POLITECNICO
MILANO 1863

AIRLAB
ARTIFICIAL INTELLIGENCE AND ROBOTICS LAB

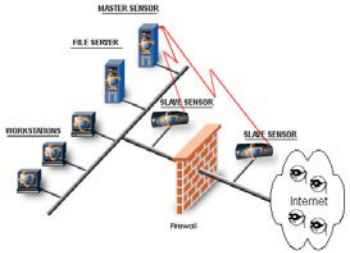
Intelligent Agents for Detecting Anomalies in Complex Systems

Francesco Amigoni
francesco.amigoni@polimi.it

Artificial Intelligence and Robotics Laboratory - Politecnico di Milano

Global vs. partial models of complex systems

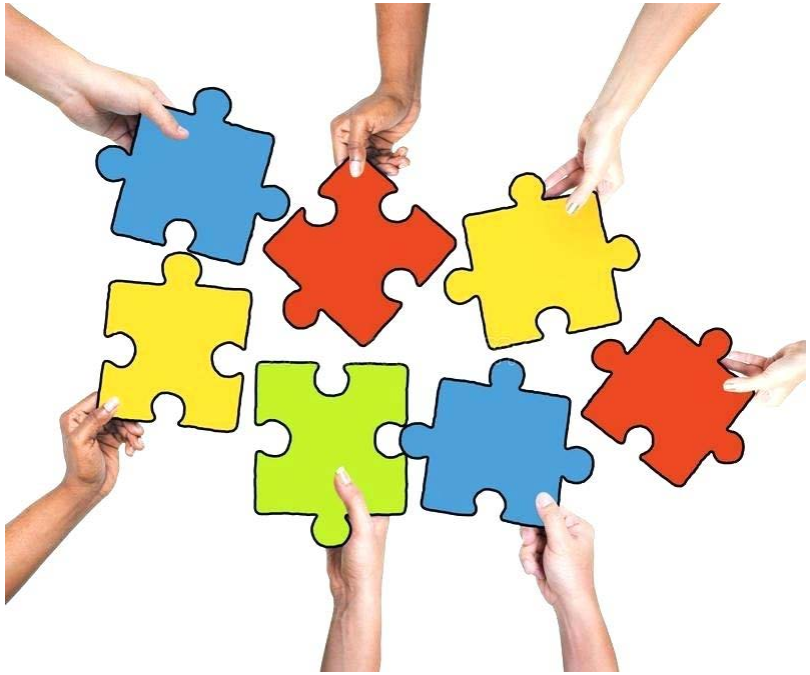
- Several complex systems do not admit **global models** capturing all their aspects but **partial models** that describe individual sub-systems or specific aspects
- Examples: heart rate, intrusion detection in computer networks, water resources, satellites, …



- Possible solutions: (black-box) data-driven approaches or **aggregating partial models**

Aggregating partial models: Overview of the idea

- Partial models are embedded in **intelligent agents**
  - Agent = independent autonomous AI system
- An agent detects only some anomalies and returns an anomaly probability
- The problem is to design the **interaction mechanisms** for aggregating anomaly probabilities returned by the agents to obtain a global anomaly probability
- Examples of interaction mechanisms: average, max or min, cooperative negotiation, voting, …



Examples

- **Heart rate**: agents relate heart rate to different physiological quantities
  [Amigoni et al., Artif Intell Med, 2003] [Amigoni et al., IEEE T Inf Technol B, 2006]

$$\mathrm{QT} = C_1 - C_2 \times \exp^{-\frac{C_2}{\mathrm{HR}}} \qquad \mathrm{HR} = \begin{cases} 65 & RR \leq 15 \\ 2,8 \times \mathrm{RR} + 25 & 15 < RR < 45 \\ 150 & RR \geq 45 \end{cases}$$

- **Intrusion detection systems in computer networks**: agents capture anomalies on different aspects
  [Amigoni et al., Proc. IAT, 2008]
  - Number of syn-flags (opening of new connections), number of reset-flags (aborted connections), most used ports, …

- **Water resources systems**: agents represent the views of different stakeholders
  [Mason et al., Water Resour Res, 2018]

Case study: Anomaly detection from data of the Cryosat-2 satellite

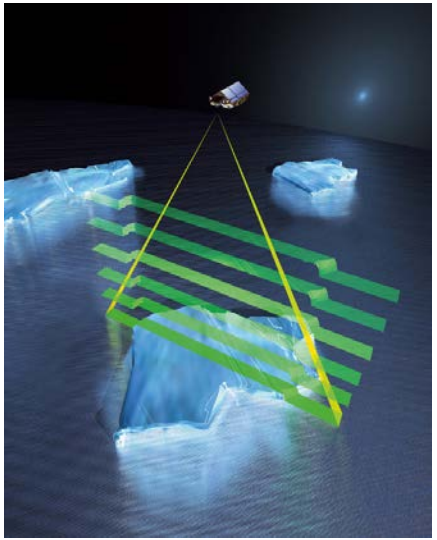
Flight-Control Team Multi-Agent System (FCTMAS) Study conducted by Politecnico di Milano, European Space Agency (ESA) - Advanced Mission Concepts and Technologies Office, and Telespazio Vega Deutschland GmbH
[Amigoni et al., Proc. IAS, 2018]

The application context

«ESA's CryoSat mission is dedicated to measuring the thickness of polar sea ice and monitoring changes in the ice sheets that blanket Greenland and Antarctica» [www.esa.int]

Cryosat-2 satellite

The flight control team receives a lot of data from the satellite and has to identify anomalous behaviors



The anomaly detection problem

Events log file

```
2013-08-03 07:08:13.553  20755 2   BEHVLimCPB    crymca  Information  Log     2013.215.05.08.50.190 DHT30304 VAL: ON STATE: ON STATUS
  limit is back to nominal
2013-08-03 07:08:11.275  13524 1   CMDHveri      crymca  Information  Log     TC: SSC09000, APID: 812, SSC: 13900,
  set stage: EV\_APP\_ACCEPT status to: PASSED
...
2013-08-22 23:32:20.754  23001 1   TPKT          crymca  Error        System  4 Missing Source Packets, APID = 68, VCID = 0,
  SSC = 14894, Time = 2013-08-22T23:32:18.710216
2013-08-22 23:32:19.511  10307 1   NCDUadmi      crymca  Warning      Log     NCDU:TM007 Data gap on TM link VC 250/0,
  Mode: Onl-TIM from KR14 . Reason: unk, Size: 15
2013-08-22 23:32:19.448  13842 1   CMDHmplx      crymca  Information  Log     Commanding link status set to: TC: NO RF, TM: GREEN
2013-08-22 23:32:19.443  14778 1   TPKT          crymca  Error        System  Unexpected spill-over data [fhp = 730]
```

Anomaly models for single variables (anomaly = deviation from nominal behavior) that return anomaly probabilities

Aggregation of anomlay probabilities returned by anomaly models for single variables to detect system-level anomalies

Anomaly models for single variables

Nominal models

Numerical values

- Neural networks
- Autoregressive Moving Average (ARMA)
- …

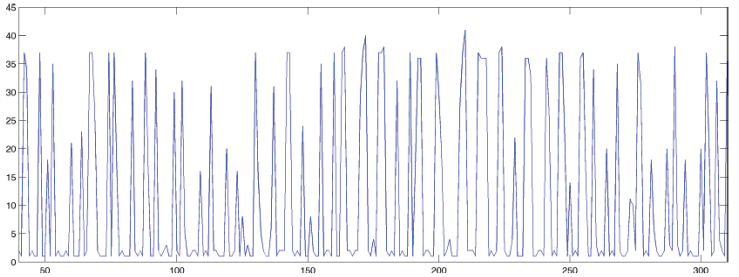
Categorical values

- Markov chains
- …

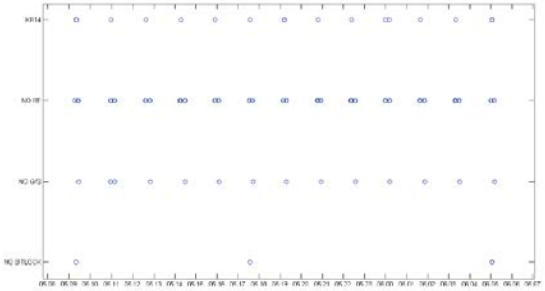
Anomaly models

$D_i: state \longmapsto anomaly\ probability$

Missing source packets



Commanding link status TC

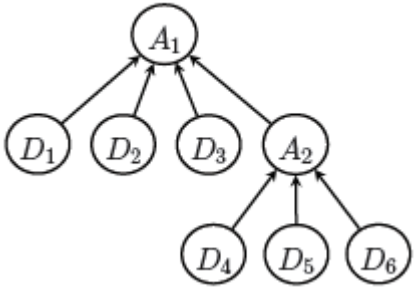


The aggregation problem

Given anomaly models for single variables $\{D_1, D_2, \dots, D_I\}$, find their best aggregation

Aggregation is a tree

- D_i are the leaves
- **Aggregation functions** A_j are other nodes
- The root returns the system's anomaly probability



The best aggregation tree maximizes the identification of anomalies *at the system level*

Aggregation functions (1)

Maximum aggregation functions

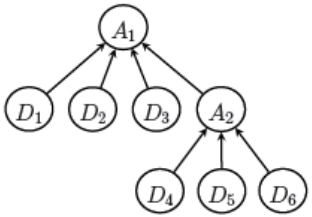
$$A_j\left(p_1, p_2, \ldots, p_{k_j}\right) = \max\{p_1, p_2, \ldots, p_{k_j}\}$$

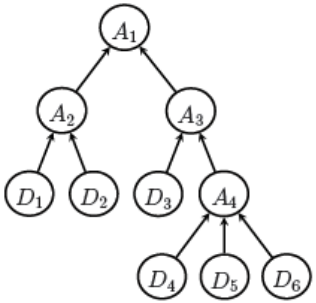
- All aggregation trees are equivalent

Average aggregation functions

$$A_j\left(p_1, p_2, \ldots, p_{k_j}\right) = \frac{p_1 + p_2 + \cdots + p_{k_j}}{k_j}$$

- Equivalence classes of aggregation trees



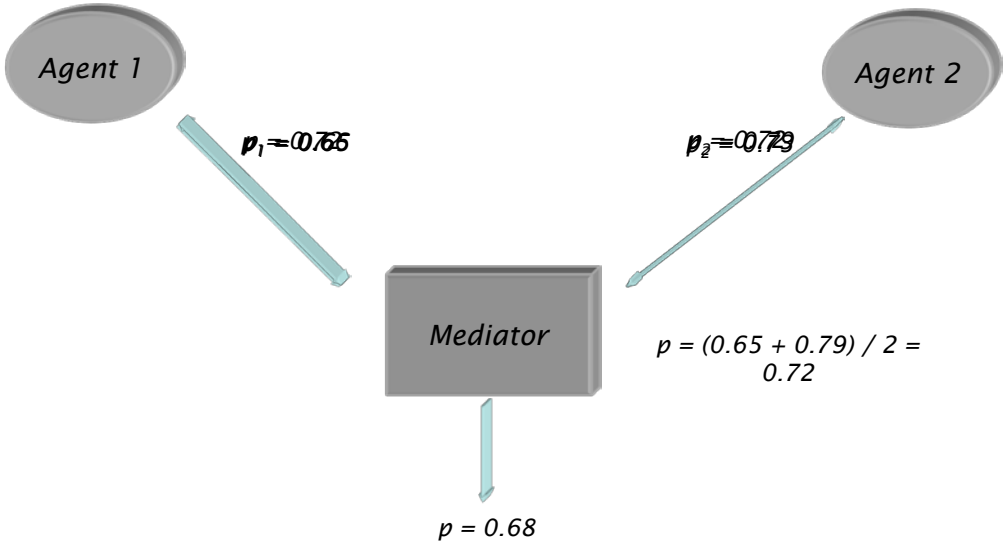


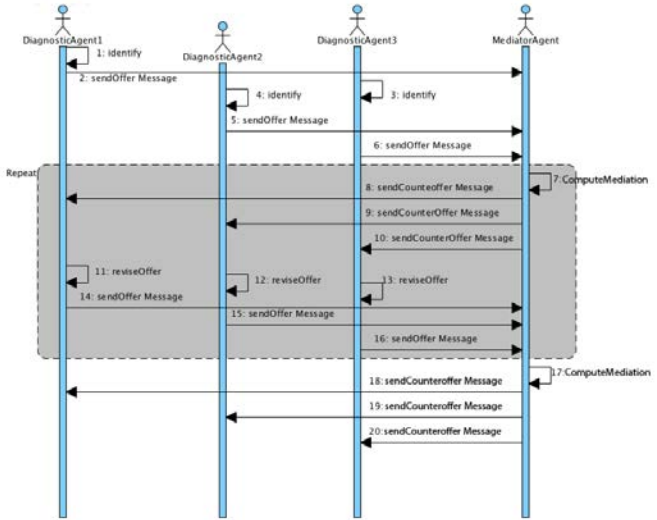
$$\frac{p_1}{4} + \frac{p_2}{4} + \frac{p_3}{4} + \frac{p_4}{12} + \frac{p_5}{12} + \frac{p_6}{12}$$

Aggregation functions (2)

Cooperative negotiation aggregation functions

- Iterative procedure to find agreements between children
- Proposals and counter-proposals [Amigoni and Gatti, JAAMAS, 2007]
- Equivalence classes of aggregation trees

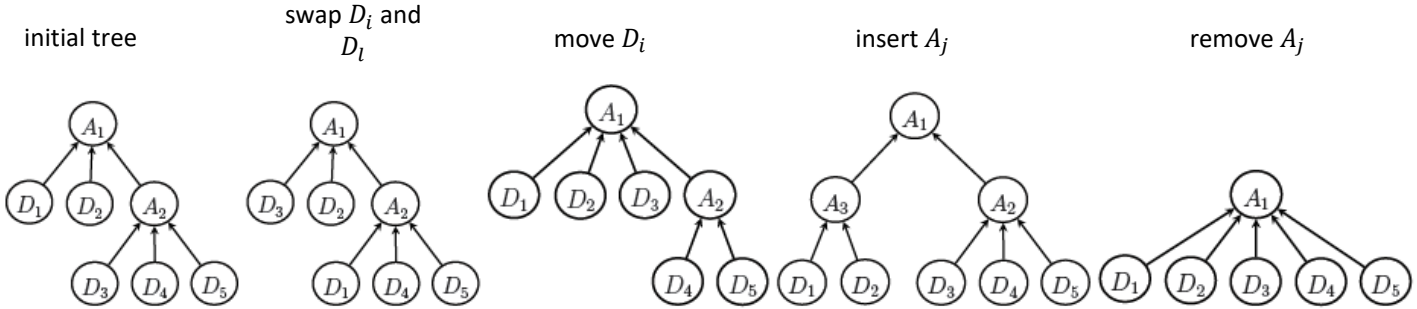




Solving algorithms (1)

Enumeration algorithm

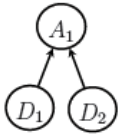
Simulated annealing algorithm: local moves with decreasing probability of accepting a move that worsens the aggregation tree



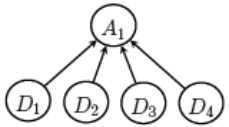
Solving algorithms (2)

Greedy algorithm: local moves that strictly improves the aggregation tree

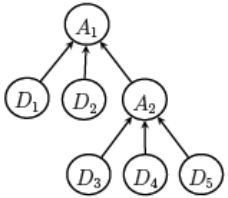
initial tree



add D_i



insert A_j



- Does not cover the entire solution space

Experimental evaluation

Implementation: each node of the aggregation tree is an independent software agent in JADE

Anomaly models of 5 variables built from data of Cryosat-2 collected in February 2013

Anomaly at the system level with probability > 0.2

Cost matrix

	System ok	System anomalous
Classified ok	0	50
Classified anomalous	10	0

Test on data of Cryosat-2 collected in July and August 2013

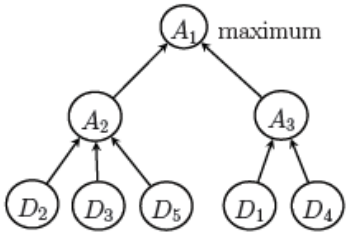
Injection of anomalies at the system level

Experiment
Single variable anomaly → System anomaly

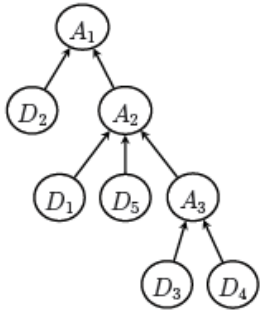
#	D_1	D_2	D_3	D_4	D_5	A/N
01	0.5					A
02	0.5					A
03		0.3				A
04		0.4				A
05		0.7				A
06		0.5				A
07			0.3			A
08			0.8			A
09			0.9			A
10				0.7		A
11				0.7		A
12					0.8	A
13					0.4	A
14					0.4	A
15					0.3	A
16	0.4					A
17	0.8					A
18	0.5					A
19		0.5				A
20		0.5				A

Best solutions

cost = 0



cost = 100

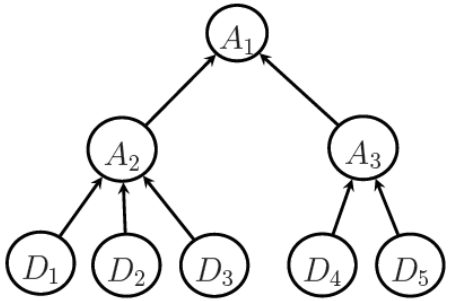


A_j: cooperative negotiation

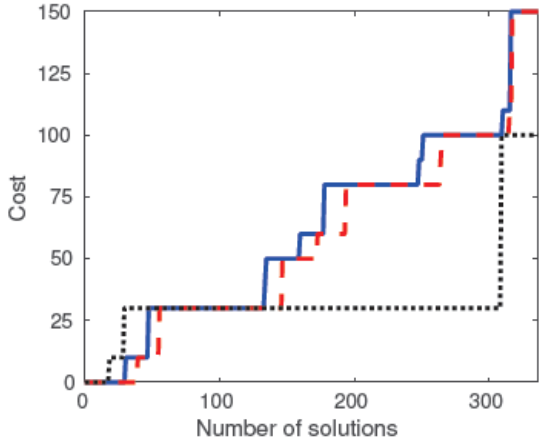
Experiment
Correlated variables

Optimal solutions with cost = 0

#	D_1	D_2	D_3	D_4	D_5	A/N
01				0.60		A
02		0.20				N
03	0.10	0.15	0.35			N
04	0.10	0.12	0.35			N
05	0.10	0.10	0.35	0.70		A
06			0.20			N
07				0.50		A
08	0.10	0.09	0.25			N
09	0.20	0.17	0.35			N
10	0.25	0.18	0.35			A
11			0.20			N
12				0.40		A
13			0.10	0.50		A
14	0.10	0.10	0.25	0.60		A
15	0.10	0.08	0.25			N
16	0.05	0.07	0.20			N
17	0.30	0.28	0.45			A
18	0.35	0.30	0.55			A
19	0.25	0.10	0.35			A
20	0.10	0.12	0.25			N



A_j: cooperative negotiation

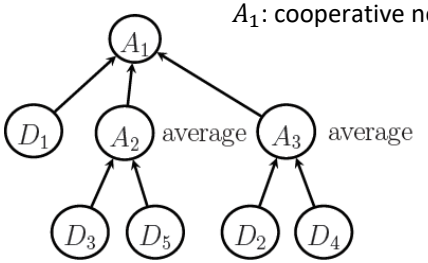


Blue: enumeration
Red: simulated annealing
Black: greedy

Experiment
Subsets of correlated variables

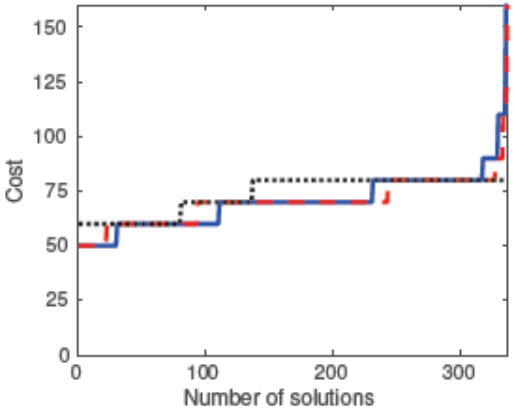
Optimal solutions produce 2 false positives (cost = 20)

#	D_1	D_2	D_3	D_4	D_5	A/N
01	0.15	0.20	0.25			N
02	0.10	0.15	0.20			N
03	0.25	0.28	0.32			N
04	0.30	0.32	0.35			A
05	0.25	0.35	0.40			A
06	0.35	0.40	0.50			A
07	0.45	0.55	0.60			A
08	0.10	0.09	0.25			N
09	0.20	0.17	0.35			N
10	0.25	0.18	0.35			N
11				0.30	0.20	N
12				0.40	0.30	N
13				0.50	0.40	N
14				0.60	0.45	A
15				0.65	0.55	A
16				0.70	0.40	A
17				0.80	0.70	A
18	0.05	0.10	0.15	0.20	0.30	N
19	0.15	0.20	0.25	0.30	0.40	N
20	0.20	0.25	0.30	0.40	0.50	N



A_1: cooperative negotiation

Blue: enumeration
Red: simulated annealing
Black: greedy



Experiment
Scalability

Enumeration algorithm

Number of anomaly models for single variables

I	5	6	7	8	9	10	11	12
trees	$3 \cdot 10^2$	$6 \cdot 10^3$	$1 \cdot 10^5$	$4 \cdot 10^6$	$2 \cdot 10^8$	$8 \cdot 10^9$	$4 \cdot 10^{11}$	$3 \cdot 10^{13}$
time	0 s	0 s	0 s	15 s	9 m	8 h	19 d*	3.7 y*

* estimated

Simulated annealing and greedy algorithms are tunable by the user (e.g., number of iterations)

Guidelines for applications

System anomalies related to single variables

→ Use maximum aggregation functions

System anomalies related to correlated variables

→ Use cooperative negotiation aggregation functions

System anomalies related to subsets of correlated variables

→ Use combinations of aggregation functions

Simulated annealing is the most effective algorithm for building aggregation trees

Conclusions

Aggregating partial models using mechanisms for intelligent agents' interaction could provide a solution for anomaly detection in complex systems

The approach is similar to some ensemble approaches, but it provides more structure and better understanding of the systems (good for diagnosis)

Thank you!