# Probabilistic Opacity

Béatrice Bérard

Université P. & M. Curie, LIP6 and Inria, LSV, ENS Cachan

Based on joint work with:
S. Haddad, O. Kouchnarenko, E. Lefaucheux, J. Mullins, M. Sassolas

AFSEC, 21 juin 2017

# Context: Information Flow

Goal: Detect/measure/compare/remove information leaks

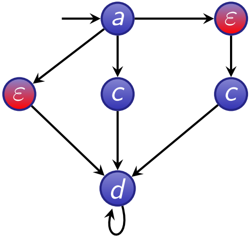
Opacity: In a partially observed transition system, it is achieved when an external observer can never be sure if a secret behaviour has occurred. [Bryans, Koutny, Mazaré, Ryan 2008]

Secret: visiting a red state hidden from observer

observing $ad^*$ dicloses the secret

$acd^*$ is ambiguous

$\mathcal{A}$:




Opacity is used to express a large variety of information flow properties, for instance: anonymity, non interference, conditional declassification.

# Outline

# Opacity framework

## Problems

- A transition system $\mathcal{A}$ with pathes $Path(\mathcal{A})$,
- Some pathes are secret: $Sec \subseteq Path(\mathcal{A})$,
- An external agent knows the system and observes its executions via a function $\mathcal{O}$ on $Path(\mathcal{A})$,

### Qualitative problem

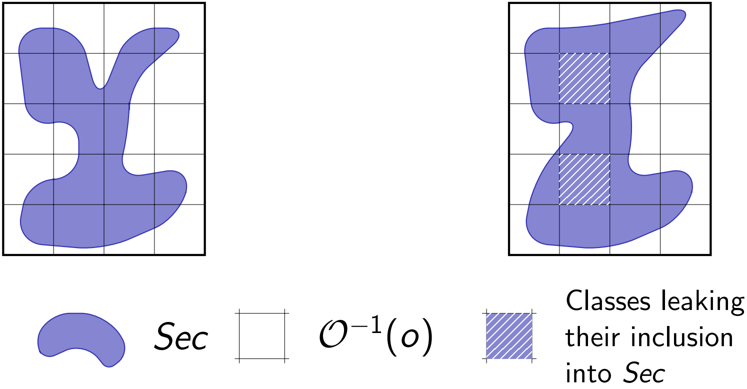
Does there exist a path $\rho$ disclosing the secret: $\mathcal{O}^{-1}(\mathcal{O}(\rho)) \subseteq Sec$ ?
i.e. all pathes with the same observation as $\rho$ are secret.

If no, all secret pathes are ambiguous and the system is **opaque**.

### Quantitative problem

What is the "measure" of disclosing pathes ?

# Illustration




With $\overline{Sec} = Path(\mathcal{A}) \setminus Sec$:
No disclosing path iff
$V = Sec \setminus \mathcal{O}^{-1}(\mathcal{O}(\overline{Sec}))$ is empty

Measuring the disclosure set $V$

# Verification and control of qualitative opacity with regular secrets

## On transition systems

- checking opacity is undecidable in general [BKMR08],
- PSPACE-complete for finite automata [Cassez, Dubreil, Marchand 09], also with opacity variants [Saboori, Hadjicostis 13], and for any functional transducer as observation [B., Mullins 14].
- Enforcement of opacity [Wu, Lafortune 12], [Marchand 11-15, with many co-authors], [Tong, Ma, Li, Seatzu, Giua 16].

## On Petri nets

- undecidable in general [BKMR08][B., Haar, Schmitz, Schwoon 17],
- ESPACE-complete for safe PNs, even when weak-fairness conditions are added. (ESPACE is the class of problems that can be solved in deterministic space $2^{O(n)}$) [BHSS17]

# Strong anonymity

**Actions of participants: $P$**

For any path $\rho \in Path(\mathcal{A})$, replacing an action in $P$ by any other one produces a path still in $Path(\mathcal{A})$.

**Translates as opacity [BKMR08]**

- $\mathcal{O}$ is the morphism into $(\Sigma \cup \{\sharp\})^*$ defined by: $\mathcal{O}(a) = \sharp$ if $a \in P$ and $\mathcal{O}(a) = a$ otherwise
- $\pi_P$ the projection on $P^*$

$\mathcal{A}$ is strongly anonymous w.r.t. $P$ iff for any $u \in P^*$,

$$Sec_u = \{\rho \in Path(\mathcal{A}) \mid \pi_P(\rho) \neq u \wedge |\pi_P(\rho)| = |u|\}$$

is opaque for $\mathcal{A}$ and $\mathcal{O}$.

# Strong anonymity

**Actions of participants: $P$**

For any path $\rho \in Path(\mathcal{A})$, replacing an action in $P$ by any other one produces a path still in $Path(\mathcal{A})$.

**Translates as opacity [BKMR08]**

- $\mathcal{O}$ is the morphism into $(\Sigma \cup \{\sharp\})^*$ defined by:
  $\mathcal{O}(a) = \sharp$ if $a \in P$ and $\mathcal{O}(a) = a$ otherwise
- $\pi_P$ the projection on $P^*$

$\mathcal{A}$ is strongly anonymous w.r.t. $P$ iff for any $u \in P^*$,

$$Sec_u = \{\rho \in Path(\mathcal{A}) \mid \pi_P(\rho) \neq u \wedge |\pi_P(\rho)| = |u|\}$$

is opaque for $\mathcal{A}$ and $\mathcal{O}$.

**But also as another inclusion problem [BM14]**

$\mathcal{O}_P(Path(\mathcal{A})) \subseteq Path(\mathcal{A})$ for the substitution defined by:
$\mathcal{O}_P(a) = P$ if $a \in P$ and $\mathcal{O}_P(a) = \{a\}$ otherwise

# Quantitative aspects

Several sources of uncertainty:

- Partial observation of executions
- Probabilities

$\hookrightarrow$ based on randomness, resolved on the fly by the environment.

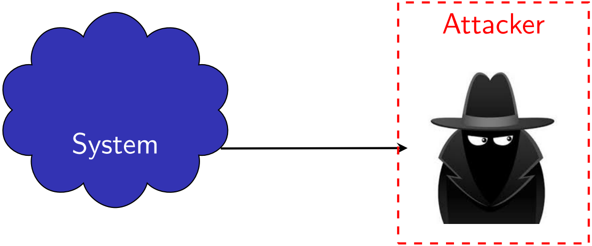
- Nondeterministic choice

$\hookrightarrow$ resolved on the fly by an internal agent.

- Underspecification

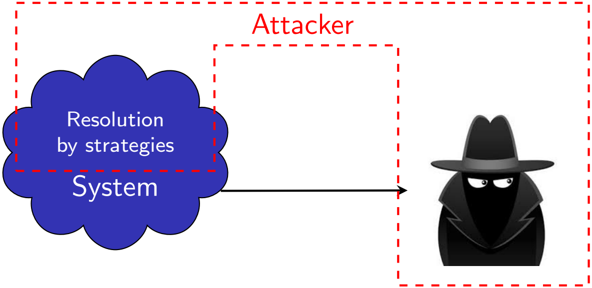
$\hookrightarrow$ resolved later on in the modeling process by refinement.

# Opacity under uncertainty




- Probabilistic choice: Markov Chains
  [B., Mullins, Sassolas 10,15] [Saboori, Hadjicostis 14]

# Opacity under uncertainty

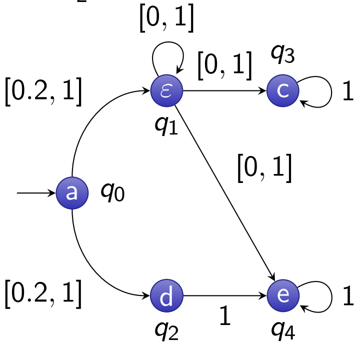



- Probabilistic choice: Markov Chains
  [B., Mullins, Sassolas 10,15] [Saboori, Hadjicostis 14]
- Combined with nondeterministic choice:
  [B., Chatterjee, Sznajder 15] for MDPs and POMDPs,
  [B., Haddad, Lefaucheux 17] for MDPs,
- Underspecification: [B., Kouchnarenko, Mullins, Sassolas 16] for IMCs.

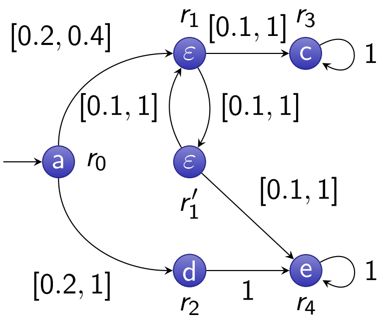
# A toy example

Access control to a database inspired from [Biondi et al. 13]

$\mathcal{M}_2$ :



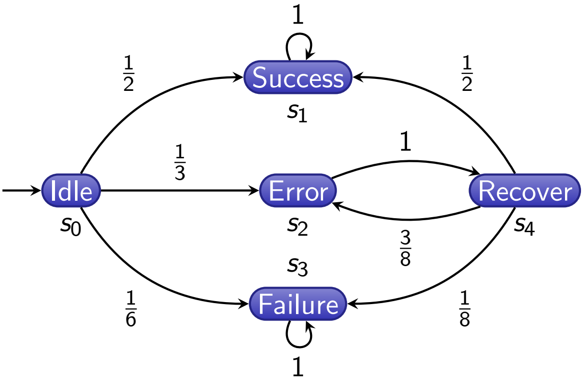

$\mathcal{M}_1$ :




0: input user name, 1: input password, 3: access granted if correct
2: not on the list of authorized users, 4: reject
$Sec = \{0.1.3^\omega\}$; All states except 1 and 1′ are observable.

# Outline

# Observable Markov chains




**A Markov Chain $\mathcal{A} = (S, \Delta, \mathcal{O})$ over $\Sigma$:**

- countable set $S$ of states,
- $\Delta : S \to \mathcal{D}ist(S)$,
- $\mathcal{O} : S \to \Sigma \cup \{\varepsilon\}$ observation function.

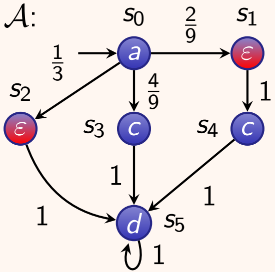
equipped with an initial distribution $\mu_0$.

# Opacity on MCs

**$\omega$-Disclosure of $Sec$ in $(\mathcal{A}, \mu_0)$:**

$Disc_\omega(\mathcal{A}, \mu_0, Sec) = \mathbf{P}_{\mathcal{A},\mu_0}(V)$ for $V = Sec \setminus \mathcal{O}^{-1}(\mathcal{O}(\overline{Sec}))$.

**Example with $Sec$: presence of $s_1$ or $s_2$, hidden by $\mathcal{O}$**




| $Path(\mathcal{A})$ | $\mathcal{O}$ | $Sec$? | $V$? | $\mathbf{P}_{\mathcal{A}}$ |
|---|---|---|---|---|
| $s_0 s_2 s_5^\omega$ | $ad^\omega$ | ✓ | ✓ | $1/3$ |
| $s_0 s_3 s_5^\omega$ | $acd^\omega$ | ✗ | ✗ | $4/9$ |
| $s_0 s_1 s_4 s_5^\omega$ | $acd^\omega$ | ✓ | ✗ | $2/9$ |

$$Disc_\omega(\mathcal{A}, \mathbf{1}_{s_0}, Sec) = \frac{1}{3}$$

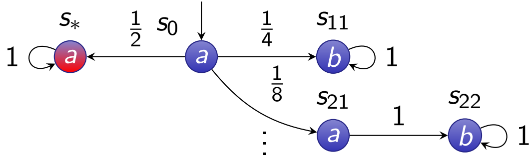
# Finite disclosure

**Restricting *Sec* to the set of pathes visiting states from a given subset**

assuming a path remains secret once a secret state has been visited.

Observation sequence $w$ in $\Sigma^*$ is:
**disclosing** if all pathes in $\mathcal{O}^{-1}(w)$ are secret,
**minimal disclosing** if disclosing with no strict disclosing prefix.

- $Disc(\mathcal{A}, \mu_0, Sec)$: probability of minimal disclosing observations,
- $Disc_n(\mathcal{A}, \mu_0, Sec)$: probability of disclosing observations of length $n$.



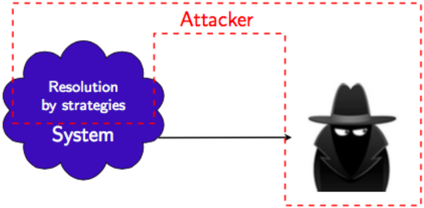

$Disc_\omega = \frac{1}{2}$

$Disc = Disc_n = 0$

$Disc \leq Disc_\omega$, equality if $\mathcal{A}$ is convergent and finitely branching.

# Outline

# Interactions with the system




## Active attacker

The attacker consists of two components:

- The passive external observer,
- Some piece of code inside the system.

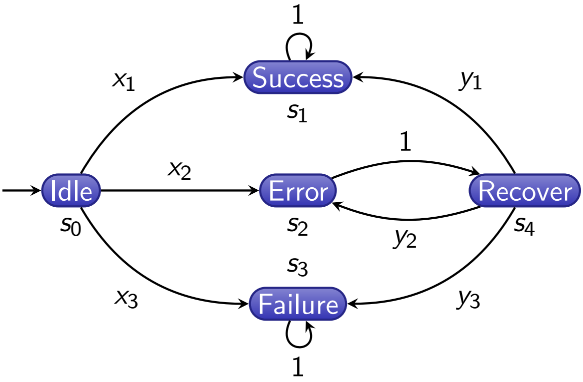
Worst case corresponds to maximal disclosure.

## System designer

The designer has provided a first version with the required functionalities.
He must develop the access policy...

... to obtain minimal disclosure.

# Constraint Markov Chains




$\mathcal{M}_1 = (S, T_1, \mathcal{O})$ :

$T_1(s_0)$ subset of:
$0 \le x_1, x_2, x_3 \le 1$
$x_1 + x_2 + x_3 = 1$

$T_1(s_4)$ subset of:
$0 \le y_1, y_2, y_3 \le 1$
$y_1 + y_2 + y_3 = 1$

**A CMC over $\Sigma$:** [Jonsson, Larsen 1991] [Caillaud et al., 2011]

$\mathcal{M} = (S, T, \mathcal{O})$ is like an OMC with

- finite set of states $S$,
- $T : S \to 2^{\mathcal{D}ist(S)}$.

# Subclasses of CMCs

**MDP: Markov Decision Processes**

For each $s \in S$, $T(s)$ is a finite set.

**LCMC: Linear CMCs**

For each $s \in S$, $T(s)$ is the set of distributions that are solutions of a linear system.

**IMC: Interval MC**

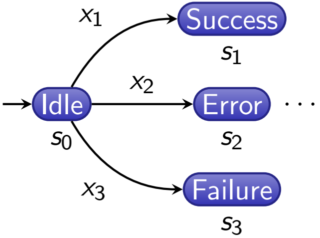
For each $s$, $T(s)$ is described by a family of intervals $(I(s, s'))_{s' \in S}$.

**Relations**

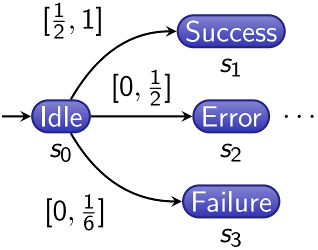
- IMC is a strict subclass of LCMC,
- Any LCMC can be transformed in an exponentially larger MDP.

# Examples

LCMC $\mathcal{M}_2$:




IMC $\mathcal{M}_3$:




$$0 \le x_1, x_2, x_3 \le 1$$
$$x_1 + x_2 + x_3 = 1$$

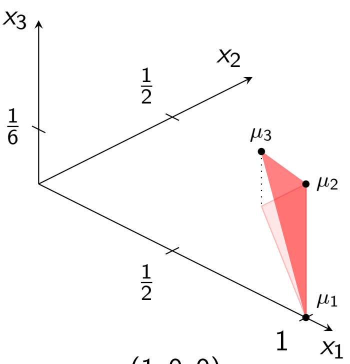
$$x_2 \ge 2x_3$$
$$x_2 + x_3 \le \frac{1}{2}$$

$$\frac{1}{2} \le x_1 \le 1$$
$$0 \le x_2 \le \frac{1}{2}$$
$$0 \le x_3 \le \frac{1}{6}$$

$$\mu_1 = (1, 0, 0)$$
$$\mu_2 = (\tfrac{1}{2}, \tfrac{1}{2}, 0)$$
$$\mu_3 = (\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{6})$$

$$\mu_4 = (\tfrac{5}{6}, 0, \tfrac{1}{6}) \in T_3(s_0)$$
$$\mu_4 \notin T_2(s_0)$$

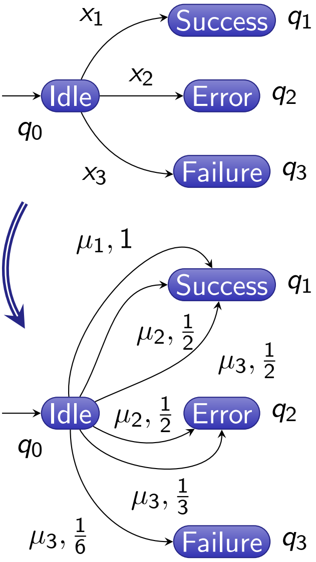
# From LCMCs to MDPs



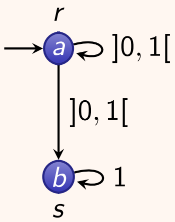

$$\mu_1 = (1, 0, 0)$$

$$\mu_2 = (\tfrac{1}{2}, \tfrac{1}{2}, 0)$$

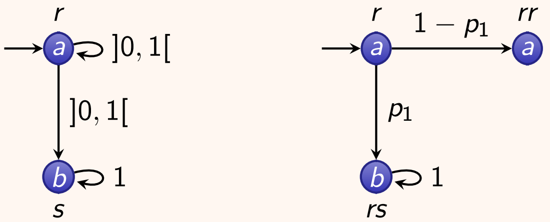
$$\mu_3 = (\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{6})$$



# Strategies on CMCs



# Strategies on CMCs

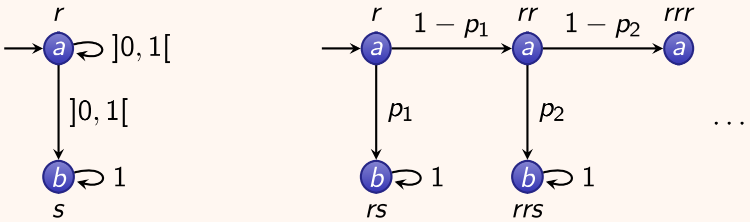



**A strategy for $\mathcal{M} = (S, T, \mathcal{O})$ with initial distribution $\mu_0$:**

$\sigma : FRuns(\mathcal{M}) \to \mathcal{D}ist(S)$
For $\rho = s_0 \xrightarrow{\mu_1} s_1 \ldots \xrightarrow{\mu_n} s_n,\ \ \sigma(\rho) \in T(s_n)$.

Scheduling $\mathcal{M}$ with $\sigma$ produces a (possibly infinite) MC $\mathcal{M}_\sigma$.

# Strategies on CMCs




**A strategy for $\mathcal{M} = (S, T, \mathcal{O})$ with initial distribution $\mu_0$:**

$\sigma : FRuns(\mathcal{M}) \to \mathcal{D}ist(S)$

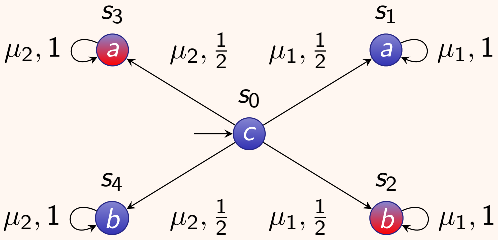
For $\rho = s_0 \xrightarrow{\mu_1} s_1 \ldots \xrightarrow{\mu_n} s_n$, $\sigma(\rho) \in T(s_n)$.

Scheduling $\mathcal{M}$ with $\sigma$ produces a (possibly infinite) MC $\mathcal{M}_\sigma$.

# Randomized strategies on MDPs

**An MDP with distributions $\mu_1$ and $\mu_2$ for $s_0$ and secret states $\{s_2, s_3\}$**

$Disc = \frac{1}{2}$ with the two strategies choosing $\mu_1$ or $\mu_2$ in $s_0$
if they are known by the observer.




But $Disc = 0$ with randomized strategies $\sigma_p$ such that
$\sigma_p(s_0) = p\mu_1 + (1-p)\mu_2$ with $0 < p < 1$. Necessary for minimisation.

**A randomized strategy associates $\sigma(\rho) \in \mathcal{D}ist(T(s_n))$**

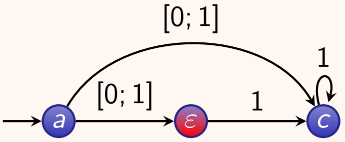
with $\rho = s_0 \xrightarrow{\mu_1} s_1 \ldots \xrightarrow{\mu_n} s_n$ (instead of $\sigma(\rho)$ in $T(s_n)$).

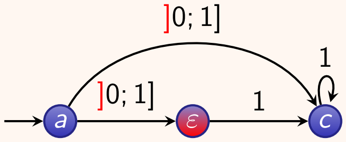
# Modal edges

**An edge $(s, s')$ is modal**

if a strategy can block it completely.

**Example on an IMC with *Sec* : presence of red, hidden by $\mathcal{O}$.**



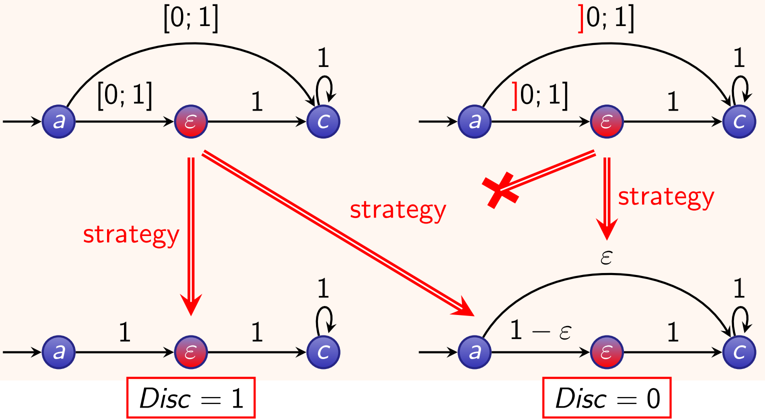



# Modal edges

**An edge $(s, s')$ is modal**

if a strategy can block it completely.

**Example on an IMC with *Sec* : presence of red, hidden by $\mathcal{O}$.**



# Maximal and minimal disclosure

For *Sec* in $\mathcal{M}$ with initial distribution $\mu_0$:

- $Disc_{\mathsf{max}}(\mathcal{M}, \mu_0, Sec) = sup_{\sigma \in Strat(\mathcal{M})} Disc(\mathcal{M}_\sigma, \mu_0, Sec)$
- $Disc_{\mathsf{min}}(\mathcal{M}, \mu_0, Sec) = inf_{\sigma \in Strat(\mathcal{M})} Disc(\mathcal{M}_\sigma, \mu_0, Sec)$

Several disclosure problems for a given $\mathcal{M}$

- Value problem: compute the disclosure $Disc_{\mathsf{max}}$ or $Disc_{\mathsf{min}}$.
- Quantitative decision problems: Given a threshold $\theta \in [0, 1]$, is $Disc_{\mathsf{max}} \geq \theta$ ? is $Disc_{\mathsf{min}} \leq \theta$ ?
- Qualitative decision problems:
  Limit-sure disclosure: the quantitative problem
  with $\theta = 1$ for maximisation and $\theta = 0$ for minimisation.

# Maximal Disclosure

[BCS15] On MDPs, if observer ignores the strategies:

- The value can be computed in polynomial time;
- All problems are decidable.

[BKMS16]: For a non modal LCMC, the value can be computed in EXPTIME.

[BHL17] On MDPs, if observer knows the strategies:

- Deterministic strategies are sufficient;
- The problem asking whether there exists a strategy producing value 1 is EXPTIME-complete;
- But the quantitative and limit-sure problems are undecidable.

Consequence:

The quantitative problem is undecidable for general LCMCs.

# Minimal Disclosure

[BHL17] On MDPs, if observer knows the strategies:

- Families of randomized strategies are necessary;
- The value can be computed in EXPTIME;
- All problems are decidable.

# Outline

A brief overview on opacity

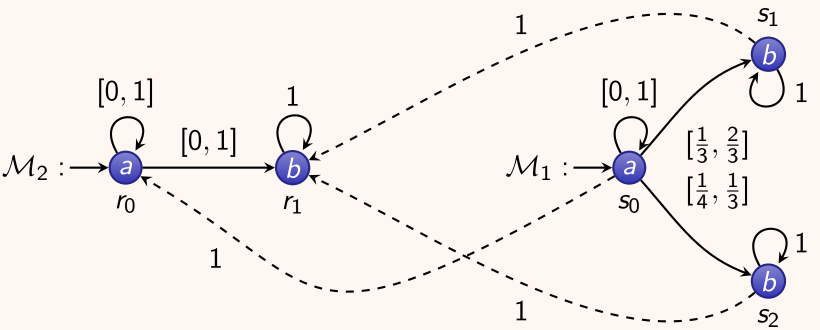
Probabilistic disclosure for Markov Chains

Disclosing a secret under strategies

**Opacity and refinement**

# Refinement for CMCs

Refinement of $\mathcal{M}_2$ by $\mathcal{M}_1$:




## Strong refinement [Jonsson, Larsen, 1991]

is a relation $\mathcal{R} \subseteq S_1 \times S_2$ compatible with labeling, containing $(s_{1,init}, s_{2,init})$ and if $s_1 \mathcal{R} s_2$ there is a mapping $\delta : S_1 \to \mathcal{D}ist(S_2)$ such that:

- all distributions in $T_1(s_1)$ translate to $S_2$ in a way compatible with $T_2(s_2)$
- if $\delta(s_1')(s_2') > 0$ then $s_1' \mathcal{R} s_2'$.

# Monotonicity of maximal disclosure

No inclusion between $\underline{sat}(\mathcal{M}_1) = \{\mathcal{M}_{1,\sigma_1} \mid \sigma_1 \in Strat(\mathcal{M}_1)\}$ and $\underline{sat}(\mathcal{M}_2) = \{\mathcal{M}_{2,\sigma_2} \mid \sigma_2 \in Strat(\mathcal{M}_2)\}$.

**Disclosure is monotonic for LCMCs:**

If $\mathcal{M}_1$ weakly refines $\mathcal{M}_2$ with initial states $s_{1,init}$ and $s_{2,init}$ then for a secret $Sec$, $Disc_{\max}(\mathcal{M}_1, \mathbf{1}_{s_{1,init}}, Sec) \leq Disc_{\max}(\mathcal{M}_2, \mathbf{1}_{s_{2,init}}, Sec)$.

**Construction of the relation**

$$\begin{tikzcd}
\mathcal{M}_2 & \mathcal{M}_1 \arrow[l, "\mathcal{R}"'] \\
 & \mathcal{M}_{1,\sigma_1} \arrow[u, "sat_1"']
\end{tikzcd}$$

# Monotonicity of maximal disclosure

No inclusion between $\underline{sat}(\mathcal{M}_1) = \{\mathcal{M}_{1,\sigma_1} \mid \sigma_1 \in \mathit{Strat}(\mathcal{M}_1)\}$ and $\underline{sat}(\mathcal{M}_2) = \{\mathcal{M}_{2,\sigma_2} \mid \sigma_2 \in \mathit{Strat}(\mathcal{M}_2)\}$.

**Disclosure is monotonic for LCMCs:**

If $\mathcal{M}_1$ weakly refines $\mathcal{M}_2$ with initial states $s_{1,init}$ and $s_{2,init}$ then for a secret $\mathit{Sec}$, $\mathit{Disc}_{\max}(\mathcal{M}_1, \mathbf{1}_{s_{1,init}}, \mathit{Sec}) \leq \mathit{Disc}_{\max}(\mathcal{M}_2, \mathbf{1}_{s_{2,init}}, \mathit{Sec})$.

**Construction of the relation**

$$\begin{tikzcd}
\mathcal{M}_2 & \mathcal{M}_1 \arrow[l, "\mathcal{R}"'] \\
 & \mathcal{M}_{1,\sigma_1} \arrow[u, "sat_1"'] \arrow[ul, dotted, red, "\mathcal{R} \circ sat_1"']
\end{tikzcd}$$

# Monotonicity of maximal disclosure

No inclusion between $\underline{sat}(\mathcal{M}_1) = \{\mathcal{M}_{1,\sigma_1} \mid \sigma_1 \in Strat(\mathcal{M}_1)\}$ and $\underline{sat}(\mathcal{M}_2) = \{\mathcal{M}_{2,\sigma_2} \mid \sigma_2 \in Strat(\mathcal{M}_2)\}$.

**Disclosure is monotonic for LCMCs:**

If $\mathcal{M}_1$ weakly refines $\mathcal{M}_2$ with initial states $s_{1,init}$ and $s_{2,init}$ then for a secret $Sec$, $Disc_{\max}(\mathcal{M}_1, \mathbf{1}_{s_{1,init}}, Sec) \leq Disc_{\max}(\mathcal{M}_2, \mathbf{1}_{s_{2,init}}, Sec)$.

**Construction of the relation**

$$\begin{tikzcd}
\mathcal{M}_2 & \mathcal{M}_1 \arrow[l, "\mathcal{R}"'] \\
\mathcal{M}_{2,\sigma_2} \arrow[u, dashed, "sat_2"] & \mathcal{M}_{1,\sigma_1} \arrow[u, "sat_1"'] \arrow[ul, dotted, "\mathcal{R} \circ sat_1"'] \arrow[l, dashed, "\mathcal{R}'"]
\end{tikzcd}$$

# Monotonicity of maximal disclosure

No inclusion between $\underline{sat}(\mathcal{M}_1) = \{\mathcal{M}_{1,\sigma_1} \mid \sigma_1 \in Strat(\mathcal{M}_1)\}$ and $\underline{sat}(\mathcal{M}_2) = \{\mathcal{M}_{2,\sigma_2} \mid \sigma_2 \in Strat(\mathcal{M}_2)\}$.

## Disclosure is monotonic for LCMCs:

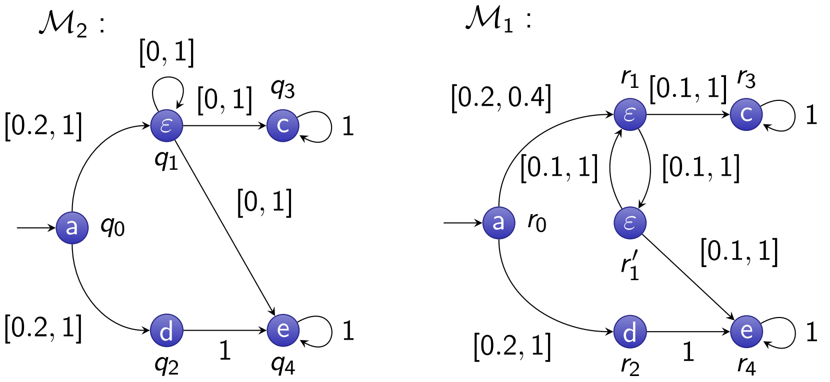
If $\mathcal{M}_1$ weakly refines $\mathcal{M}_2$ with initial states $s_{1,init}$ and $s_{2,init}$ then for a secret $Sec$, $Disc_{\max}(\mathcal{M}_1, \mathbf{1}_{s_{1,init}}, Sec) \leq Disc_{\max}(\mathcal{M}_2, \mathbf{1}_{s_{2,init}}, Sec)$.

## Construction of the relation

$$\begin{tikzcd}
\mathcal{M}_2 & & \mathcal{M}_1 \arrow[ll, "\mathcal{R}"'] \\
\mathcal{M}_{2,\sigma_2} \arrow[u, dashed, "sat_2"] & & \mathcal{M}_{1,\sigma_1} \arrow[u, "sat_1"'] \arrow[ull, dotted, "\mathcal{R} \circ sat_1"'] \arrow[ll, dashed, "\mathcal{R}'"]
\end{tikzcd}$$

If $\mathcal{M}_1$ weakly refines $\mathcal{M}_2$ then for any strategy $\sigma_1$ of $\mathcal{M}_1$, there is a strategy $\sigma_2$ of $\mathcal{M}_2$ such that $\mathcal{M}_{1,\sigma_1}$ refines $\mathcal{M}_{2,\sigma_2}$.

# Example




$\mathcal{M}_2$ is refined by $\mathcal{M}_1$,

$$Disc_{\max}(\mathcal{M}_2, \mathbf{1}_{r_0}, Sec) = 0.8 \text{ and } Disc_{\max}(\mathcal{M}_1, \mathbf{1}_{q_0}, Sec) = 0.$$

# A consequence for modeling

IMCs are not closed under conjunction but:

**The conjunction of two IMCs $\mathcal{M}_1$ and $\mathcal{M}_2$ is an LCMC**

Using results from [Caillaud et al, 2011]:

**For LCMCs $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$**

- $\mathcal{M}_1 \wedge \mathcal{M}_2$ weakly refines both $\mathcal{M}_1$ and $\mathcal{M}_2$, hence:

$$Disc_{\mathsf{max}}(\mathcal{M}_1 \wedge \mathcal{M}_2) \leq min(Disc_{\mathsf{max}}(\mathcal{M}_1), Disc_{\mathsf{max}}(\mathcal{M}_2)).$$

- If $\mathcal{M}_3$ refines both $\mathcal{M}_1$ and $\mathcal{M}_2$ then it also weakly refines $\mathcal{M}_1 \wedge \mathcal{M}_2$, hence:

$$Disc_{\mathsf{max}}(\mathcal{M}_3) \leq Disc_{\mathsf{max}}(\mathcal{M}_1 \wedge \mathcal{M}_2).$$

# Conclusion

Opacity is a flexible way to express information flow properties

not necessarily preserved under arbitrary refinement.

Linear CMCs form a good class for compact specifications of probabilistic systems with:

- nice closure properties;
- an increased security criterion with schedulers as adversaries;
- monotonicity of maximal disclosure;
- But the quantitative problem is undecidable in general, like for MDPs, unless the structure is fixed.

Minimisation on MDPs

- require randomized strategies;
- and all quantitative problems are decidable.

# Conclusion

Opacity is a flexible way to express information flow properties

not necessarily preserved under arbitrary refinement.

Linear CMCs form a good class for compact specifications of probabilistic systems with:

- nice closure properties;
- an increased security criterion with schedulers as adversaries;
- monotonicity of maximal disclosure;
- But the quantitative problem is undecidable in general, like for MDPs, unless the structure is fixed.

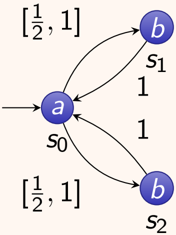
Minimisation on MDPs

- require randomized strategies;
- and all quantitative problems are decidable.

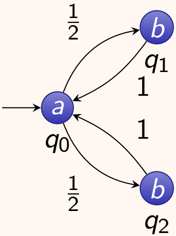
**Thank you**

# Strict inclusion of $\underline{sat}(\mathcal{M})$ in $sat(\mathcal{M})$

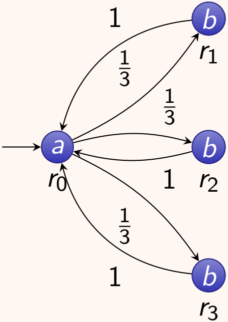
## An implementation not obtained by strategies




Specification $\mathcal{M}$




$\mathcal{A}_0$ with single strategy




$\mathcal{A}_1$ implementation of $\mathcal{M}$