Detection and Quantification of Events in Stochastic Systems

PhD defense, Hugo Bazille

Under the supervision of Eric Fabre and Blaise Genest

December 2nd, 2019

Context

Rise of the machines...

We rely more and more on automatized processes:

Banking

Transportation

Communication

Health

... and the problems they induce

- Security
- Efficiency
- Confidentiality
- ...

Imperfect information

In most systems, exact state is not known

- Cost of sensors,
- Opacity...

Imperfect information

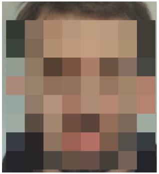
In most systems, exact state is not known

- Cost of sensors,
- Opacity...

Imperfect information

In most systems, exact state is not known

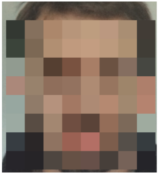
- Cost of sensors,
- Opacity...



Imperfect information

In most systems, exact state is not known

- Cost of sensors,
- Opacity...



Challenge

We want to know if it is possible to recover some hidden information!

Historically: qualitative verification

Can one always recover some hidden information on this system?

A huge background literature on...

- Deductive verification
- Testing
- **Model-checking**

Each approach has its advantages/drawbacks:
For Model-checking, fully automated but **need for a model**.

Recently: quantitative verification

Important quantitative properties

- **How likely** can one recover some hidden information on this system?
- **How fast**?

Recently: quantitative verification

Important quantitative properties

- **How likely** can one recover some hidden information on this system?
- **How fast**?

Quantification using stochastic models

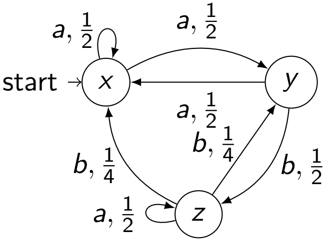
- Natural way to quantify for questions such as "how likely": a probability,
- Natural representation for many real systems, eg telecommunications.

Model: Labeled Markov Chains

In this talk: discrete states, discrete time.

LMC

- Markov chain with labels representing observations,
- Sum of outgoing transition probabilities is 1.



- Same expressive power as Hidden Markov Models (used in control community).
- The model is assumed to be known, but it may not be easy to obtain.

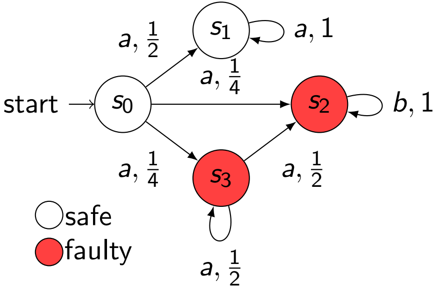
Plan

Plan

Diagnosability

Diagnosability

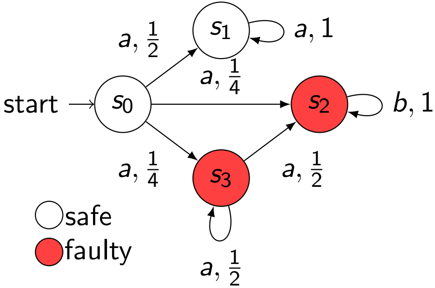
Ability to retrieve a binary information (occurrence of a "fault") from an observation of the system. In this talk: permanent faults.



Diagnosability

Diagnosability

Ability to retrieve a binary information (occurrence of a "fault") from an observation of the system. In this talk: permanent faults.



aaaaab is faulty and non ambiguous: can diagnose.

aaaaa is ambiguous: cannot diagnose.

Can we detect any fault occurrence in bounded time?

Faults and diagnosis

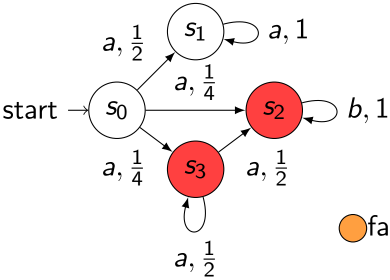
An LMC is

- diagnosable if there is no faulty ambiguous infinite execution,
- A-diagnosable if the probability of faulty ambiguous infinite execution is 0.

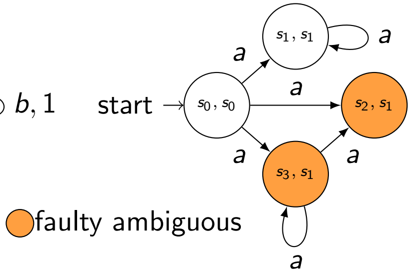
Diagnosability and twin machine

Twin machine

Synchronized (unprobabilized) product $A_1 \times A_{1\,C}$.



LMC A_1

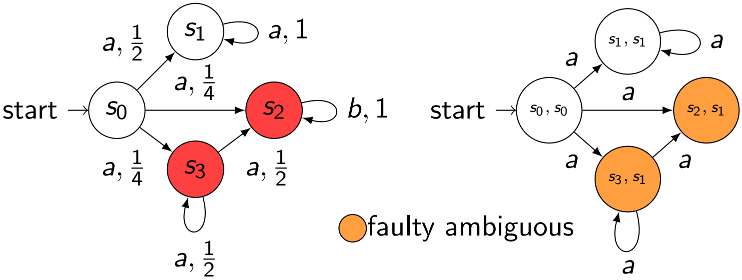


its twin machine

Diagnosability and twin machine

Twin machine

Synchronized (unprobabilized) product $A_1 \times A_{1\,C}$.



LMC A_1

its twin machine

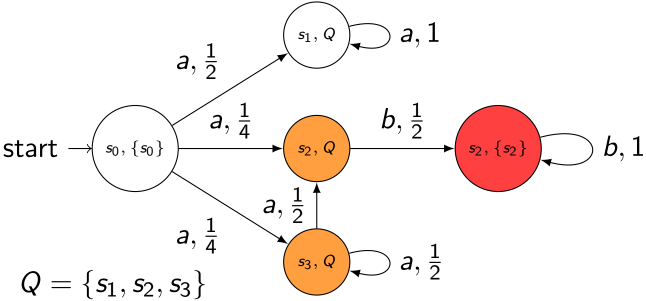
Theorem[YL02]

A_1 is diagnosable if there is no faulty ambiguous loop in the twin machine. (NLOGSPACE complete)

A-diagnosability and diagnoser

Diagnoser

Synchronized product $A_1 \times 2^{A_1}$.

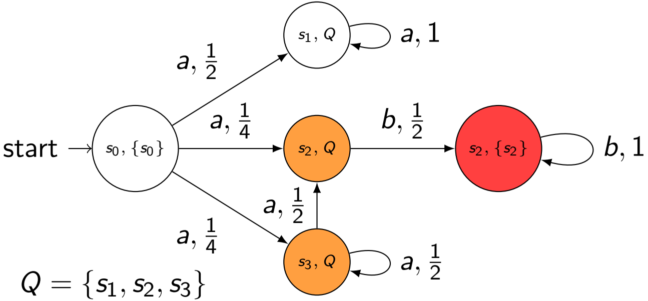


$Q = \{s_1, s_2, s_3\}$

A-diagnosability and diagnoser

Diagnoser

Synchronized product $A_1 \times 2^{A_1}$.



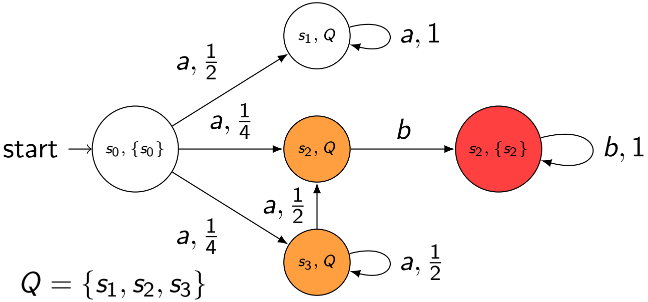
Theorem[BHL14]

A_1 is A-diagnosable if there is no faulty ambiguous loop in a BSCC of the diagnoser. (PSPACE-complete)

Plan

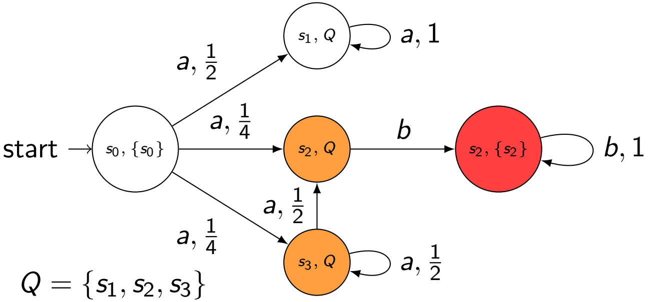
Quantitative diagnosis

What can we say when not all faulty executions are diagnosable?
What can we say about the time between a fault and its detection?



Quantitative diagnosis

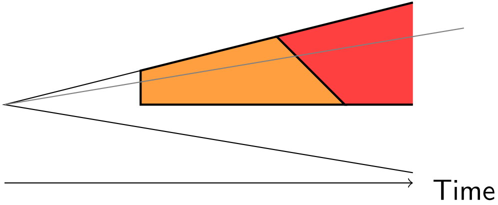
What can we say when not all faulty executions are diagnosable?
What can we say about the time between a fault and its detection?



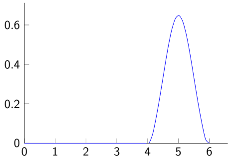
Probability of diagnosis: 1
Mean time before detection (conditionally to detection occurring): 2

Probability distribution

Be more precise than mean time?
Can we have the whole probability distribution of time to diagnosis?



Space of trajectories



Probability distribution of detection delay

Moments

A computable quantity: moments

Moment of order n: $\mu_n = \mathbb{E}[X^n] = \sum x^n \, \mathbb{P}(x)$

- Mean time: μ_1. Variance: $\mu_2 - \mu_1^2$...

Moments

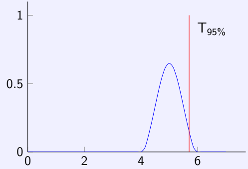
A computable quantity: moments

Moment of order n: $\mu_n = \mathbb{E}[X^n] = \sum x^n \, \mathbb{P}(x)$

- Mean time: μ_1. Variance: $\mu_2 - \mu_1^2$...

Concentration bounds

Markov's inequality: $\mathbb{P}(|X| \geq \alpha) \leq \frac{\mu_n}{\alpha^n}$



Moments

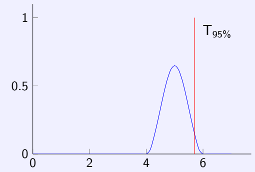
A computable quantity: moments

Moment of order n: $\mu_n = \mathbb{E}[X^n] = \sum x^n \, \mathbb{P}(x)$

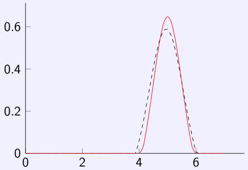
- Mean time: μ_1. Variance: $\mu_2 - \mu_1^2$...

Concentration bounds

Markov's inequality: $\mathbb{P}(|X| \geq \alpha) \leq \frac{\mu_n}{\alpha^n}$



Approximate the distribution?



Moments

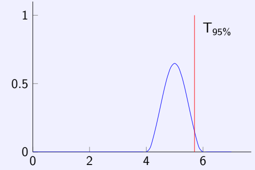
A computable quantity: moments

Moment of order n: $\mu_n = \mathbb{E}[X^n] = \sum x^n \, \mathbb{P}(x)$

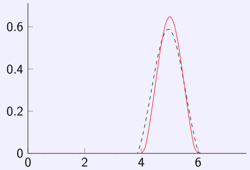
- Mean time: μ_1. Variance: $\mu_2 - \mu_1^2$...

Concentration bounds

Markov's inequality: $\mathbb{P}(|X| \geq \alpha) \leq \frac{\mu_n}{\alpha^n}$



Approximate the distribution?



Theorem [CDC17,FoSSaCS18]

One can compute the n first moments of the detection time distribution.

Plan

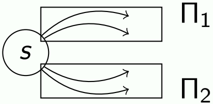
Combination of trajectories

Moment of order n on paths length of Π:

$\mu_n(\Pi) = \sum_{\pi \in \Pi} |\pi|^n \, \mathbb{P}(\pi)$

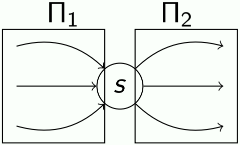
For disjoint union of paths

$$\mu_n(\Pi_1 \uplus \Pi_2) = \mu_n(\Pi_1) + \mu_n(\Pi_2)$$



For concatenated paths

$$\mu_n(\Pi_1 . \Pi_2) = \sum_{i=0}^{n} \binom{n}{i} \mu_i(\Pi_1) \mu_{n-i}(\Pi_2)$$



Deducing an appropriate semi-ring

Using adapted semiring: $(\mathbb{R}, \oplus, \otimes, \overline{0}, \overline{1})$
Should represent :
$w(\Pi) = (\sum_{\pi \in \Pi} P(\pi), \quad \sum_{\pi \in \Pi} P(\pi)|\pi|, \quad \sum_{\pi \in \Pi} P(\pi)|\pi|^2, \dots)$

Example for first two moments:

$$(x_1, y_1, z_1) \oplus (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$
$$(x_1, y_1, z_1) \otimes (x_2, y_2, z_2) = (x_1 x_2, x_1 y_2 + x_2 y_1, x_1 z_2 + 2 y_1 y_2 + x_2 z_1)$$

Deducing an appropriate semi-ring

Using adapted semiring: $(\mathbb{R}, \oplus, \otimes, \overline{0}, \overline{1})$
Should represent :
$w(\Pi) = (\sum_{\pi \in \Pi} P(\pi), \quad \sum_{\pi \in \Pi} P(\pi)|\pi|, \quad \sum_{\pi \in \Pi} P(\pi)|\pi|^2, \ldots)$

Example for first two moments:

$(x_1, y_1, z_1) \oplus (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$
$(x_1, y_1, z_1) \otimes (x_2, y_2, z_2) = (x_1 x_2, x_1 y_2 + x_2 y_1, x_1 z_2 + 2 y_1 y_2 + x_2 z_1)$

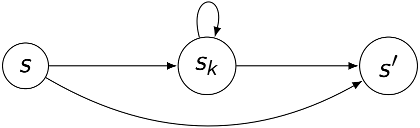
Can be generalized for any number of moments with extended semirings.

Integration over a set of paths

$w(\pi) = \bigotimes_t w(t)$
$w(\Pi) = \bigoplus_{\pi \in \Pi} w(\pi)$

Designing a recursive algorithm based on this information.

Adaptation of Floyd Warshall algorithm

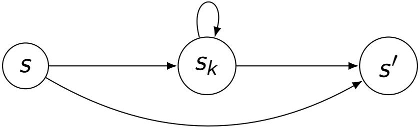


Including all states one by one:

$w(\Pi_k(s, s')) =$

$w(\Pi_{k-1}(s, s')) \oplus w(\Pi_{k-1}(s, s_k)) \otimes w(\Pi_{k-1}(s_k, s_k))^* \otimes w(\Pi_{k-1}(s_k, s'))$

Adaptation of Floyd Warshall algorithm



Including all states one by one:
$w(\Pi_k(s,s')) =$
$w(\Pi_{k-1}(s,s')) \oplus w(\Pi_{k-1}(s,s_k)) \otimes w(\Pi_{k-1}(s_k,s_k))^* \otimes w(\Pi_{k-1}(s_k,s'))$

Theorem [CDC17,FoSSaCS18]

There is a polynomial algorithm that computes the m first moments of the detection time distribution in a diagnoser with $|S|$ states.
Complexity : $O(m^2|S|^3)$

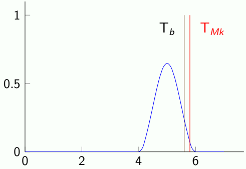
CDC17: Diagnosability Degree of Stochastic Discrete Event Systems, with Eric Fabre and Blaise Genest
FoSSaCS18: Symbolically Quantifying Response Time in Stochastic Models using Moments and Semirings, with Eric Fabre and Blaise Genest

Results on the use of moments

Concentration bounds [FoSSaCS18]

Given any two moments, one can compute optimal concentration bounds.

Ex: $T_b = \mu_1 + \sqrt{\frac{1-\alpha}{\alpha}(\mu_2 - \mu_1^2)}$.

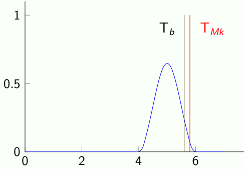


Results on the use of moments

Concentration bounds [FoSSaCS18]

Given any two moments, one can compute optimal concentration bounds.

Ex: $T_b = \mu_1 + \sqrt{\frac{1-\alpha}{\alpha}(\mu_2 - \mu_1^2)}$.



Approximate the time distribution [FoSSaCS18]

- The detection time distribution is totally determined by its moments.
- One can approximate this distribution.

FoSSaCS18: Symbolically Quantifying Response Time in Stochastic Models using Moments and Semirings, with Eric Fabre and Blaise Genest

Plan

Classification

Classification

Being able to retrieve the source of an observation among several choices.

Classification

Classification

Being able to retrieve the source of an observation among several choices.

Let us take a randomly generated sequence:

"Despite the constant negative press covfefe"

Which stochastic system produced it?

Classification

Classification

Being able to retrieve the source of an observation among several choices.

Let us take a randomly generated sequence:

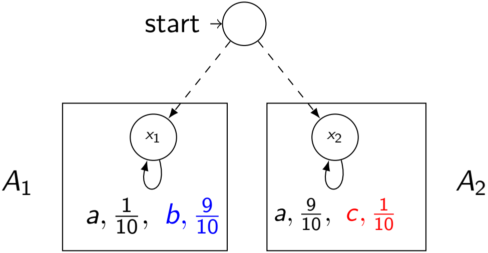
"Despite the constant negative press covfefe"

Which stochastic system produced it?

Classification: more formally

Classification

Given one system chosen at random between A_1, A_2 and an observation w produced by an execution of this system, decide which one was chosen.



- $aaaab \to A_1$.

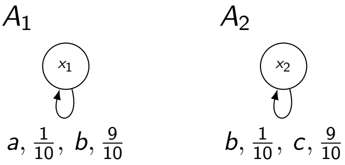
Different classifications

Classifier

Function $f : \Sigma^* \to \{1, 2\}$

Does there exist f that answers correctly...

- For sure: eventually



not sure

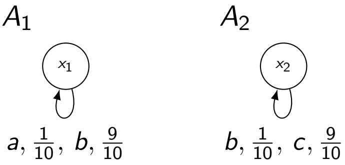
Different classifications

Classifier

Function $f : \Sigma^* \to \{1, 2\}$

Does there exist f that answers correctly...

- For sure: eventually
- Almost sure: eventually with probability 1



almost sure

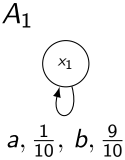
Different classifications

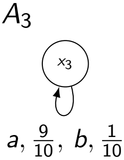
Classifier

Function $f : \Sigma^* \to \{1, 2\}$

Does there exist f that answers correctly...

- For sure: eventually
- Almost sure: eventually with probability 1





not almost sure

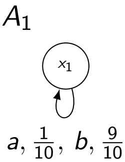
Different classifications

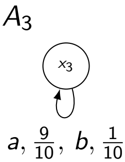
Classifier

Function $f : \Sigma^* \to \{1, 2\}$

Does there exist f that answers correctly...

- For sure: eventually
- Almost sure: eventually with probability 1
- Limit sure: with arbitrarily high confidence





"limit sure"

Sure and almost-sure classification: easy problems

Theorem [YL02]

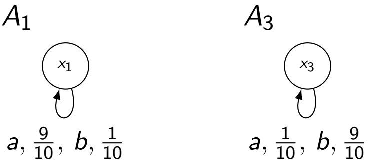
Sure classification is decidable in NLOGSPACE.

Theorem [BHL14]

Almost sure classification is PSPACE-complete.

Limit-sure classification

Two LMCs A_1, A_2 are limit-sure classifiable iff there exists a classifier, f such that $P(\rho \text{ run of } A_1 \text{ of size } k \mid f(obs(\rho)) = 2) \rightarrow_{k\rightarrow\infty} 0$, and similarly for ρ run of A_2.



a limit-sure classifier f: outputs A_1 if the proportion of a is greater than $1/2$, A_2 else.

In general, use Maximum A Posteriori: $MAP(w) = 1 \Leftrightarrow P_1(w) > P_2(w)$.

Plan

Language equivalence

Equivalence between stochastic languages

$A_1 \equiv A_2$ iff for all $w \in \Sigma^*$, $P_1(w) = P_2(w)$.

Equivalence $\Rightarrow$ non-classifiability.

Theorem [Bal93]

Checking equivalence between languages of two LMCs is PTIME.

Similar to [Tze89] for equivalence of PFAs.

Related work: 1) Distinguishability

Monitor [KP16]

Function $Mon : \Sigma^* \to \{\perp, 1\}$ such that if $Mon(u) = 1$ then for all v, $Mon(uv) = 1$.

$L(Mon) = \{w, Mon(w) = 1\} \subseteq \Sigma^\infty$.

Distinguishability for LMCs [KP16]

The LMCs A_1, A_2 are distinguishable if for all $\varepsilon > 0$ there exists a monitor Mon such that $P_{A_1}(L(Mon)) > 1 - \varepsilon$ and $P_{A_2}(L(Mon)) < \varepsilon$.

Equivalent to limit-sure classification for LMCs.

Related work: 1) Distinguishability

Theorem [CK14,KP16]

Distinguishability is PTIME.

Related work: 1) Distinguishability

Theorem [CK14,KP16]

Distinguishability is PTIME.

Total variation metric between two LMCs:

$$TVM(A_1, A_2) = sup_{W \subseteq \Sigma^\infty}(P_1(W) - P_2(W))$$

Theorem [KP16]

Checking distinguishability $\Leftrightarrow$ checking $TVM(A_1, A_2) = 1$.

Related work: 1) Distinguishability

Theorem [CK14,KP16]

Distinguishability is PTIME.

Total variation metric between two LMCs:

$$TVM(A_1, A_2) = sup_{W \subseteq \Sigma^\infty}(P_1(W) - P_2(W))$$

Theorem [KP16]

Checking distinguishability $\Leftrightarrow$ checking $TVM(A_1, A_2) = 1$.

Theorem [CK14]

Checking if $TVM(A_1, A_2) = 1$ is PTIME.

Idea: find two equivalent (reachable) subdistributions.

Related work: 2) initial step opacity

Probabilistic system opacity [KH18]

$\sum_{w \in \Sigma^n} min(P_1(w), P_2(w)) \rightarrow_n 0$?, ie the probability to make an error by using the MAP decreases to 0 with the size of the observation.

Equivalent to limit-sure classification for LMCs.

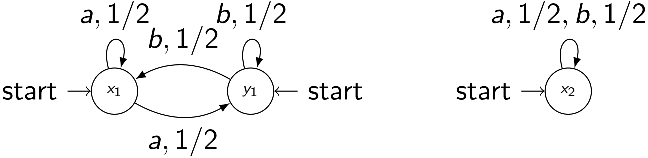
Related work: 2) initial step opacity

Probabilistic system opacity [KH18]

$\sum_{w \in \Sigma^n} min(P_1(w), P_2(w)) \rightarrow_n 0$?, ie the probability to make an error by using the MAP decreases to 0 with the size of the observation.

Equivalent to limit-sure classification for LMCs.

Focus on the stationary distribution of the underlying Markov Chain.



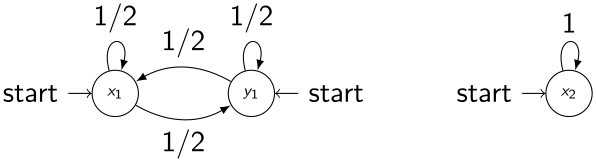
Related work: 2) initial step opacity

Probabilistic system opacity [KH18]

$\sum_{w \in \Sigma^n} min(P_1(w), P_2(w)) \rightarrow_n 0$?, ie the probability to make an error by using the MAP decreases to 0 with the size of the observation.

Equivalent to limit-sure classification for LMCs.

Focus on the stationary distribution of the underlying Markov Chain.



Stationary distribution: $(1/2, 1/2)$ Stationary distribution: (1)

Related work: 2) initial step opacity

Theorem [KH18]

Suppose A_1, A_2 start in their stationary distribution and are ergodic:
A_1 and A_2 are not limit-sure classifiable iff $A_1 \equiv A_2$.

Related work: 2) initial step opacity

Theorem [KH18]

Suppose A_1, A_2 start in their stationary distribution and are ergodic:
A_1 and A_2 are not limit-sure classifiable iff $A_1 \equiv A_2$.
Also, if for all state $s, \sigma_1(s) > 0, \sigma_2(s) > 0$ then

- A_1 and A_2 are not classifiable iff
- A_1 and A_2 are equivalent from stationary distribution.

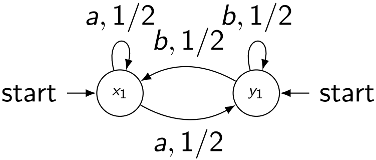
Related work: 2) initial step opacity

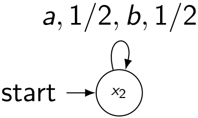
Theorem [KH18]

Suppose A_1, A_2 start in their stationary distribution and are ergodic:
A_1 and A_2 are not limit-sure classifiable iff $A_1 \equiv A_2$.
Also, if for all state $s, \sigma_1(s) > 0, \sigma_2(s) > 0$ then

- A_1 and A_2 are not classifiable iff
- A_1 and A_2 are equivalent from stationary distribution.

A_1



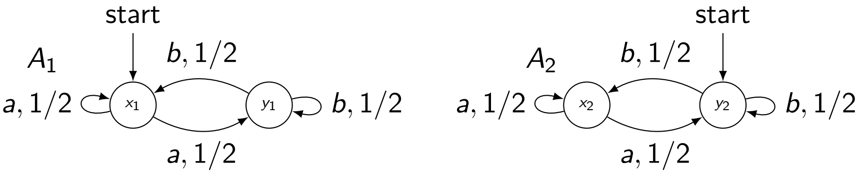


A_2

A_1 from stationary distribution $\equiv A_2$ from stationary distribution.
Hence, cannot limit-sure classify between them.

Related work: 2) initial step opacity

In general, the assumption that all states are initial is crucial.



Stationary distribution: $(1/2, 1/2)$ Stationary distribution: $(1/2, 1/2)$

A_1 from stationary distribution $\equiv$ A_2 from stationary distribution.
But the first letter is enough to classify!

Plan

Our goal

Our goal

- Generalize the idea of [KH18],
- Obtain a general and efficient algorithm and compare with [CK14,KP16].

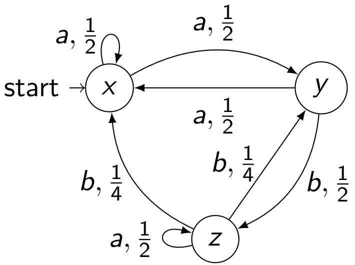
Problem: all states of the LMC are not always reachable from one observation!

Our idea: consider stationary distributions **given the set of states the system can be in after the observation**.

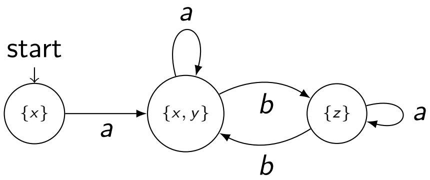
$\mathcal{B}_w$: the possible states after observation w

Consider beliefs $\mathcal{B}_w = \{s \mid s_0 \to^w s\}$ for all observation w.
Ex: $\mathcal{B}_a = \{x, y\}$

We will consider statistics knowing we are in belief $\mathcal{B}$.



A

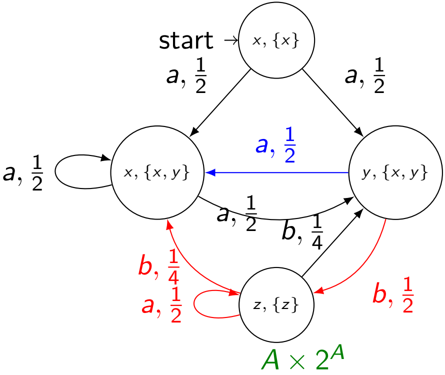


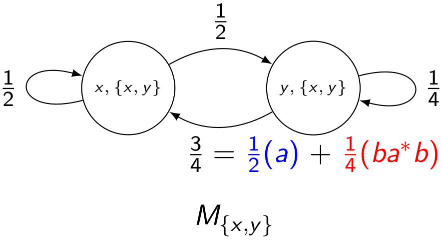
2^A =Observer(A)

Markov chain $M_{\mathcal{B}}$ induced by a belief $\mathcal{B}$

Markov chain $M_{\mathcal{B}}$

$M_{\mathcal{B}}(y,x)$ is the probability in $A \times 2^A$ to reach $(x,\mathcal{B})$ from $(y,\mathcal{B})$ without seeing $(\_,\mathcal{B})$ in-between.

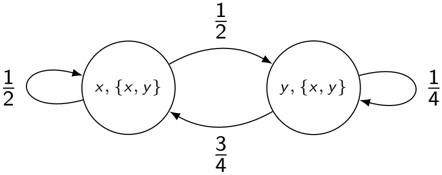




Stationary distribution wrt a belief

Stationary distribution wrt a belief

Let $\sigma_{\mathcal{B}}$ be the stationary distribution of $M_{\mathcal{B}}$.

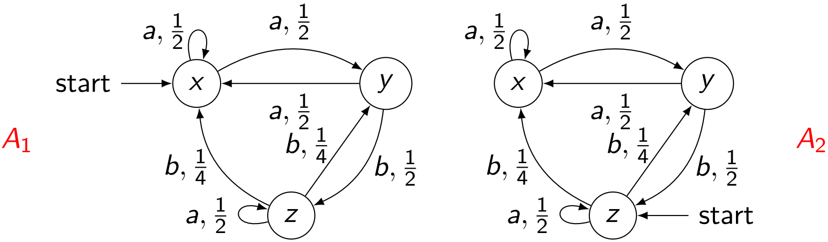


$M_{\{x,y\}}$

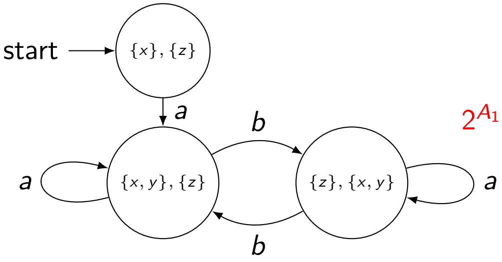
$$\sigma_{\{x,y\}} = (3/5, 2/5).$$

Compared to [KH18], consider the stochastic languages starting from $\sigma_{\{x,y\}}$ instead of σ_{stat}.

Main result

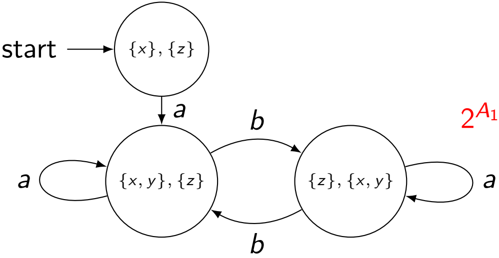

A_1
start
$a, \frac{1}{2}$
$a, \frac{1}{2}$
x
y
$a, \frac{1}{2}$
$b, \frac{1}{4}$
$b, \frac{1}{4}$
$b, \frac{1}{2}$
$a, \frac{1}{2}$
z
A_2
$a, \frac{1}{2}$
$a, \frac{1}{2}$
x
y
$a, \frac{1}{2}$
$b, \frac{1}{4}$
$b, \frac{1}{4}$
$b, \frac{1}{2}$
$a, \frac{1}{2}$
z
start

Main result



$2^{A_1 \times A_2}$

Main result



$2^{A_1 \times A_2}$

Theorem [FSTTCS19]

- One cannot limit-surely classify between A_1, A_2 iff
- There is $\mathcal{B}$ belief of $A_1 \times A_2$ such that $(A_1, \sigma^1_{\mathcal{B}}) \equiv (A_2, \sigma^2_{\mathcal{B}})$.

Problem: exponential number of beliefs.

We use linear programming (similar to [CK14]) to find such a plausible $\mathcal{B}$.

FSTTCS19: Classification among Labeled Markov Chains, with S. Akshay, Eric Fabre and Blaise Genest

Results

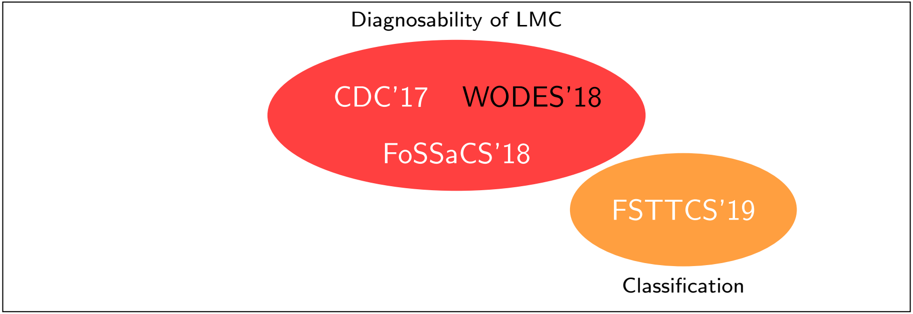
Algorithm [FSTTCS19]

- Polynomial time algorithm to solve limit-sure classification,
- Based on finding a plausible $\mathcal{B}$ with equivalent stochastic languages in A_1 and A_2.

- Idea is an extension of [KH18],
- the method is very different from [CK14],
- but the resulting algorithm is similar to [CK14].
- However, less variables in the Linear Program (search only in BSCCs).
- Stationary distributions on beliefs allow one to solve additional problems, eg classification in a security context.

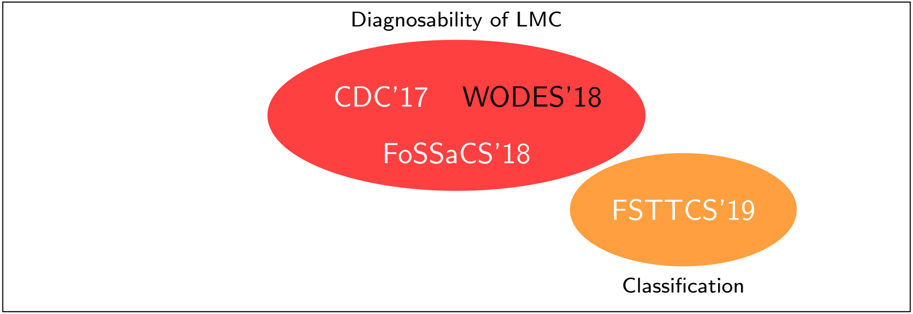
Contributions on diagnosability and classification

Markovian models



Contributions on diagnosability and classification

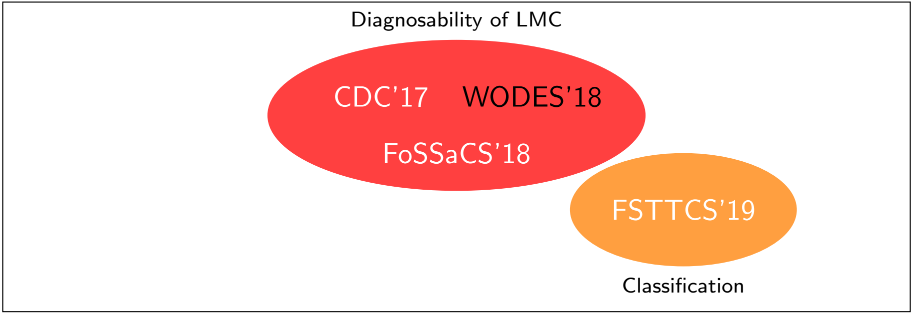
Markovian models



- Analysis of quantified diagnosability
- Algorithm to compute the moments of detection time distribution

Contributions on diagnosability and classification

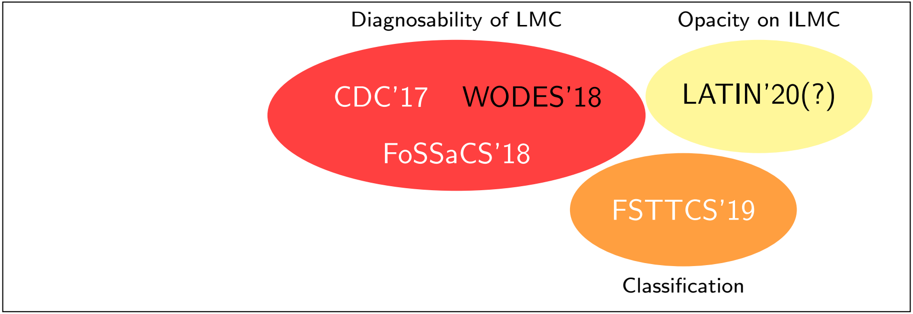
Markovian models



- Analysis of quantified diagnosability
- Algorithm to compute the moments of detection time distribution
- Limit-sure classification: another approach for a PTIME algorithm
- Use a notion of stationary distribution on LMCs

Contributions on diagnosability and classification

Markovian models

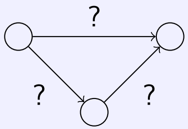


- Analysis of quantified diagnosability
- Algorithm to compute the moments of detection time distribution
- Limit-sure classification: another approach for a PTIME algorithm
- Use a notion of stationary distribution on LMCs
- Opacity of Interval-LMCs

Plan

Learning a model

Obtaining a stochastic model is hard.



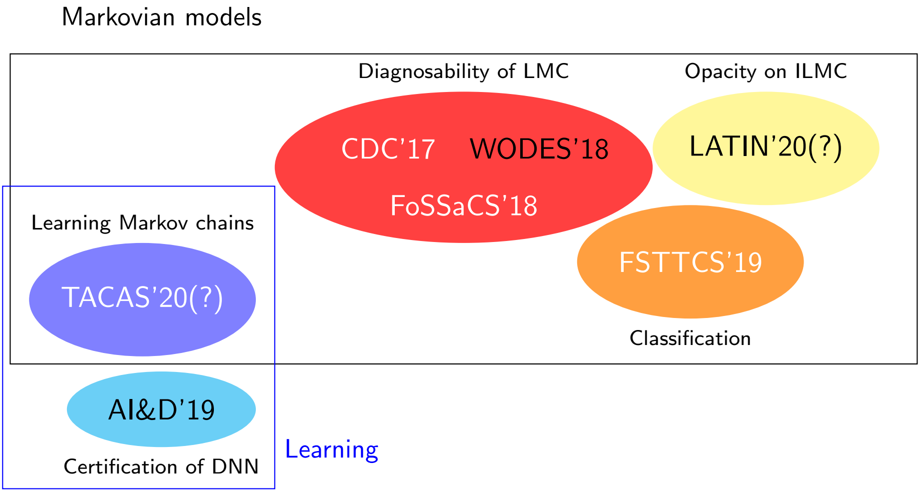
Our goal

- Learn transition probabilities by observing the system,
- Being able to give guarantees on the result,
- Focus on global properties with CTL logic.

TACAS'20(?): Global PAC Bounds for Learning Discrete Time Markov Chains, with Blaise Genest, Cyrille Jegourel and Sun Jun.

Plan

Contributions



Perspectives

What is the runtime on practical models?

- Worst case complexity analysis, and some heuristics (WODES'18)
- Experiments on use-cases?
- More heuristics?

Perspectives

What is the runtime on practical models?

- Worst case complexity analysis, and some heuristics (WODES'18)
- Experiments on use-cases?
- More heuristics?

How to deal with uncertainty on probabilities?

- Internship of K. Garg I co-supervised on Interval-LMCs and opacity (LATIN'20(?)),
- Many problems are harder with uncertain probabilities,
- Same questions as in this talk but with imprecise probabilities?

Perspectives

What is the runtime on practical models?

- Worst case complexity analysis, and some heuristics (WODES'18)
- Experiments on use-cases?
- More heuristics?

How to deal with uncertainty on probabilities?

- Internship of K. Garg I co-supervised on Interval-LMCs and opacity (LATIN'20(?)),
- Many problems are harder with uncertain probabilities,
- Same questions as in this talk but with imprecise probabilities?

Guarantees for learning?

- Use formal methods to obtain guarantees for learning MC (TACAS'20(?)),
- Survey over verification of DNNs (AI&D'19),
- How to more efficiently give different guarantees on different systems?

Thank you!

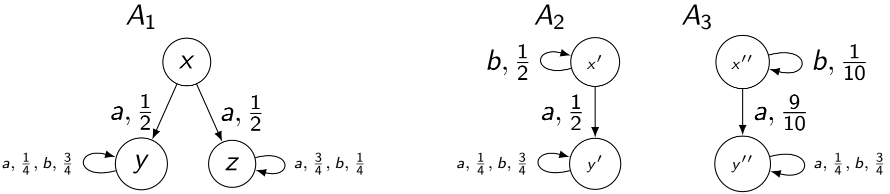
In a security context

Attacker

- Add some power to an attacker, by allowing him to reset the system,
- Verify if there is a strategy for the attacker to be able to decide.

A_1, A_2 is limit-sure (resp. $1-\varepsilon$) attack classifiable iff

1. there is a *reset strategy* $\tau : \Sigma^* \to \{\bot, reset\}$ telling when to reset, and which eventually stops resetting, with probability 1 on the reset runs, and
2. a limit-sure (resp. $1-\varepsilon$) classifier for u, where $u \in \Sigma^*$ denotes the suffix of observations since last reset.



Results on attack-classification

Theorem

Limit-sure attack-classification is PSPACE-complete.

Theorem

$1 - \varepsilon$ attack-classification is undecidable.