Belief bisimulation for hidden Markov models

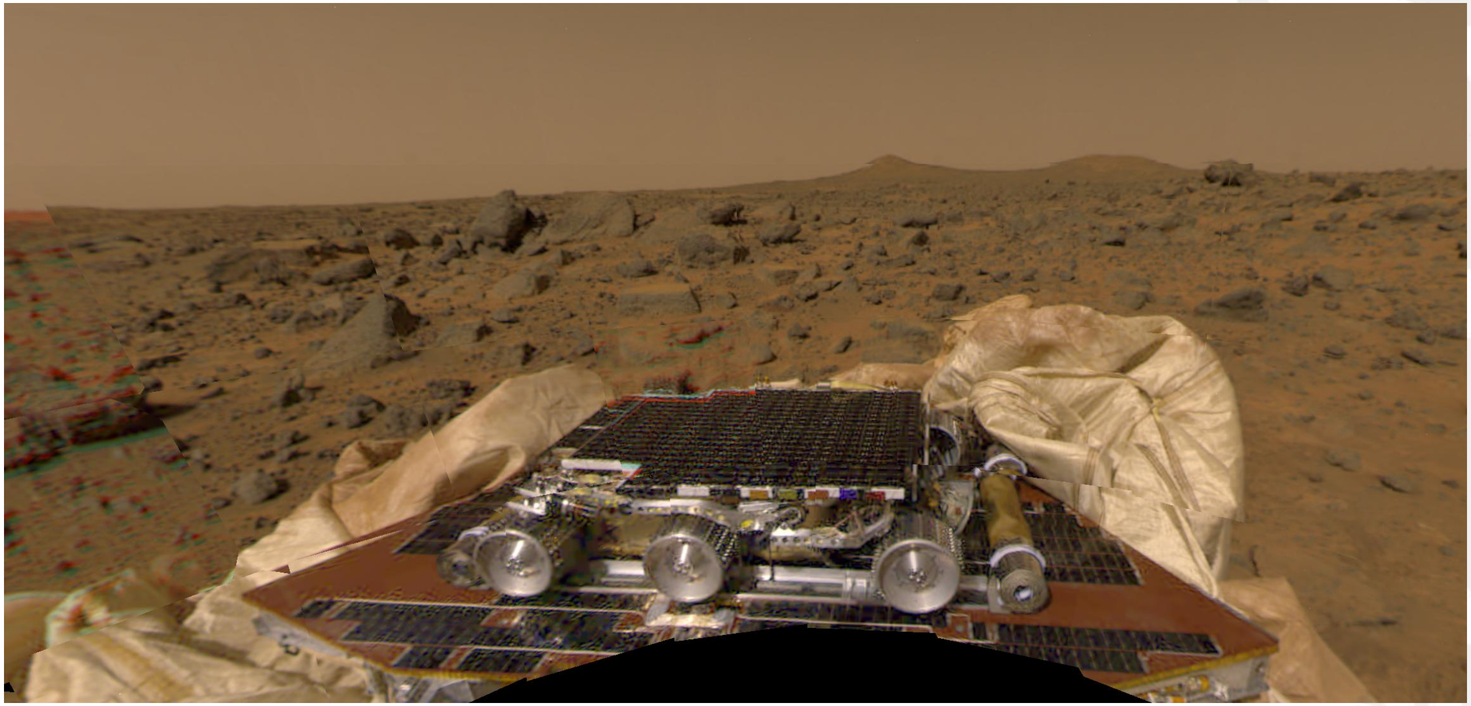
David N. Jansen
Flemming Nielson
Lijun Zhang

Radboud University Nijmegen, The Netherlands
Danish Technical University, Denmark

Holy Wednesday 2012

NASA's current Mars rover...

- How does a Mars rover find its position and orientation?
- Current: camera image sent to Earth, judged by humans

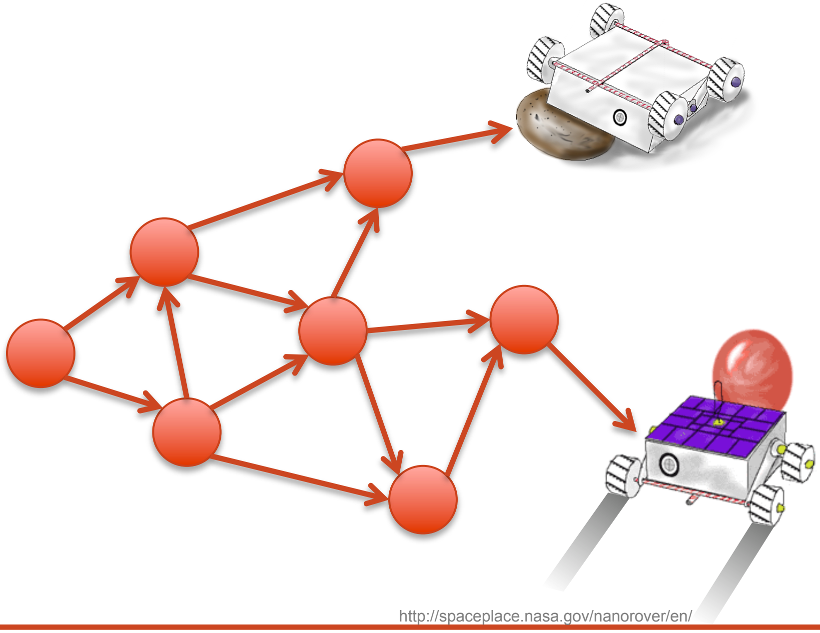


http://mars.jpl.nasa.gov/MPF/ops/sol2.html

NASA’s next Mars rover...

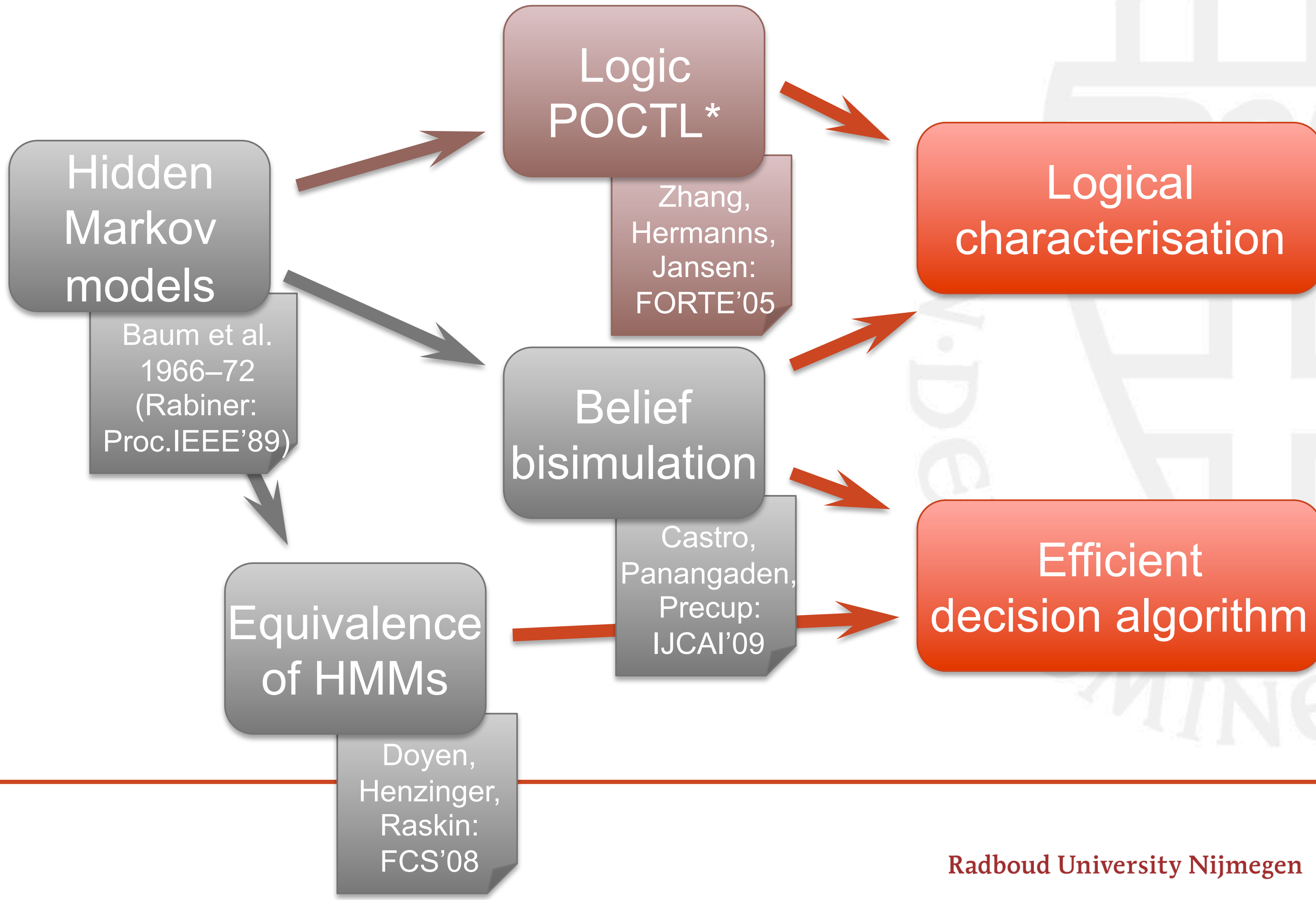
- How does a Mars rover find its position and orientation?
- Current: camera image sent to Earth, judged by humans
- Future possibility: use sensors and infer state from measurements

- **Given a sequence of measurements, what is the current state?**

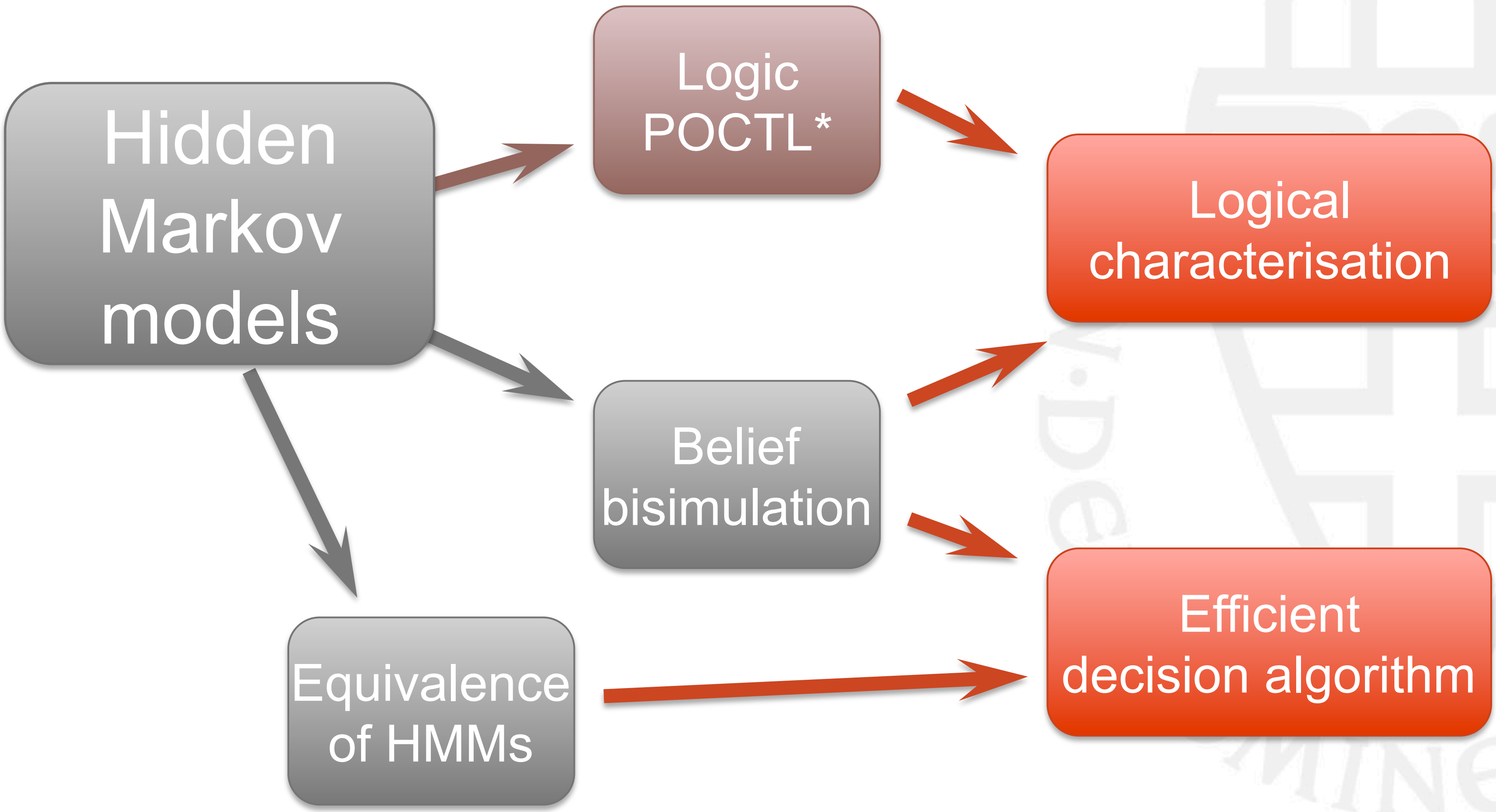


http://spaceplace.nasa.gov/nanorover/en/

Context and Overview



Context and Overview



What is a hidden Markov model?

- describes behaviour of a probabilistic process

- current state is hidden
- observations provide partial information on current state

- Definition:
  - set of states S (labelled with atomic propositions)
  - set of observations Ω
  - probabilistic transition relation $P: S \times S \rightarrow [0,1]$ (labelled with observations)

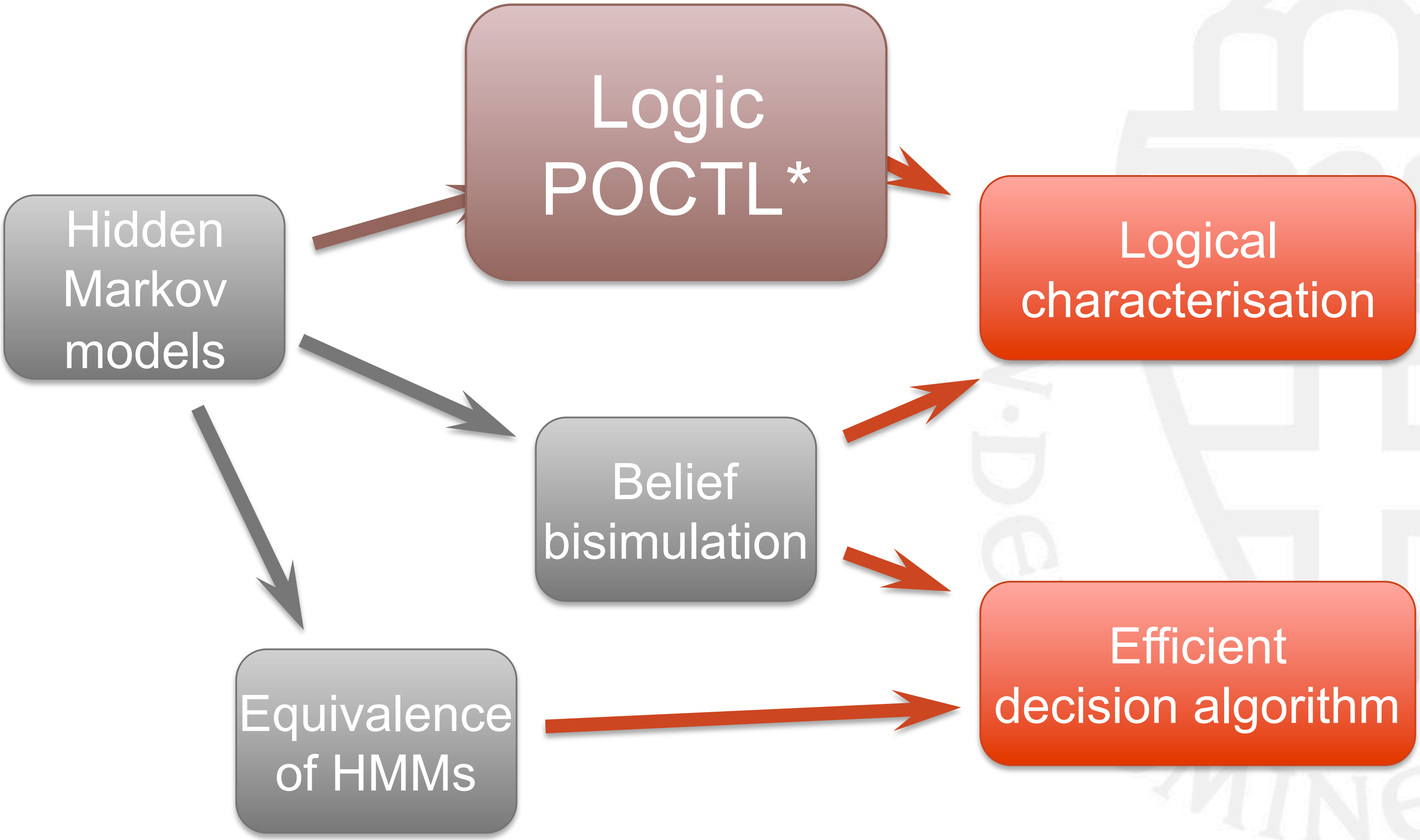
- Belief state or information state:
  what is known about the current state, based on observations

What is a HMM good for?

- speech recognition
  - What has been said, given a certain sequence of sounds?
- adaptive communication channel
  - What is the current reliability of the channel, given the recent pattern of breakdowns?
- biological sequence analysis
  - Is the probability of a match high enough, given the two sequences of acids/bases?
- card game
  - Which cards does the player still hold, given a certain play?

- Global question:
  What state has a high/the highest probability, given a certain sequence of observations?
- (Problem 2 in Rabiner '89)

Context and Overview



A logic for HMMs

- POCTL* = probabilistic and observation-CTL*
- extends CTL* and PCTL
- allows to specify constraints on observations

- Example formulas:
  - The probability is at least 0.5 that
    after observations ω1, ω2 and ω3, the system is in state *turned*.

$$P_{\geq 0.5}(X_{\omega 1}\ X_{\omega 2}\ X_{\omega 3}\ \mathit{turned})$$

  - The probability to get observation sequence (υ1 υ2 υ3) or (v1 v2)
    is smaller than 0.1.

$$P_{<0.1}(\ (X_{\upsilon 1}\ X_{\upsilon 2}\ X_{\upsilon 3}\ \text{true}) \vee (X_{v1}\ X_{v2}\ \text{true})\ )$$

- Model checking algorithm for these formulas exists
  [Zhang, Hermanns, Jansen: FORTE'05]

POCTL* syntax

state formulas

$$\Phi ::= \text{true} \mid a \mid \neg\Phi \mid \Phi \wedge \Phi \mid \varepsilon$$

path formulas

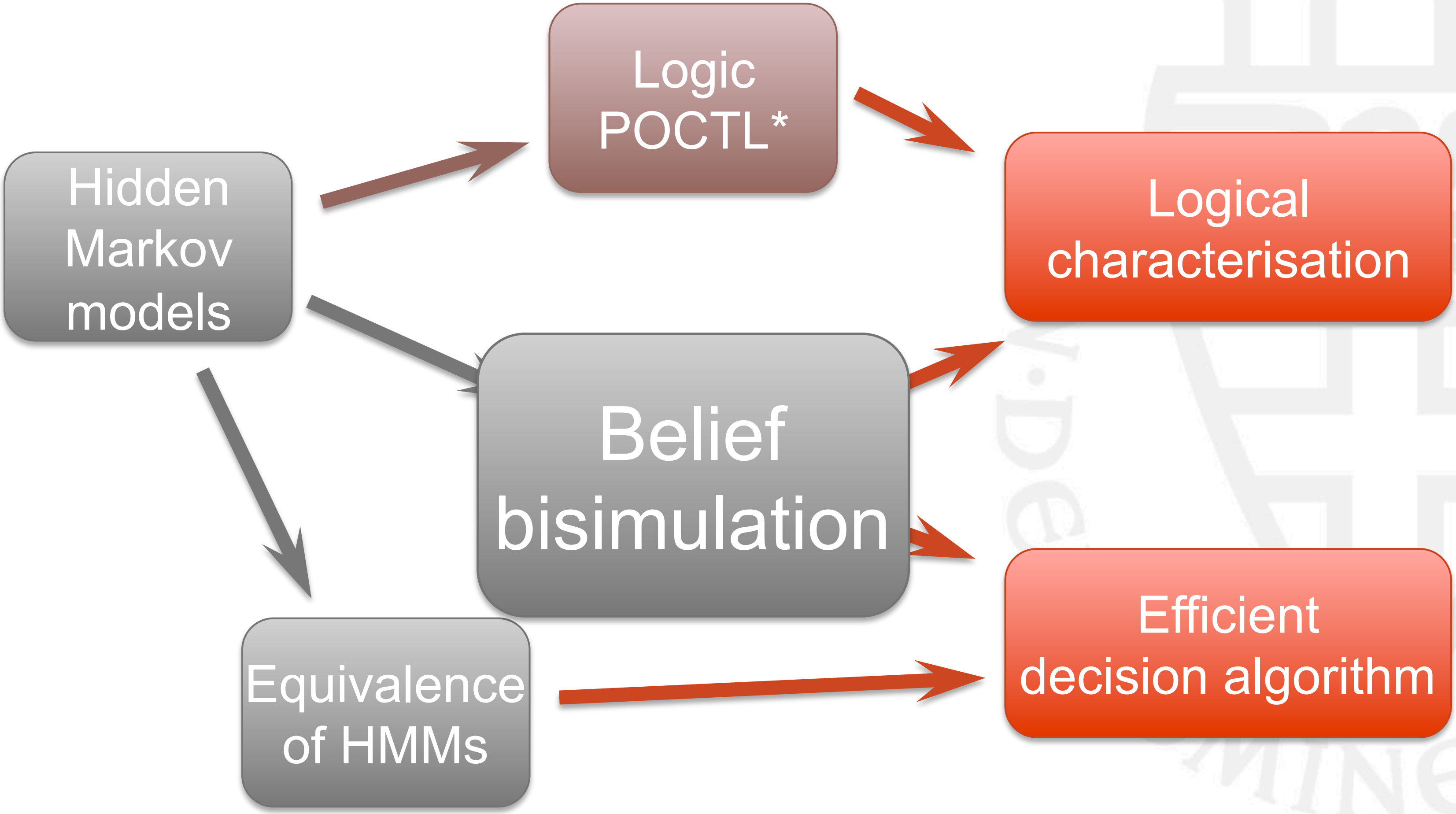
$$\varphi ::= \Phi \mid \neg\varphi \mid \varphi \wedge \varphi \mid X_{\Upsilon}\, \varphi \mid \varphi\ U^{\leq n}\ \varphi$$

belief state formulas

$$\varepsilon ::= \neg\varepsilon \mid \varepsilon \wedge \varepsilon \mid P_{\bowtie p}(\varphi)$$

Constraint on probability of paths with a given property

Context and Overview



What is bisimulation?

- Identify states that show equivalent behaviour
- Goal: smaller model → simpler model checking

- standard definition for Markov chains, extended to HMM:
- Equivalence relation $R \subseteq S \times S$
  is a bisimulation if
  - it respects the **labelling:**
    for all atomic propositions a, $L(s,a) = L(t,a)$
  - it respects the **observations:**
    for all $\omega \in \Omega$, $\mathrm{Prob}_s(\omega) = \mathrm{Prob}_t(\omega)$
  - it respects the **transitions** (conditional on ω):
    for all $\omega \in \Omega$ and equivalence classes $C \in S/R$, $\mathrm{Prob}(C \mid s \xrightarrow{\omega}) \ R \ \mathrm{Prob}(C \mid t \xrightarrow{\omega})$

  for all $(s,t) \in R$

- notation: $s \sim t$

strong belief bisimulation

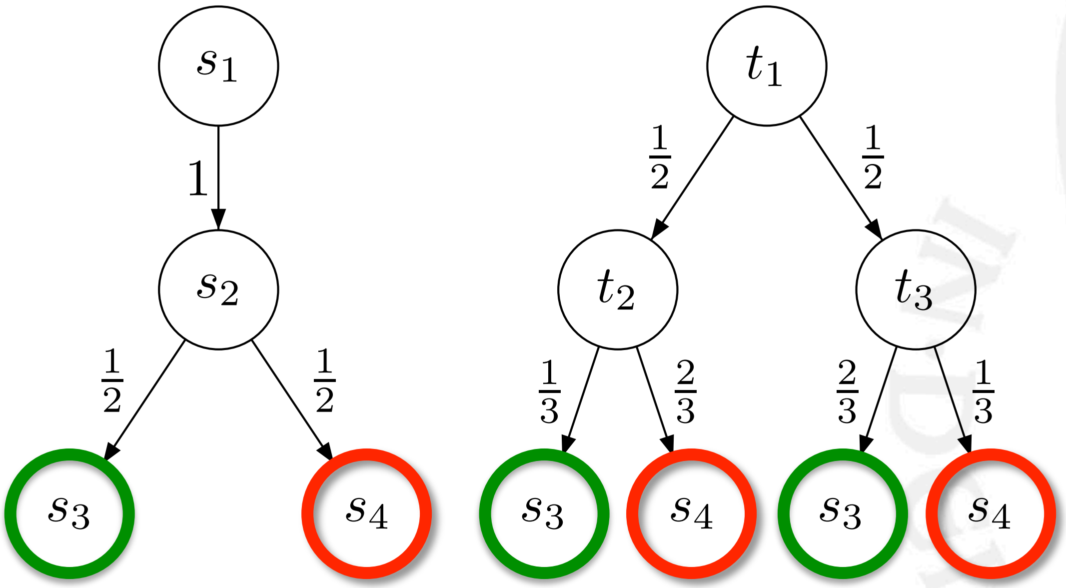
- Problem of (standard) bisimulation:
  relation between states, but states are invisible
- Solution:
  define relation between belief states

- Equivalence relation $R \subseteq \mathcal{B} \times \mathcal{B}$
  is a strong belief bisimulation if
  - it respects the **labelling:**
    for all atomic propositions a, $L(b,a) = L(c,a)$
  - it respects the **observations:**
    for all $\omega \in \Omega$, $\mathrm{Prob}_b(\omega) = \mathrm{Prob}_c(\omega)$
  - it respects the **transitions** (conditional on ω):
    for all $\omega \in \Omega$, $\mathrm{Prob}(* \mid b \xrightarrow{\omega})\ R\ \mathrm{Prob}(* \mid c \xrightarrow{\omega})$

  for all $(b,c) \in R$

- notation: $b \sim_{sb} c$

bisimulation vs. belief bisimulation



$s_1 \not\sim t_1$

$\{s_1 \mapsto 1\} \sim_{sb} \{t_1 \mapsto 1\}$

weak belief bisimulation

- Equivalence relation $R \subseteq \mathcal{B} \times \mathcal{B}$
  is a weak belief bisimulation if
  - it respects the **labelling:**
    for all atomic propositions a, $L(b,a) = L(c,a)$
  - it respects the **observations:**
    for all $\omega \in \Omega$, $\mathrm{Prob}_b(\omega) = \mathrm{Prob}_c(\omega)$
  - it respects the **transitions:**
    for all equivalence classes $B \in \mathcal{B}/R$,
    $\mathrm{Prob}(B \mid b \rightarrow) = \mathrm{Prob}(B \mid c \rightarrow)$

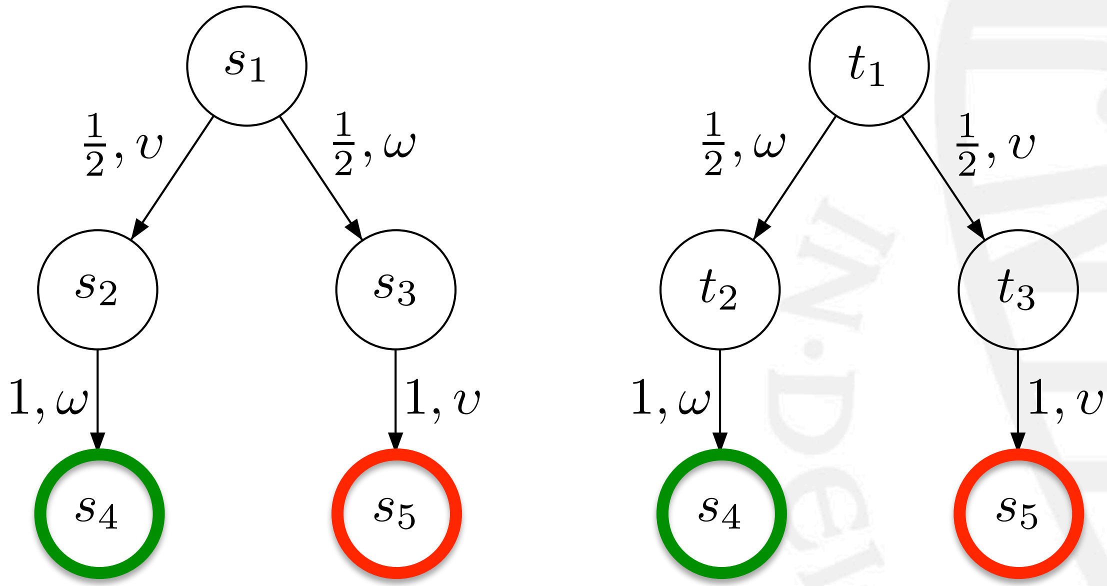
  for all $(b,c) \in R$

- notation: $b \sim_{\mathrm{wb}} c$

in strong belief bisimulation:

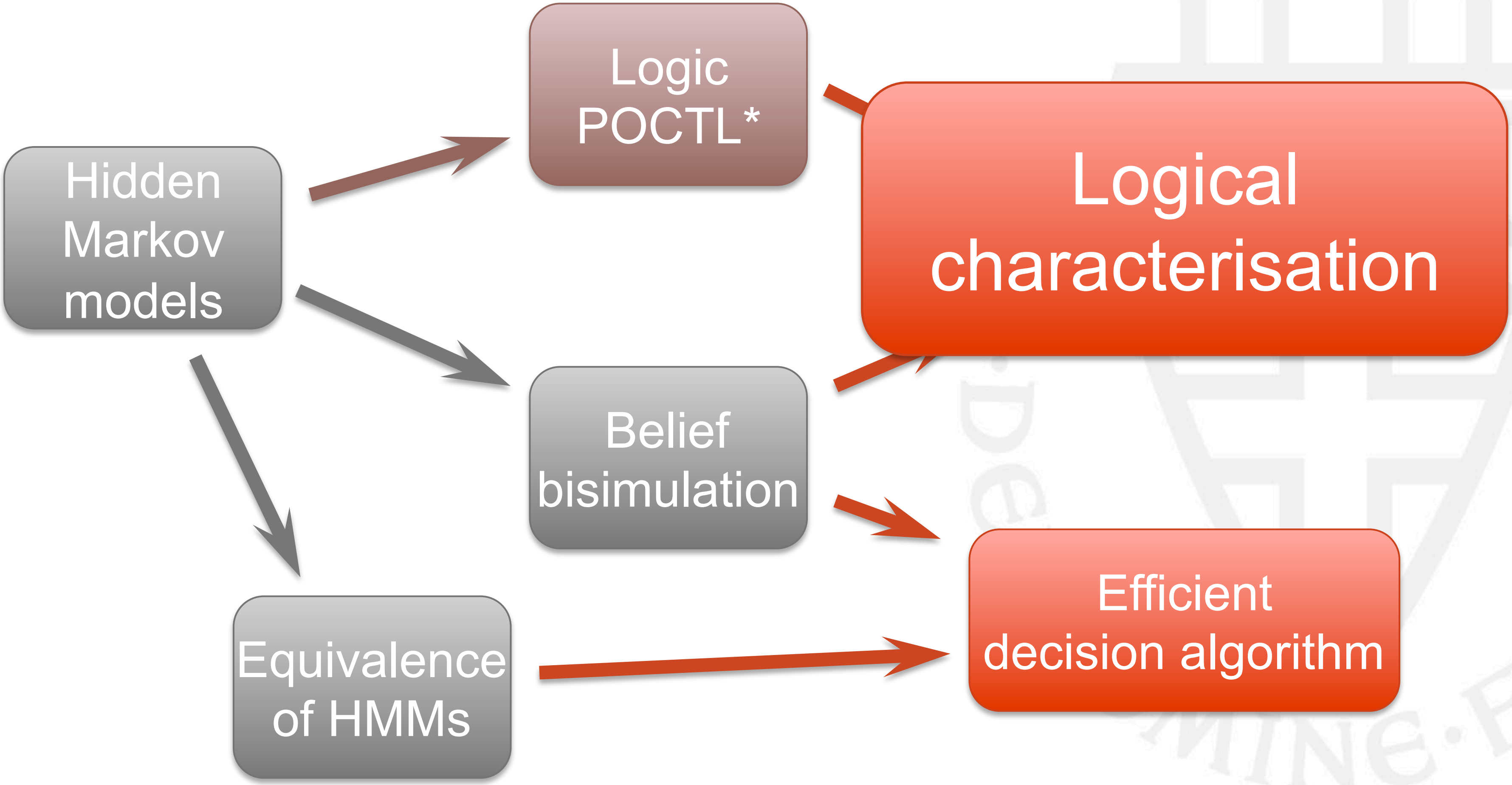
it respects the transitions (conditional on ω):
for all $\omega \in \Omega$, $\mathrm{Prob}(* \mid b \xrightarrow{\omega})\ R\ \mathrm{Prob}(* \mid c \xrightarrow{\omega})$

strong vs. weak belief bisimulation



$\{s_1 \mapsto 1\} \not\sim_{sb} \{t_1 \mapsto 1\}$ $\{s_1 \mapsto 1\} \sim_{wb} \{t_1 \mapsto 1\}$

Context and Overview

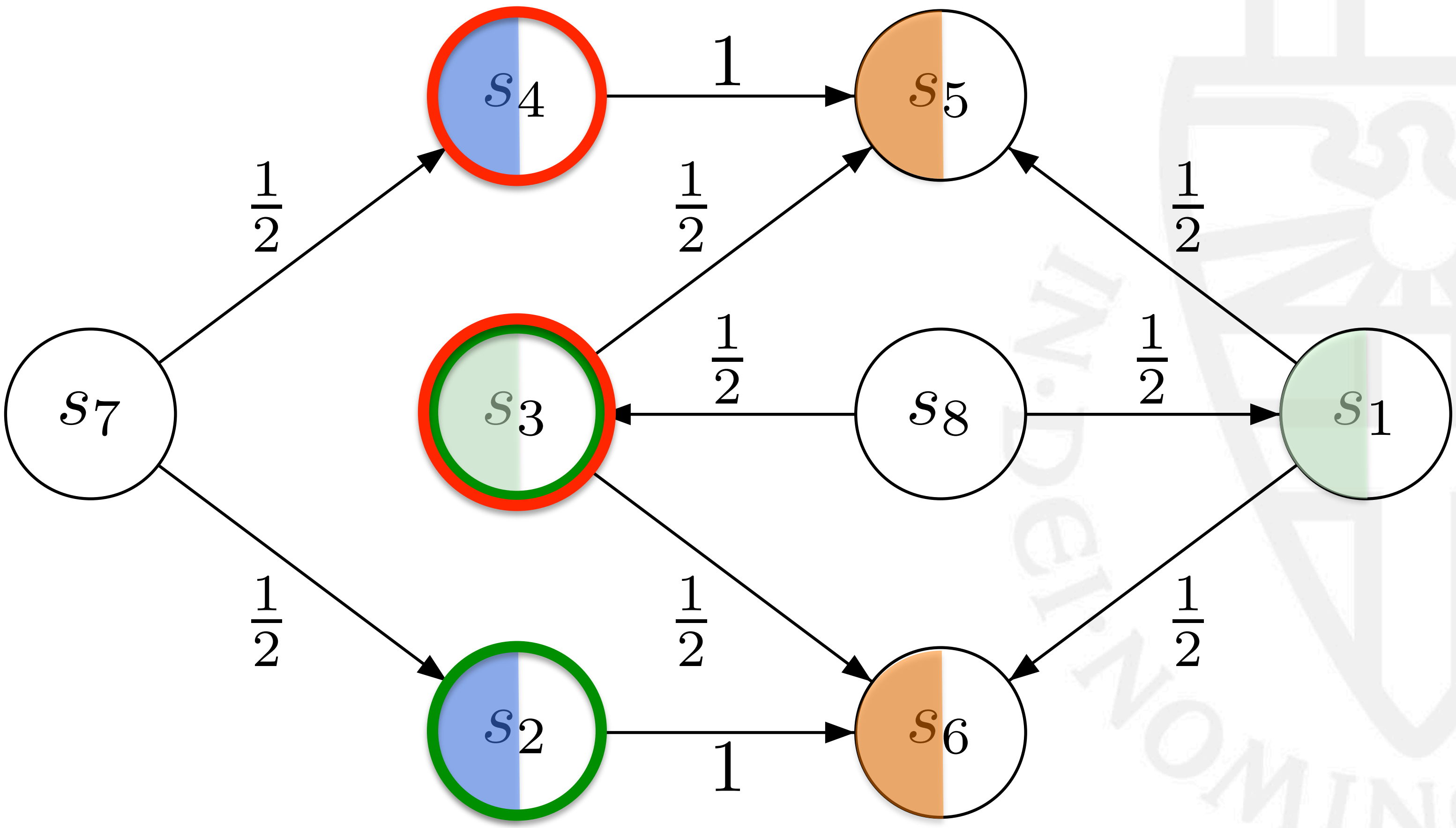


logical characterisation

Goal of logical characterisation:
Find a logic with the property
"Two states are bisimilar ($\sim$, $\sim_{sb}$, $\sim_{wb}$) iff they satisfy the same formulas."

Relation	Characterising logic
$\sim$	POCTL*
$\sim_{sb}$	SBBL*
$\sim_{wb}$	WBBL*

POCTL* is too strong for $\sim_{sb}$



POCTL* is too strong for $\sim_{sb}$

• ● $\models$	● $\not\models$	$P_{\geq 0.5}(P_{\geq 1}(X\ a_3))$
• ● $\not\models$	● $\models$	$P_{\geq 0.5}(a_1 \wedge a_2)$
• $\{s_7 \mapsto 1\} \not\models$	$\{s_8 \mapsto 1\} \models$	$P_{\geq 0.5}(X\ (a_1 \wedge a_2))$
• $\{s_7 \mapsto 1\} \models$	$\{s_8 \mapsto 1\} \not\models$	$P_{=0}((X\ a_1)\ U^{\leq \infty}\ a_2)$
• $\{s_7 \mapsto 1\} \models$	$\{s_8 \mapsto 1\} \not\models$	$P_{=0}(\neg a_1\ U^{\leq \infty}\ (a_2\ U^{\leq \infty}\ a_3))$

SBBL*

state formulas

$$\Phi ::= \text{true} \mid a \mid \neg\Phi \mid \Phi \wedge \Phi \mid \varepsilon$$

path formulas

$$\varphi ::= \Phi \mid \neg\varphi \mid \varphi \wedge \varphi \mid X_Y\, \varphi \mid \varphi\, U^{\leq n}\, \varphi$$

belief state formulas

$$\varepsilon ::= \neg\varepsilon \mid \varepsilon \wedge \varepsilon \mid P_{\bowtie p}(\varphi) \mid P_{\bowtie p}(\Phi\, U^{\leq n}\, \Phi)$$

SBBL* characterises strong belief bisimilarity.

WBBL*

state formulas

$$\Phi ::= \text{true} \mid a \mid \neg\Phi$$

path formulas

$$\varphi ::= \Phi \mid X_Y\ \varphi \mid X_Y\ \text{true}$$

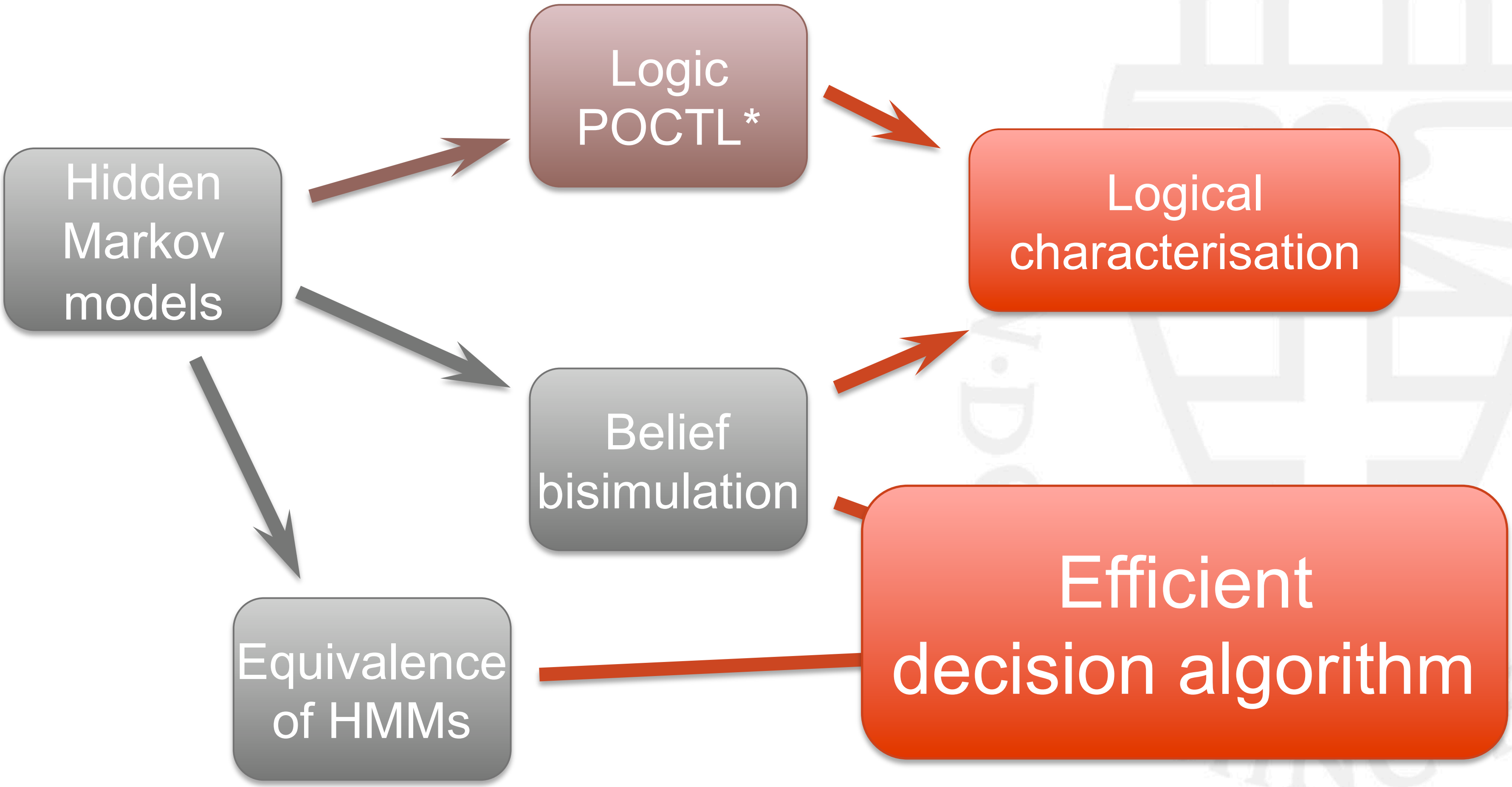
Example formulas are not in WBBL*

belief state formulas

$$\varepsilon ::= \neg\varepsilon \mid \varepsilon \wedge \varepsilon \mid P_{\bowtie p}(\varphi) \mid P_{\bowtie p}(\Phi\ U^{\leq n}\ \Phi)$$

WBBL* characterises weak belief bisimilarity.

Context and Overview



Deciding belief bisimilarity

- problem: $\mathcal{B}$ is uncountable
  $\rightarrow$ no complete description of $\mathcal{B}/\sim_{sb}$ or $\mathcal{B}/\sim_{wb}$

- solution:
  - equation system over variables
    $b_1 := b(s_1)$, $b_2 := b(s_2)$, ..., $b_n := b(s_n)$,
    $c_1 := c(s_1)$, $c_2 := c(s_2)$, ..., $c_n := c(s_n)$
  - equality holds $\Leftrightarrow$ $b \sim_{sb} c$

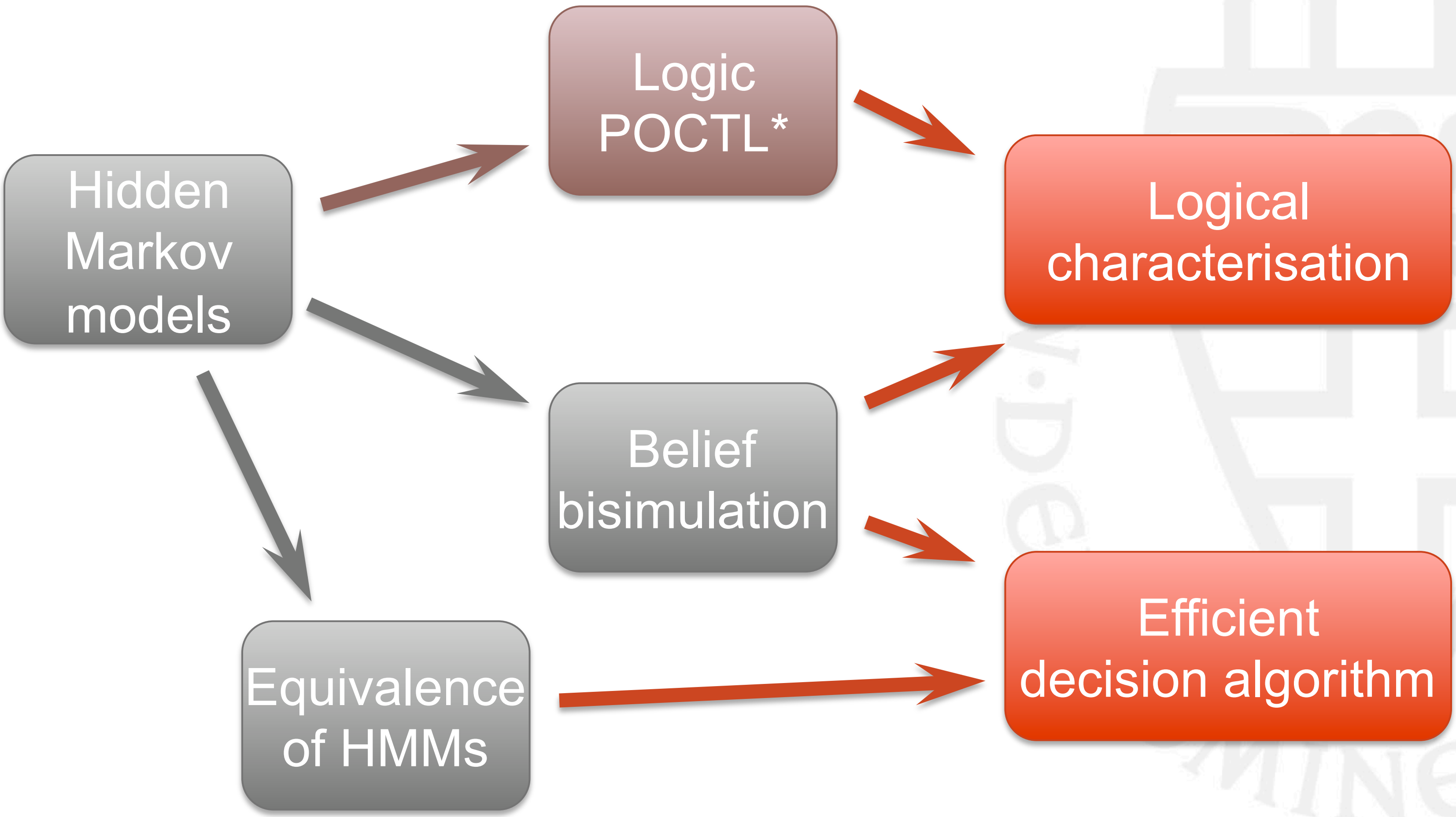
Iterative generation of equation system

- basis of iteration:
  belief states b and c should have same probability of label or observation
  $$\sum_{s_i \models a} b_i = L(b,a) = L(c,a) = \sum_{s_i \models a} c_i$$
- induction step:
  + regard one more transition
  + bring into upper triangular form
  + delete linearly dependent equations
- equation system with $2|S|$ variables
  - special form: coefficient for c_i is always negative of coefficient for b_i
    → at most $|S|$ linearly independent equations

time complexity

- bringing to upper triangular form is most costly step
- weak belief bisimilarity:
  - generate $\leq |S| + |AP| + |\Omega|$ equations with $2|S|$ unknowns
  - time complexity is $\in O(|S|^3)$
- strong belief bisimilarity:
  - generate $\leq \max\{|AP| + |\Omega|, |S|\cdot|\Omega|\}$ equations with $2|S|$ unknowns
  - time complexity is $\in O(|S|^3|\Omega|)$
- better than $\in O(|S|^4)$
  [Doyen, Henzinger, Raskin: FCS'08]

Wrapping up



Hidden Markov models
Logic POCTL*
Belief bisimulation
Equivalence of HMMs
Logical characterisation
Efficient decision algorithm

Conclusions

- WBBL* is too weak to describe standard property on HMMs: “What state has a high/the highest probability, given a certain sequence of observations?”
- Weak belief bisimulation does not preserve this property

- Improved time complexity from $O(|S|^4)$ to $O(|S|^3)$