

Decidability of Parameterized Probabilistic Information Flow

Danièle Beauquier¹, Marie Duflot¹ and
Yury Lifshits^{2,3}

¹Université Paris 12

²Steklov Institute of Mathematics at St.Petersburg

³California Institute of Technology

CSR 2007

- Assume we have a **system**
- And somebody observes a **part of its behavior**
- We fix some **property** of the system

Can the observer recover that property?

Our result: there is an algorithm answering the above question given any system and property

Outline

- 1 Information Flow: Definitions
 - System
 - Observation
 - System Properties
 - Definition of Information Flow
- 2 Decidability Results

Part I

What is a system?

What is a partial observation of its behavior?

What is a property of the system?

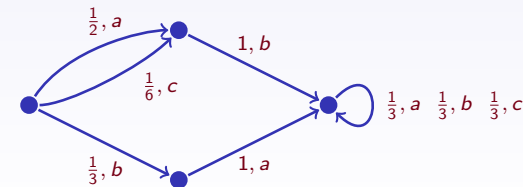
When does a system have **information flow**?

System is a Probability Distribution

- A system is a probability distribution over traces
- A trace is a finite or infinite sequence of alphabet characters
- Today: $\Sigma = L \cup H$ (low-level and high-level events)
- The distribution is described by **Finite Markov Automaton**

Finite Markov Automaton

- Finite number of states
- Edges are labelled by alphabet characters
- Every edge has a probability
- For every edge the sum of probabilities over all outgoing edges is equal to 1



Observation model

For any trace α observation is a projection to low-level events $\alpha|_L$

Projection is just deleting all characters from H from the sequence

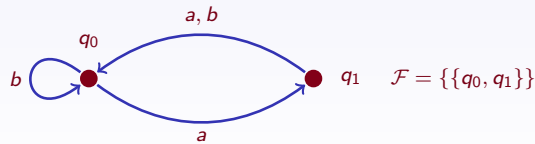
Defining a System Property

We describe any property on traces by **recognizing automaton**: property holds \Leftrightarrow automaton accepts

Today we restrict ourselves to properties recognized by Muller automaton

Muller Automaton

- Finite number of states
- Initial state, family \mathcal{F} of “accepting” sets of states
- Every edge is labelled by alphabet character
- The automaton is complete and deterministic: for every pair (v, a) there exist a unique outgoing edge from the vertex v with that label a
- Muller automaton accepts trace if during “reading” it the set of states visited infinitely many times belongs to \mathcal{F}



Property-Specific Information Flow

$\mathcal{P}_S(P)$ denotes the probability measure of the set of all traces from S satisfying P

The conditional probability $\mathcal{P}_S(P|u)$ denotes the probability measure of the set of all traces which S satisfy P and whose projection to L is starting from u

System S has **no information flow for property P** if

$$\forall u \quad \mathcal{P}_S(P|u) = \mathcal{P}_S(P)$$

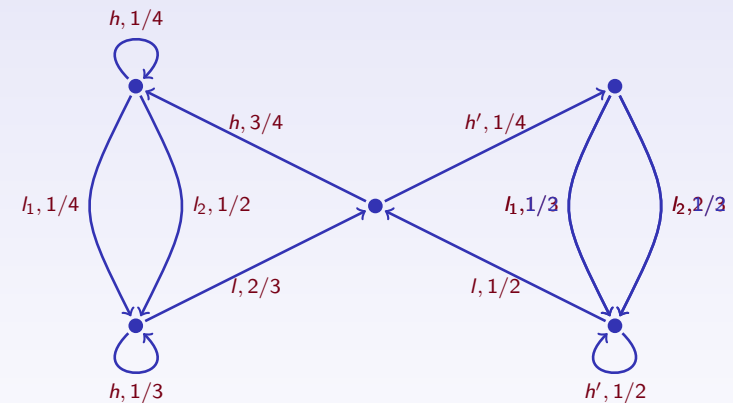
General Information Flow

A system is **without information flow** iff it has no flow for every (defined by Muller automaton) property P

We call property to be **sequential** iff it's Muller automaton treats every low-level event in precisely the same way

A system is **without sequential information flow** iff it has no flow for every sequential property P

Information Flow: Example



The Markov chain above **has no** sequential information flow
Now the Markov chain above **has** sequential information flow

Part II

Given system/property can we determine the existence of information flow?

Deciding Property-Specific Information Flow

Theorem

There is an algorithm deciding property-specific information flow for every pair of system/property (i.e. for pair of Muller automaton and Markov chain)

Reduction to linear algebra:

- 1 Compute a composition of Markov automaton and Büchi automaton
- 2 Simplify it by the rule " $H^*I \rightarrow I$ "

Algorithm Inside (1/2)

We reduce property-specific information flow to the following mathematical problem:

Input: vectors a, c , matrices M_1, \dots, M_n

Question: does there exist a finite sequence of indices such that $aM_{i_1} \dots M_{i_k} c \neq 0$?

Algorithm Inside (2/2)

For every k we will compute basis for linear hull of $V_k = \{aM_{i_1} \dots M_{i_k}\} \cup V_{k-1}$

- 1 a is a basis for V_0
- 2 In order to get basis for V_{k+1} from V_k we multiply all basis vectors by all matrices and keep the maximal linearly independent subset
- 3 **Stopping condition:** $\dim(V_{k+1}) = \dim(V_k)$
- 4 Check whether $V_k \perp c$

Deciding General Information Flow

Theorem

There is an algorithm deciding general information flow for every system described by a Markov chain

Theorem

There is an algorithm deciding sequential information flow for every system described by a Markov chain

Highlights

- System is a Markov probability distribution over traces
- Property is described by Muller automaton
- We can determine the existence of information flow by linear algebra tricks

Future work

- More general models for systems and properties
- Quantitative measure for information flow

Thanks for your attention! Questions?

Danièle Beauquier <http://www.univ-paris12.fr/lacl/beauquier/>

Marie DufLOT <http://www.univ-paris12.fr/lacl/duflot/>

Yury Lifshits <http://yury.name>

Some related work:



D. Beauquier, M. DufLOT, Y. Lifshits
Decidability of Parameterized Probabilistic Information Flow. CSR'07.
<http://yury.name/papers/beauquier2007decidability.pdf>



D. Beauquier, M. DufLOT, M. Minea
A Probabilistic Property-Specific Approach to Information Flow. MMM-ACNS'05.
<http://www.univ-paris12.fr/lacl/Rapports/publications/TR-2005-02.pdf>



A. Slissenko
Complexity problems in the analysis of information systems security. MMM=ACNS'03.
<http://www.springerlink.com/index/WKDENGHGBAFE28KNC.pdf>