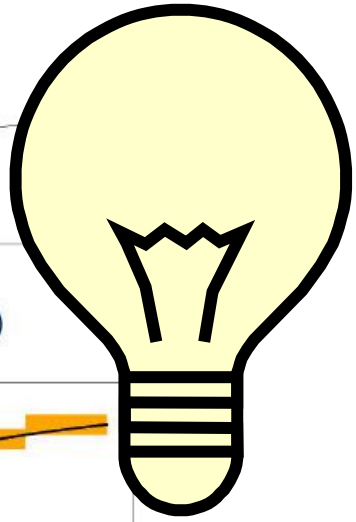
Interval arithmetics to consistently overapproximate activation functions

Logistic function $\sigma : \mathbb{R} \to (0, 1)$



Abstract logistic function $\tilde{\sigma}_p : [\mathbb{R}] \to [[0, 1]]$ $(p \in \mathbb{R}^+)$

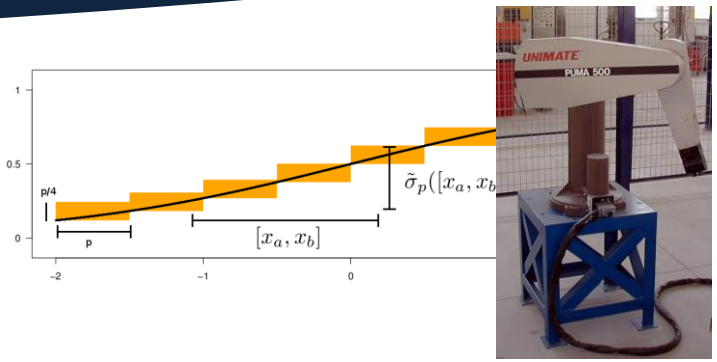




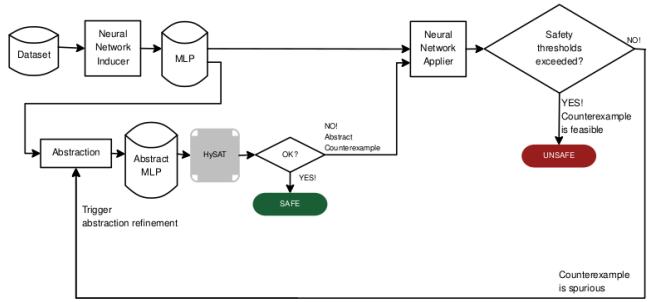
Satisfiability Modulo Theory (SMT) solvers do the «dirty work»

Analysis of Neural Nets

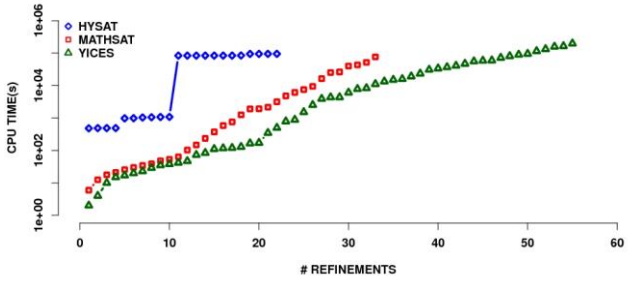
Luca Pulina, Armando Tacchella:
An Abstraction-Refinement Approach to Verification of Artificial Neural Networks. CAV 2010: 243-257



Luca Pulina, Armando Tacchella:
NeVer: a tool for artificial neural networks verification. Ann. Math. Artif. Intell. 62(3-4): 403-425 (2011)



Luca Pulina, Armando Tacchella:
Challenging SMT solvers to verify neural networks. AI Commun. 25(2): 117-135 (2012)



More recently...

Intriguing properties of neural networks

Christian Szegedy
Google Inc.

Wojciech Zaremba
New York University

Ilya Sutskever
Google Inc.

Joan Bruna
New York University

Dumitru Erhan
Google Inc.

Ian Goodfellow
University of Montreal

Rob Fergus
New York University
Facebook Inc.

Abstract

Deep neural networks are highly expressive models that have recently achieved state of the art performance on speech and visual recognition tasks. While their expressiveness is the reason they succeed, it also causes them to learn uninterpretable solutions that could have counter-intuitive properties. In this paper we report two such properties.

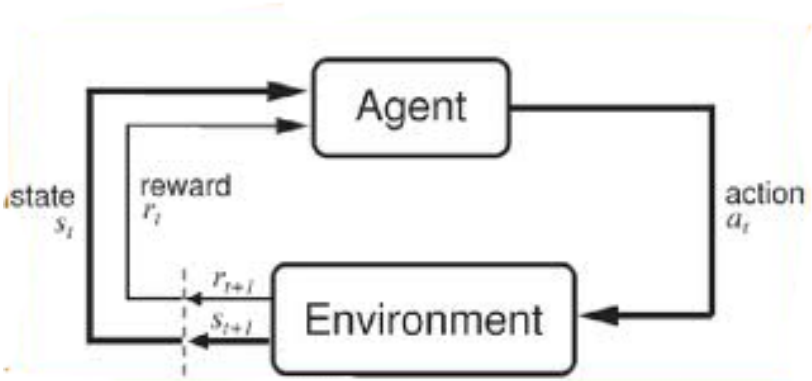
First, we find that there is no distinction between individual high level units and random linear combinations of high level units, according to various methods of unit analysis. It suggests that it is the space, rather than the individual units, that contains the semantic information in the high layers of neural networks.

Second, we find that deep neural networks learn input-output mappings that are fairly discontinuous to a significant extent. We can cause the network to misclassify an image by applying a certain hardly perceptible perturbation, which is found by maximizing the network's prediction error. In addition, the specific nature of

At some point the Machine Learning community realized that there might be a problem...

Citations of our CAV 2010 paper (source: Google Scholar)

Safe Reinforcement Learning



Reinforcement learning has been out there for quite some time

Recent advancements made it «fashionable» again

Safe Reinforcement Learning

Reinforcement learning is based on «trial and error»

- Fruitful actions are rewarded
- Wrong ones are penalized

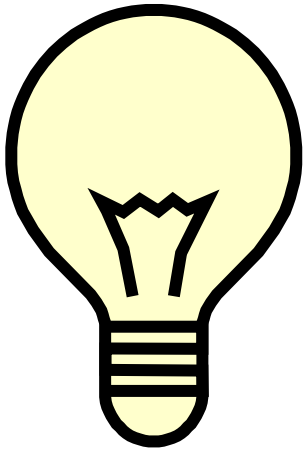
What if the robot breaks while trying?

Safe Reinforcement Learning

Markov Decision Problem

State-to-action mapping

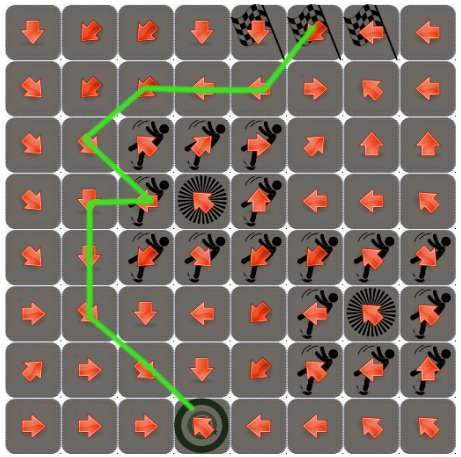
Markov Chain



Check the probability of reaching «bad» states with a probabilistic model checker

Safe Reinforcement Learning

Shashank Pathak, Luca Pulina, Armando Tacchella:
Verification and repair of control policies for safe reinforcement learning. Appl. Intell. 48(4): 886-908 (2018)

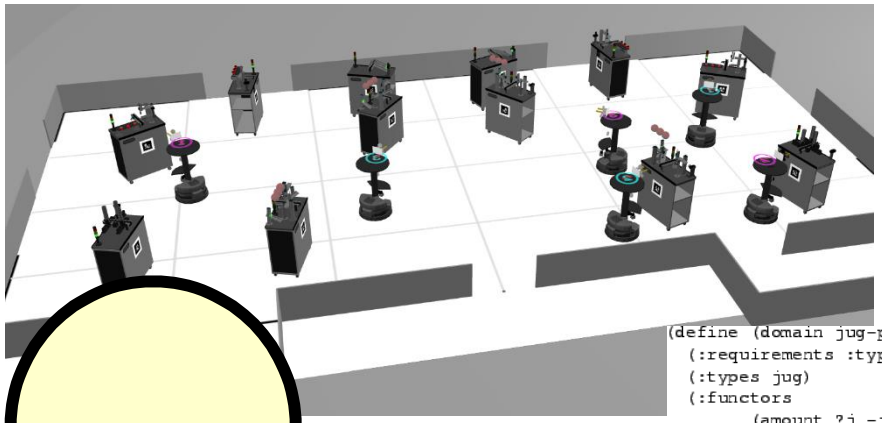


Francesco Leofante, Simone Vuotto, Erika Ábrahám, Armando Tacchella, Nils Jansen:
Combining Static and Runtime Methods to Achieve Safe Standing-Up for Humanoid Robots. ISoLA (1) 2016: 496-514

Shashank Pathak, Erika Ábrahám, Nils Jansen, Armando Tacchella, Joost-Pieter Katoen:
A Greedy Approach for the Efficient Repair of Stochastic Models. NFM 2015: 295-309

Optimal Planning

Optimal Planning



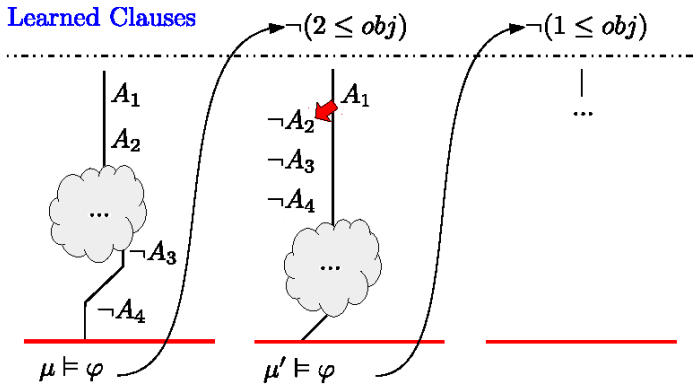
A formal model of the domain can be produced...

```
(define (domain jug-pouring)
  (:requirements :typing :fluents)
  (:types jug)
  (:functors
        (amount ?j -jug)
        (capacity ?j -jug)
          - (fluent number))
  (:action empty
        :parameters (?jug1 ?jug2 - jug)
        :precondition (fluent-test
                (>= (- (capacity ?jug2) (amount ?jug2))
                          (amount ?jug1)))
        :effect (and (change (amount ?jug1) 0)
                     (change (amount ?jug2)
                           (+ (amount ?jug1) (amount ?jug2)))))
)
```

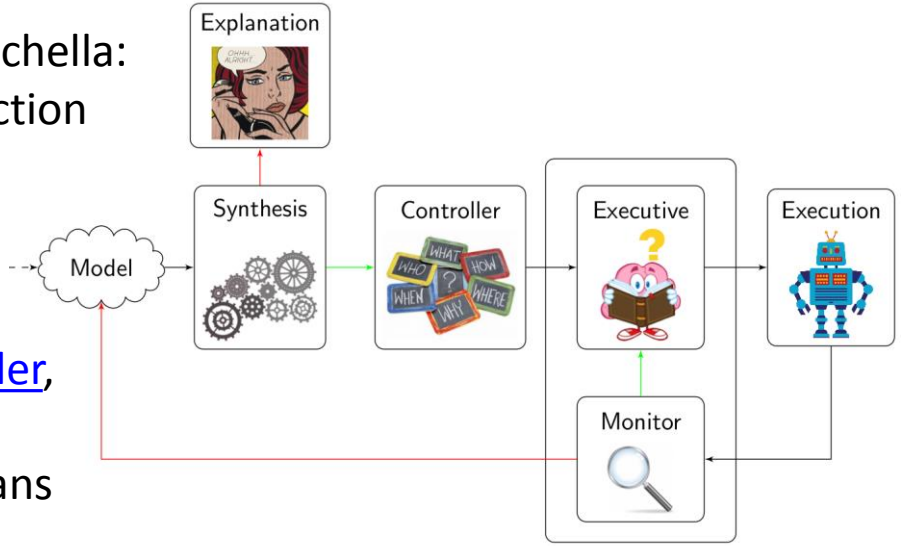
The solution of the model can be encoded to Optimization Modulo Theory

Optimal Planning

Francesco Leofante, Erika Ábrahám, Tim Niemueller, Gerhard Lakemeyer, Armando Tacchella: On the Synthesis of Guaranteed-Quality Plans for Robot Fleets in Logistics Scenarios via Optimization Modulo Theories. IRI 2017: 403-410



Francesco Leofante, Erika Ábrahám, Armando Tacchella: Task Planning with OMT: An Application to Production Logistics. IFM 2018: 316-325



Francesco Leofante, Erika Ábrahám, Tim Niemueller, Gerhard Lakemeyer, Armando Tacchella: Integrated Synthesis and Execution of Optimal Plans for Multi-Robot Systems in Logistics. Information Systems Frontiers 21(1): 87-107 (2019)

Security of Critical Infrastructure

Energy Sector

The U.S. energy infrastructure fuels the economy of the 21st century. The Department of Energy is the Sector-Specific Agency for the Energy Sector.

Chemical Sector

The Department of Homeland Security is designated as the Sector-Specific Agency for the Chemical Sector.

Water and Wastewater Systems Sector

The Environmental Protection Agency is designated as the Sector-Specific Agency for the Water and Wastewater Systems Sector.

Healthcare and Public Health Sector

The Department of Health and Human Services is designated as the Sector-Specific Agency for the Healthcare and Public Health Sector.

Nuclear Reactors, Materials, and Waste Sector

The Department of Homeland Security is designated as the Sector-Specific Agency for the Nuclear Reactors, Materials, and Waste Sector.

Transportation Systems Sector

The Department of Homeland Security and the Department of Transportation are designated as the Co-Sector-Specific Agencies for the Transportation Systems Sector.

Information Technology Sector

The Department of Homeland Security is designated as the Sector-Specific Agency for the Information Technology Sector.

Dams Sector

The Department of Homeland Security is designated as the Sector-Specific Agency for the Dams Sector. The Dams Sector comprises dam projects, navigation locks, levees, hurricane barriers, mine tailings impoundments, and other similar water retention and/or control facilities.

Critical Manufacturing Sector

The Department of Homeland Security is designated as the Sector-Specific Agency for the Critical Manufacturing Sector.

Food and Agriculture Sector

The Department of Agriculture and the Department of Health and Human Services are designated as the co-Sector-Specific Agencies for the Food and Agriculture Sector.

Defense Industrial Base Sector

The U.S. Department of Defense is the Sector-Specific Agency for the Defense Industrial Base Sector. The Defense Industrial Base Sector enables research, development, design, production, delivery, and maintenance of military weapons systems, subsystems, and components or parts to meet U.S. military requirements.

Financial Services Sector

The Department of the Treasury is designated as the Sector-Specific Agency for the Financial Services Sector.

Government Facilities Sector

The Department of Homeland Security and the General Services Administration are designated as the Co-Sector-Specific Agencies for the Government Facilities Sector.

Commercial Facilities Sector

The Department of Homeland Security is designated as the Sector-Specific Agency for the Commercial Facilities Sector, which includes a diverse range of sites that draw large crowds of people for shopping, business, entertainment, or lodging.

hcare and Public Health Sector

nications Sector

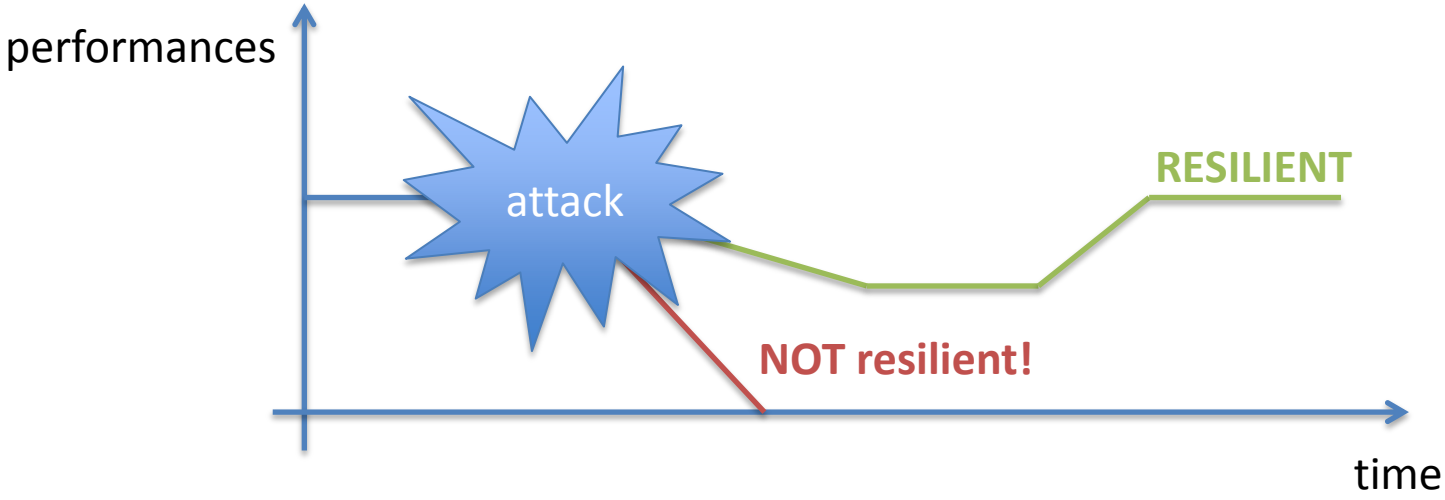
The Communications Sector is an integral component of the U.S. economy, underlying the operations of all businesses, public safety organizations, and government. The Department of Homeland Security is the Sector-Specific Agency for the Communications Sector.

Emergency Services Sector

The Department of Homeland Security is designated as the Sector-Specific Agency for the Emergency Services Sector. The sector provides a wide range of prevention, preparedness, response, and recovery services during both day-to-day operations and incident response.

Security of Critical Infrastructure

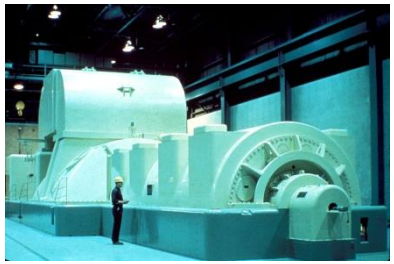
*"PPD-21 defines resilience as the ability to **prepare for** and **adapt** to changing conditions and **withstand** and **recover** rapidly from **disruptions**. Resilience includes the ability to **withstand and recover from deliberate attacks**, accidents, or naturally occurring threats or incidents."* [https://www.dhs.gov/what-security-and-resilience]



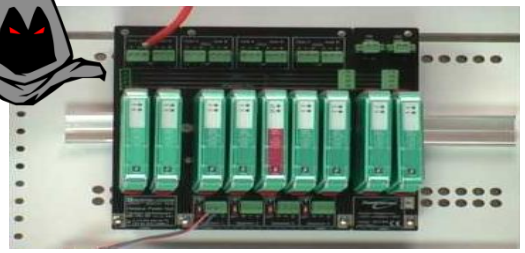
Security of Critical Infrastructure

Modification of set points

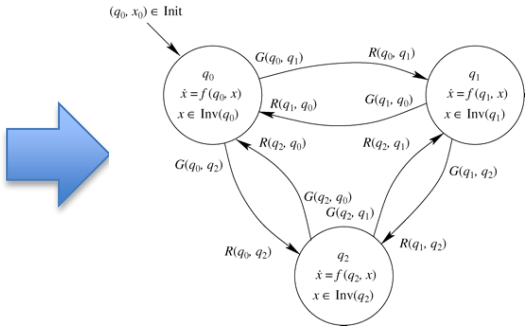
Modification of regulation parameters



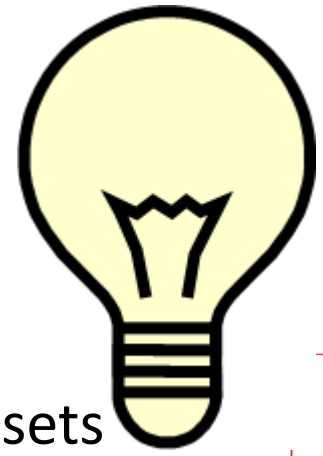
Modification of state estimation



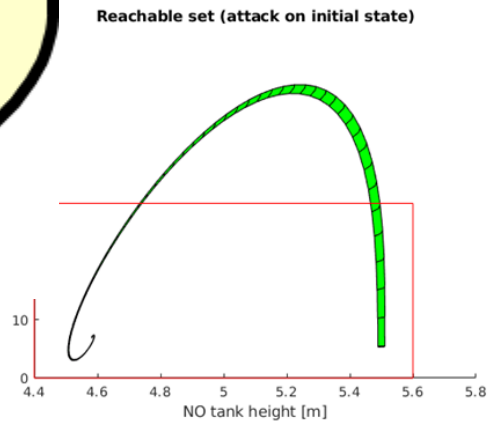
Security of Critical Infrastructure

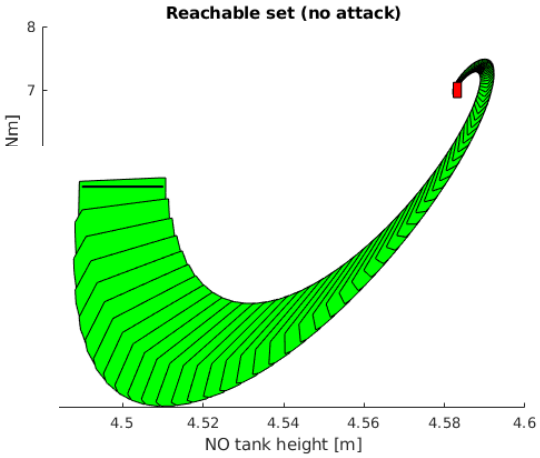


Model systems
as hybrid automata



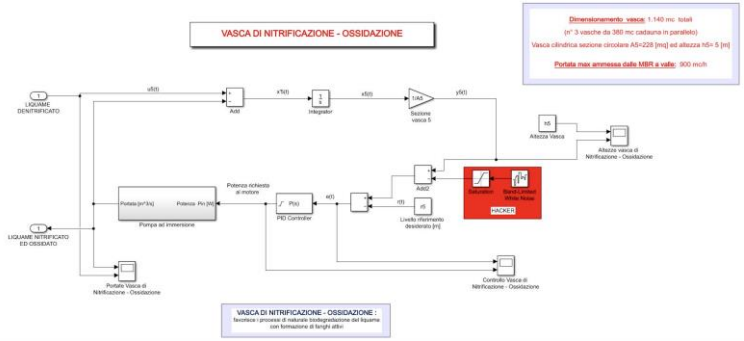
Compare reachable sets
under different hypotheses



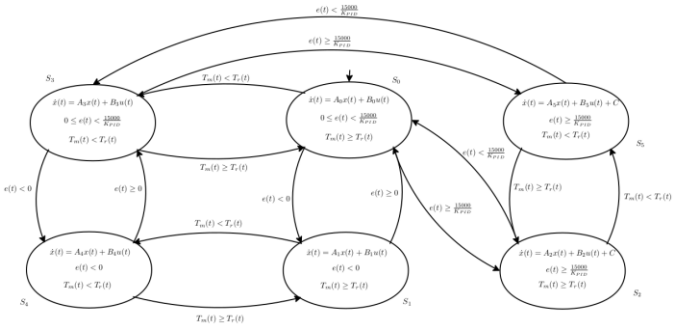


Security of Critical Infrastructure

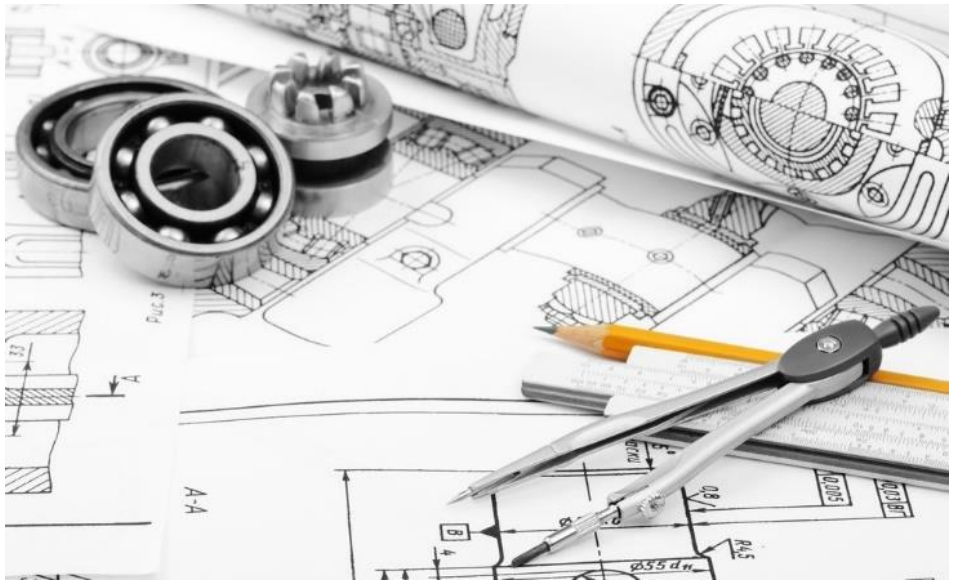
Giuseppina Murino, Alessandro Armando, Armando Tacchella: Resilience of Cyber-Physical Systems: an Experimental Appraisal of Quantitative Measures. CyCon 2019: 1-19



Giuseppina Murino, Armando Tacchella: Concrete vs. Symbolic Simulation To Assess Cyber-Resilience Of Control Systems. ECMS 2018: 433-439

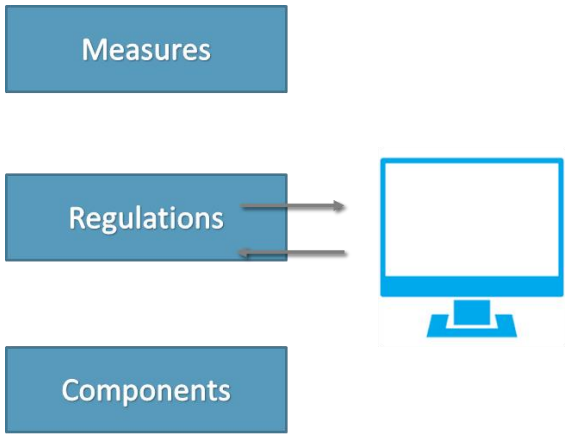


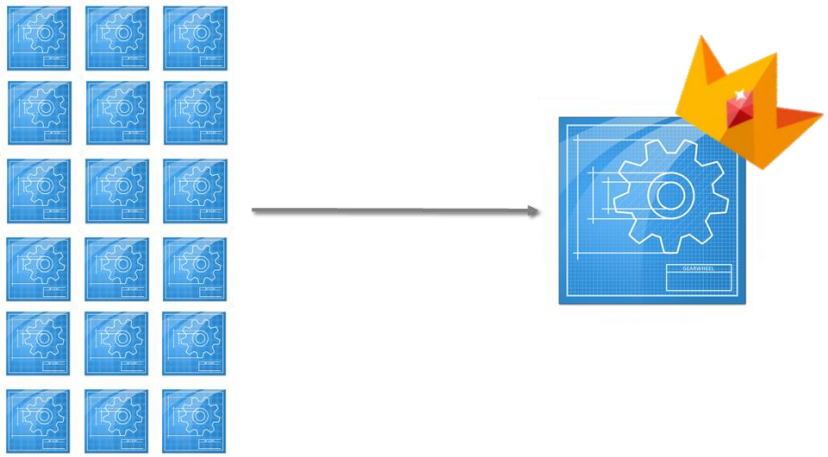
Computer-automated design



Produce designs given system specifications

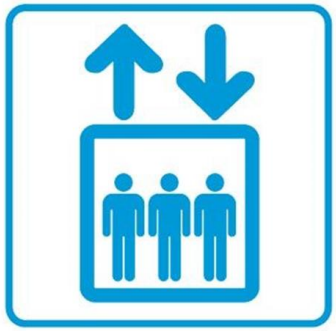
Product configuration is a subtask





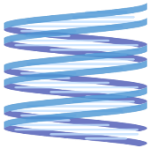
Computer-automated design

Elevator Design

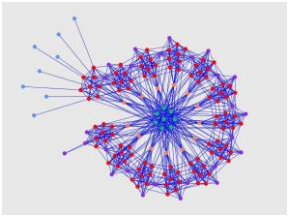


Heuristic approach

Genetic algorithms

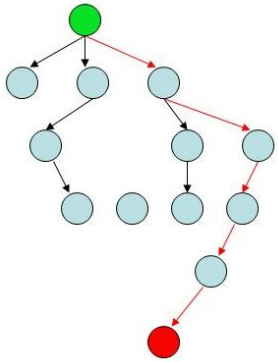


Satisfiability Modulo Theory



Computer-automated design

Leopoldo Annunziata, Marco Menapace, Armando Tacchella:
Computer Intensive Vs. Heuristic Methods In Automated Design Of Elevator Systems. ECMS 2017: 543-549

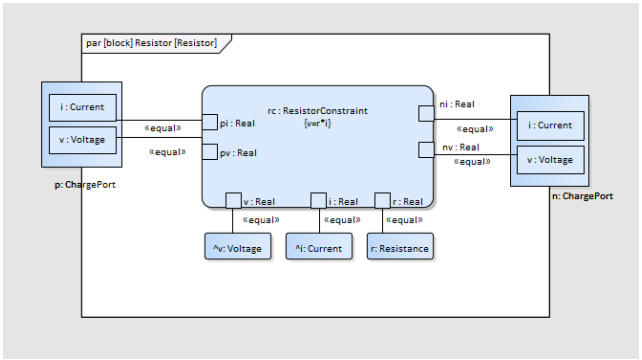


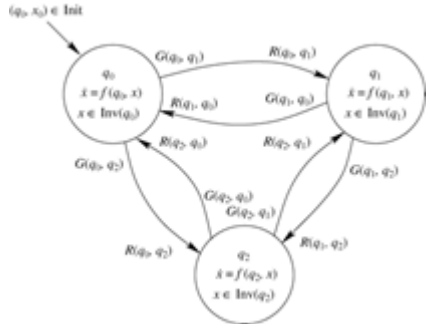
Stefano Demarchi, Marco Menapace, Armando Tacchella:
Automating Elevator Design with Satisfiability Modulo Theories ICTAI 2019 (to appear)

$$\begin{cases} y_{cf} - h_{br} \geq y_{shaft} \\ y_{cf} + d_{cf} + d_{br} \leq y_{shaft} + d_{shaft} \\ y_{cf} + y_{gear} - y_{car} \geq 0 \\ y_{car} + d_{car} - y_{cf} - y_{gear} - dcr \geq 0 \\ y_{cf} + y_{gear} - y_{car} < max_{oh} \\ y_{car} + d_{car} - y_{cf} - y_{gear} - dcr < max_{oh} \end{cases}$$

$$\begin{cases} x_{cd} \geq x_{shaft} \\ x_{ld} \geq x_{shaft} \\ x_{cd} + la_{cd} + ra_{cd} \leq x_{shaft} + w_{shaft} \\ x_{ld} + la_{ld} + ra_{ld} \leq x_{shaft} + w_{shaft} \\ x_{cd} + la_{cd} - \frac{opening}{2} \geq x_{car} \\ x_{cd} + la_{cd} + \frac{opening}{2} \leq x_{car} + w_{car} \\ x_{ld} + la_{ld} + \frac{opening}{2} + w_{frame} \leq x_{shaft} + w_{shaft} \end{cases}$$

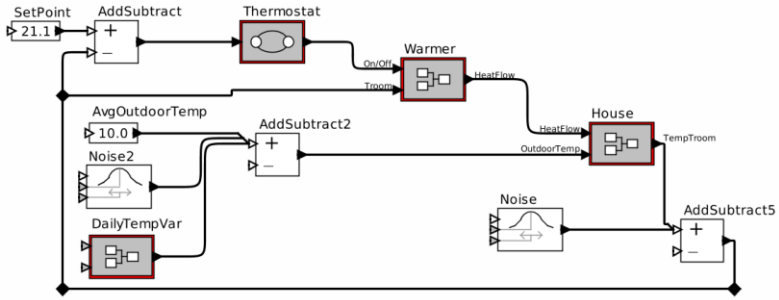
Edge of Tomorrow: go formal





$$
\begin{cases}
y_{cf} - h_{br} \geq y_{shaft} \\
y_{cf} + d_{cf} + d_{br} \leq y_{shaft} + d_{shaft} \\
y_{cf} + y_{gear} - y_{car} \geq 0 \\
y_{car} + d_{car} - y_{cf} - y_{gear} - dcr \geq 0 \\
y_{cf} + y_{gear} - y_{car} < max_{oh} \\
y_{car} + d_{car} - y_{cf} - y_{gear} - dcr < max_{oh}
\end{cases}
$$

```
(define (domain jug-pouring)
  (:requirements :typing :fluents)
  (:types jug)
  (:functors
        (amount ?j -jug)
        (capacity ?j -jug)
          - (fluent number))
  (:action empty
        :parameters (?jug1 ?jug2 - jug)
        :precondition (fluent-test
                (>= (- (capacity ?jug2) (amount ?jug2))
                          (amount ?jug1)))
        :effect (and (change (amount ?jug1) 0)
                     (change (amount ?jug2)
                           (+ (amount ?jug1) (amount ?jug2)))))
)
```



Edge of Tomorrow: fail

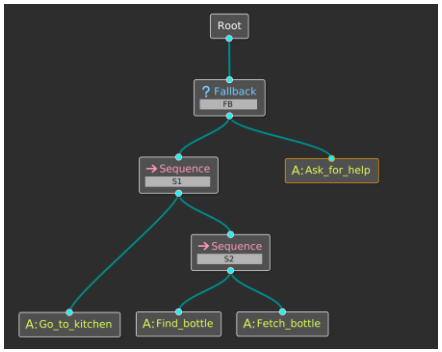
Poor integration

What did you say?
Formal Semantics?

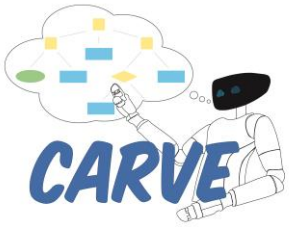
Computational Complexity

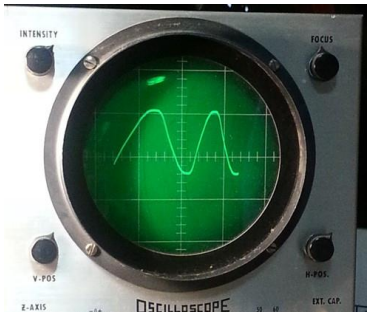
«Dark side of the moon » effect

Edge of Tomorrow: repeat



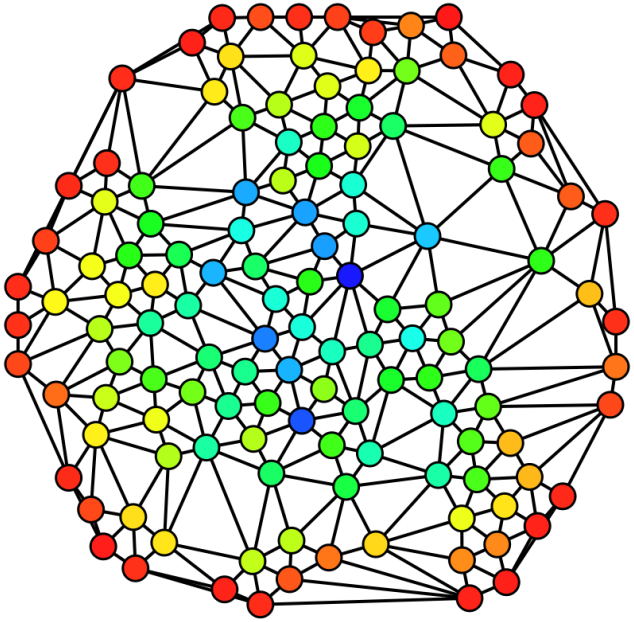
RobMoSys





SCOPE

Formal methods to improve predictability of robots



BRAIN-ISAR

Formal methods to improve predictability of graph-based models

Questions?

Photo courtesy of Dario Guidotti

If you don’t ask now, you may have troubles catching me later...