

# A Pre-Expectation Calculus for Probabilistic Sensitivity

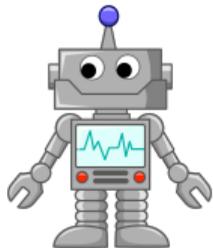
Alejandro Aguirre, Gilles Barthe,  
Justin Hsu\*, Benjamin Kaminski,  
Joost-Pieter Katoen, Christoph Matheja

Originally appeared at POPL 2021

# Reinforcement learning: a quick overview

# Reinforcement learning: a quick overview

State  $s_1 \in \mathcal{S}$

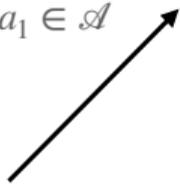
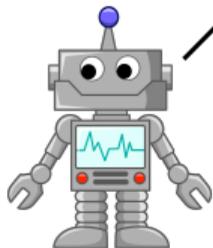


# Reinforcement learning: a quick overview

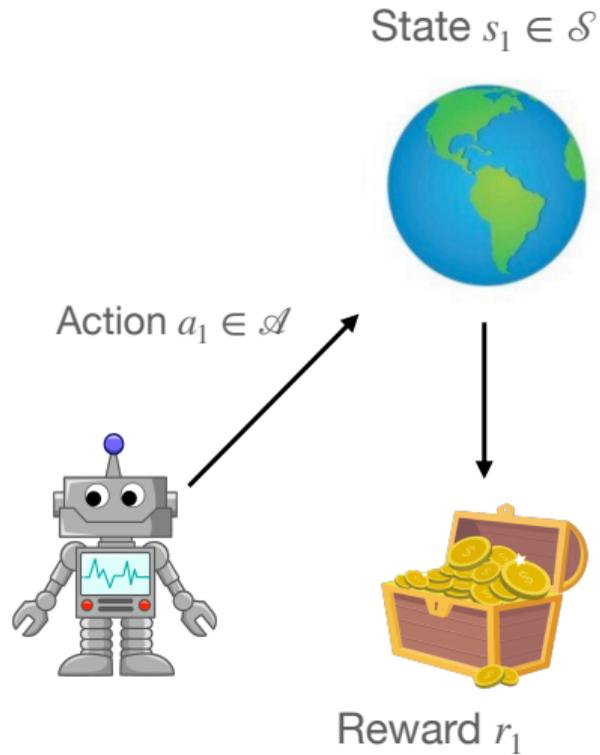
State  $s_1 \in \mathcal{S}$



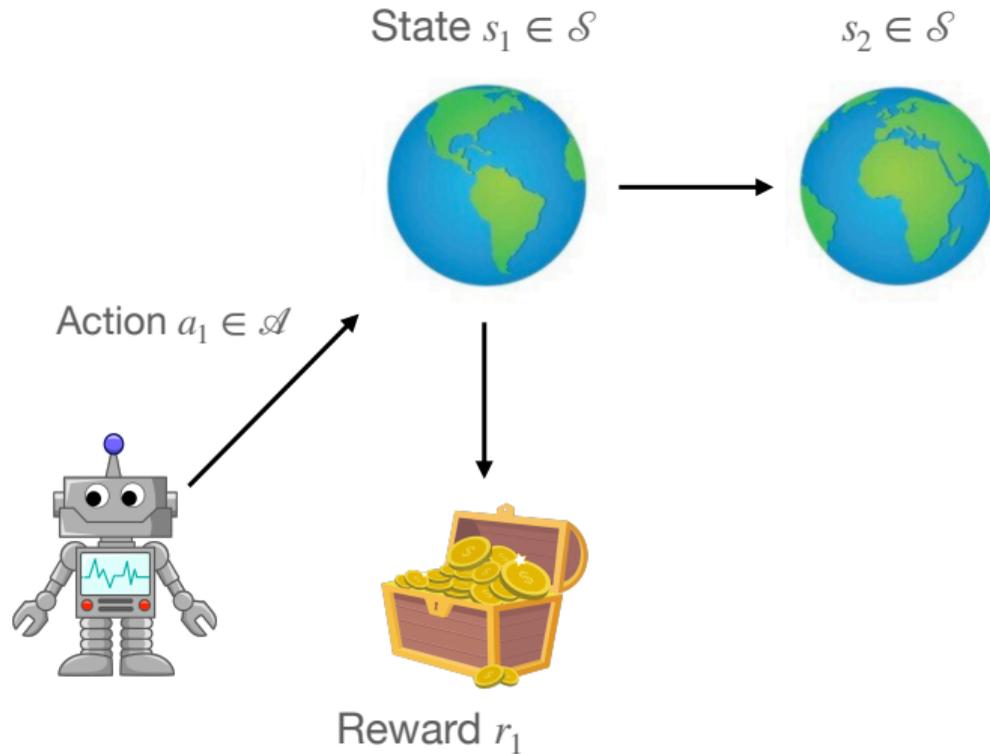
Action  $a_1 \in \mathcal{A}$



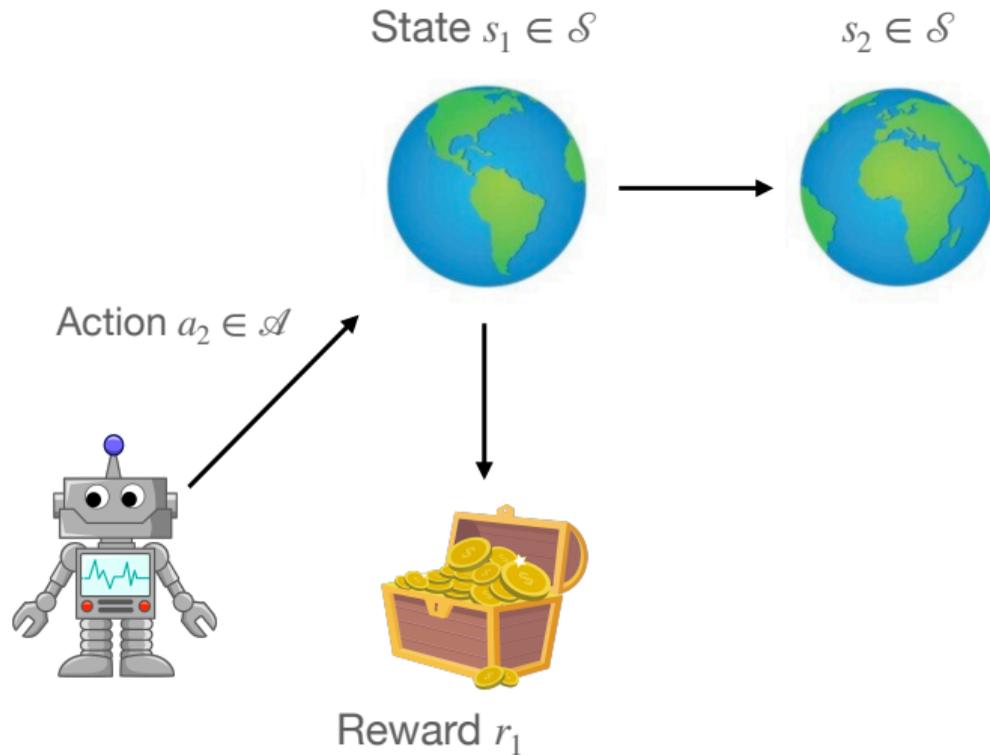
# Reinforcement learning: a quick overview



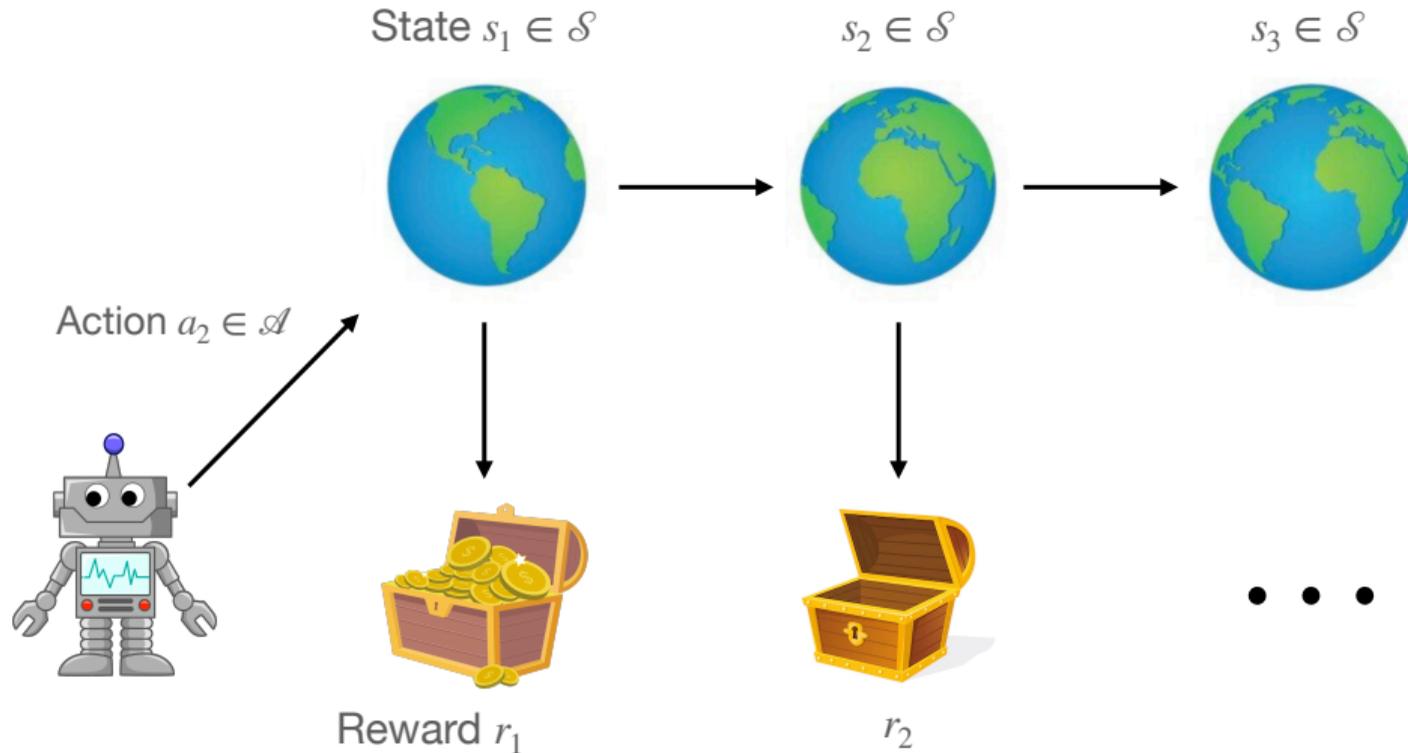
# Reinforcement learning: a quick overview



# Reinforcement learning: a quick overview



# Reinforcement learning: a quick overview



# Some terminology

## State transition function $\mathcal{P}$

- ▶ Maps state  $s$  and action  $a$  to random new state  $s'$
- ▶ Learner doesn't know this function, can only draw samples

## Reward function $\mathcal{R}$

- ▶ Maps state  $s$  and action  $a$  to random reward  $r \in [0, 1]$
- ▶ Learner doesn't know this function, can only draw samples

## Policy function $\pi$

- ▶ Maps state  $s$  to an action  $a$  to play

**Reinforcement learning:** find optimal policy  $\pi$  to maximize total expected reward

# Task: Estimating the value of a policy $\pi$

## Example: TD(0) algorithm

TD0( $V$ )

$n \leftarrow 0$ ;

**while**  $n < N$  **do**

$i \leftarrow 0$ ;

**while**  $i < |\mathcal{S}|$  **do**

$a \xleftarrow{\$} \pi(i); r \xleftarrow{\$} \mathcal{R}(i, a); j \xleftarrow{\$} \mathcal{P}(i, a)$ ;

$W[i] \leftarrow (1 - \alpha) \cdot V[i] + \alpha \cdot (r + \gamma \cdot V[j])$ ;

$i \leftarrow i + 1$

$V \leftarrow W; n \leftarrow n + 1$ ;

# Task: Estimating the value of a policy $\pi$

## Example: TD(0) algorithm

**TD0**( $V$ )

$n \leftarrow 0$ ;

**while**  $n < N$  **do**

$i \leftarrow 0$ ;

**while**  $i < |\mathcal{S}|$  **do**

$a \xleftarrow{\$} \pi(i); r \xleftarrow{\$} \mathcal{R}(i, a); j \xleftarrow{\$} \mathcal{P}(i, a)$ ;

$W[i] \leftarrow (1 - \alpha) \cdot V[i] + \alpha \cdot (r + \gamma \cdot V[j])$ ;

$i \leftarrow i + 1$

$V \leftarrow W; n \leftarrow n + 1$ ;

## Input

- ▶ Initial guess  $V$ : value of each state

## Output

- ▶ Estimated value of each state
- ▶ Final estimate is randomized

# Task: Estimating the value of a policy $\pi$

## Example: TD(0) algorithm

TD0( $V$ )

$n \leftarrow 0$ ;

**while**  $n < N$  **do**

$i \leftarrow 0$ ;

**while**  $i < |\mathcal{S}|$  **do**

$a \xleftarrow{\$} \pi(i); r \xleftarrow{\$} \mathcal{R}(i, a); j \xleftarrow{\$} \mathcal{P}(i, a);$

$W[i] \leftarrow (1 - \alpha) \cdot V[i] + \alpha \cdot (r + \gamma \cdot V[j]);$

$i \leftarrow i + 1$

$V \leftarrow W; n \leftarrow n + 1;$

## Input

- ▶ Initial guess  $V$ : value of each state

## Output

- ▶ Estimated value of each state
- ▶ Final estimate is randomized

# Task: Estimating the value of a policy $\pi$

## Example: TD(o) algorithm

**TD0(V)**

$n \leftarrow 0;$

**while**  $n < N$  **do**

$i \leftarrow 0;$

**while**  $i < |\mathcal{S}|$  **do**

$a \overset{\$}{\leftarrow} \pi(i); r \overset{\$}{\leftarrow} \mathcal{R}(i, a); j \overset{\$}{\leftarrow} \mathcal{P}(i, a);$

$W[i] \leftarrow (1 - \alpha) \cdot V[i] + \alpha \cdot (r + \gamma \cdot V[j]);$

$i \leftarrow i + 1$

$V \leftarrow W; n \leftarrow n + 1;$

## Input

- ▶ Initial guess  $V$ : value of each state

## Output

- ▶ Estimated value of each state
- ▶ Final estimate is randomized

Our goal

**Verify:** the output of  $TD(o)$  doesn't depend "too much" on the input  $V$

More formally, want to verify:

If  $V$  and  $V'$  are any two possible inputs:

$$\text{Dist}(TD(0)(V), TD(0)(V')) \leq \epsilon$$

Here,  $\text{Dist}$  is a distance between pairs of outputs (**distributions**).

More formally, want to verify:

If  $V$  and  $V'$  are any two possible inputs:

$$Dist(TD(0)(V), TD(0)(V')) \leq \epsilon$$

Here,  $Dist$  is a distance between pairs of outputs (**distributions**).

Even better: verify rate of convergence

$$Dist(TD(0)(V), TD(0)(V')) \leq (1 - \epsilon)^N \cdot dist(V, V')$$

Here,  $dist$  is a distance between pairs of inputs (**not** distributions).

More generally: want to verify probabilistic sensitivity

$$\text{Dist}(\text{Prog}(in), \text{Prog}(in')) \leq \text{dist}(in, in')$$

More generally: want to verify probabilistic sensitivity

$$\text{Dist}(\text{Prog}(in), \text{Prog}(in')) \leq \text{dist}(in, in')$$

Intuition: **small changes** in the input memory  
lead to **small changes** in the output distribution

# Our Verification Method: Relational Pre-Expectations

# Technical contributions, in three steps

## Technical contributions, in three steps

- Define relational pre-expectation transformer  $rpe$

## Technical contributions, in three steps

- Define relational pre-expectation transformer  $rpe$
- Propose a set of proof rules for bounding  $rpe$

## Technical contributions, in three steps

- Define relational pre-expectation transformer  $rpe$
- Propose a set of proof rules for bounding  $rpe$
- Prove soundness: bounding  $rpe$  implies probabilistic sensitivity property

## Step 1: Defining the relational pre-expectation transformer

## Step 1: Defining the relational pre-expectation transformer

**Given:** distance  $dist : M \times M \rightarrow \mathbb{R}$  and probabilistic program  $c$

## Step 1: Defining the relational pre-expectation transformer

**Given:** distance  $dist : M \times M \rightarrow \mathbb{R}$  and probabilistic program  $c$

**Define:** distance  $rpe(c, dist) : M \times M \rightarrow \mathbb{R}$  in terms of  $rpe$  for **subprograms** of  $c$

## Step 1: Defining the relational pre-expectation transformer

**Given:** distance  $dist : M \times M \rightarrow \mathbb{R}$  and probabilistic program  $c$

**Define:** distance  $rpe(c, dist) : M \times M \rightarrow \mathbb{R}$  in terms of  $rpe$  for **subprograms** of  $c$

$$\widetilde{rpe}(\text{skip}, \mathcal{E}) \triangleq \mathcal{E}$$

$$\begin{aligned}\widetilde{rpe}(x \leftarrow e, \mathcal{E}) &\triangleq \mathcal{E}\{e\langle 1 \rangle, e\langle 2 \rangle / x\langle 1 \rangle, x\langle 2 \rangle\} \\ &\triangleq \lambda s_1 s_2. \mathcal{E}(s_1[x \mapsto e\langle 1 \rangle], s_2[x \mapsto e\langle 2 \rangle])\end{aligned}$$

$$\widetilde{rpe}(x \stackrel{\$}{\leftarrow} d, \mathcal{E}) \triangleq \lambda s_1 s_2. \mathcal{E}^\#(\llbracket x \stackrel{\$}{\leftarrow} d \rrbracket_{s_1}, \llbracket x \stackrel{\$}{\leftarrow} d \rrbracket_{s_2}), \text{ where } \mathcal{E}^\#(\mu_1, \mu_2) \triangleq \inf_{\mu \in \Gamma(\mu_1, \mu_2)} \mathbb{E}_\mu[\mathcal{E}]$$

$$\widetilde{rpe}(c; c', \mathcal{E}) \triangleq \widetilde{rpe}(c, \widetilde{rpe}(c', \mathcal{E}))$$

$$\widetilde{rpe}(\text{if } e \text{ then } c \text{ else } c', \mathcal{E}) \triangleq [e\langle 1 \rangle \wedge e\langle 2 \rangle] \cdot \widetilde{rpe}(c, \mathcal{E}) + [\neg e\langle 1 \rangle \wedge \neg e\langle 2 \rangle] \cdot \widetilde{rpe}(c', \mathcal{E}) + [e\langle 1 \rangle \neq e\langle 2 \rangle] \cdot \infty$$

$$\widetilde{rpe}(\text{while } e \text{ do } c, \mathcal{E}) \triangleq \text{lfp} X. \Phi_{\mathcal{E}, c}(X),$$

$$\text{where } \Phi_{\mathcal{E}, c}(X) \triangleq [e\langle 1 \rangle \wedge e\langle 2 \rangle] \cdot \widetilde{rpe}(c, X) + [\neg e\langle 1 \rangle \wedge \neg e\langle 2 \rangle] \cdot \mathcal{E} + [e\langle 1 \rangle \neq e\langle 2 \rangle] \cdot \infty$$

# Step 1: Defining the relational pre-expectation transformer

**Given:** distance  $dist : M \times M \rightarrow \mathbb{R}$  and probabilistic program  $c$

**Define:** distance  $rpe(c, dist) : M \times M \rightarrow \mathbb{R}$  in terms of  $rpe$  for **subprograms** of  $c$

$$\begin{aligned} \widetilde{rpe}(\text{skip}, \mathcal{E}) &\triangleq \mathcal{E} \\ \widetilde{rpe}(x \leftarrow e, \mathcal{E}) &\triangleq \mathcal{E}\{e\langle 1 \rangle, e\langle 2 \rangle / x\langle 1 \rangle, x\langle 2 \rangle\} \\ &\triangleq \lambda s_1 s_2. \mathcal{E}(s_1[x \mapsto e\langle 1 \rangle], s_2[x \mapsto e\langle 2 \rangle]) \\ \widetilde{rpe}(x \stackrel{\$}{\leftarrow} d, \mathcal{E}) &\triangleq \lambda s_1 s_2. \mathcal{E}^\#(\llbracket x \stackrel{\$}{\leftarrow} d \rrbracket_{s_1}, \llbracket x \stackrel{\$}{\leftarrow} d \rrbracket_{s_2}), \text{ where } \mathcal{E}^\#(\mu_1, \mu_2) \triangleq \inf_{\mu \in \Gamma(\mu_1, \mu_2)} \mathbb{E}_\mu[\mathcal{E}] \\ \widetilde{rpe}(c; c', \mathcal{E}) &\triangleq \widetilde{rpe}(c, \widetilde{rpe}(c', \mathcal{E})) \\ \widetilde{rpe}(\text{if } e \text{ then } c \text{ else } c', \mathcal{E}) &\triangleq [e\langle 1 \rangle \wedge e\langle 2 \rangle] \cdot \widetilde{rpe}(c, \mathcal{E}) + [\neg e\langle 1 \rangle \wedge \neg e\langle 2 \rangle] \cdot \widetilde{rpe}(c', \mathcal{E}) + [e\langle 1 \rangle \neq e\langle 2 \rangle] \cdot \infty \\ \widetilde{rpe}(\text{while } e \text{ do } c, \mathcal{E}) &\triangleq \text{lfp} X. \Phi_{\mathcal{E}, c}(X), \\ &\text{ where } \Phi_{\mathcal{E}, c}(X) \triangleq [e\langle 1 \rangle \wedge e\langle 2 \rangle] \cdot \widetilde{rpe}(c, X) + [\neg e\langle 1 \rangle \wedge \neg e\langle 2 \rangle] \cdot \mathcal{E} + [e\langle 1 \rangle \neq e\langle 2 \rangle] \cdot \infty \end{aligned}$$

## Step 2: Bounding relational pre-expectations

Recall our goal: verify probabilistic sensitivity

$$\mathit{Dist}(c(in), c(in')) \leq \mathit{dist}(in, in')$$

## Step 2: Bounding relational pre-expectations

Recall our goal: verify probabilistic sensitivity

$$\text{Dist}(c(in), c(in')) \leq \text{dist}(in, in')$$

Strategy: verify something a bit different

$$\text{rpe}(c, d)(in, in') \leq \text{dist}(in, in')$$

## Step 2: Bounding relational pre-expectations

### Lots of proof rules

$$\frac{\mathcal{E} \leq \mathcal{E}'}{\widetilde{rpe}(c, \mathcal{E}) \leq \widetilde{rpe}(c, \mathcal{E}')} \text{ MONO}$$

$$\frac{FV(\mathcal{E}') \cap MV(c) = \emptyset}{\widetilde{rpe}(c, \mathcal{E} + \mathcal{E}') \leq \widetilde{rpe}(c, \mathcal{E}) + \mathcal{E}'} \text{ CONST}$$

$$\frac{}{\widetilde{rpe}(c, \mathcal{E}) + \widetilde{rpe}(c, \mathcal{E}') \leq \widetilde{rpe}(c, \mathcal{E} + \mathcal{E}')} \text{ SUPADD}$$

$$\frac{f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \text{ linear, with } f(\infty) \triangleq \infty}{\widetilde{rpe}(c, f \circ \mathcal{E}) = f \circ \widetilde{rpe}(c, \mathcal{E})} \text{ SCALE}$$

$$\frac{M : \text{State} \times \text{State} \rightarrow \Gamma(\llbracket d \rrbracket, \llbracket d \rrbracket)}{\widetilde{rpe}(x \stackrel{\$}{\leftarrow} d, \mathcal{E}) \leq \mathbb{E}_{(v_1, v_2) \sim M(-, -)}[\mathcal{E}\{v_1, v_2/x\langle 1 \rangle, x\langle 2 \rangle\}]} \text{ SAMP}$$

$$\frac{f : \text{State} \times \text{State} \rightarrow (D \rightarrow D) \text{ bijection}}{\widetilde{rpe}(x \stackrel{\$}{\leftarrow} U(D), \mathcal{E}) \leq \frac{1}{|D|} \sum_{v \in D} \mathcal{E}\{v, f(-, -)(v)/x\langle 1 \rangle, x\langle 2 \rangle\}} \text{ UNIF}$$

$$\frac{[e\langle 1 \rangle \wedge e\langle 2 \rangle] \cdot \widetilde{rpe}(c, \mathcal{I}) + [\neg e\langle 1 \rangle \wedge \neg e\langle 2 \rangle] \cdot \mathcal{E} + [e\langle 1 \rangle \neq e\langle 2 \rangle] \cdot \infty \leq \mathcal{I}}{\widetilde{rpe}(\text{while } e \text{ do } c, \mathcal{E}) \leq \mathcal{I}} \text{ INV}$$

# Step 2: Bounding relational pre-expectations

## Lots of proof rules

$$\frac{\mathcal{E} \leq \mathcal{E}'}{\widetilde{rpe}(c, \mathcal{E}) \leq \widetilde{rpe}(c, \mathcal{E}')} \text{ MONO}$$

$$\frac{FV(\mathcal{E}') \cap MV(c) = \emptyset}{\widetilde{rpe}(c, \mathcal{E} + \mathcal{E}') \leq \widetilde{rpe}(c, \mathcal{E}) + \mathcal{E}'} \text{ CONST}$$

$$\frac{}{\widetilde{rpe}(c, \mathcal{E}) + \widetilde{rpe}(c, \mathcal{E}') \leq \widetilde{rpe}(c, \mathcal{E} + \mathcal{E}')} \text{ SUPADD}$$

$$\frac{f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \text{ linear, with } f(\infty) \triangleq \infty}{\widetilde{rpe}(c, f \circ \mathcal{E}) = f \circ \widetilde{rpe}(c, \mathcal{E})} \text{ SCALE}$$

$$M : \text{State} \times \text{State} \rightarrow \Gamma(\llbracket d \rrbracket, \llbracket d \rrbracket)$$

$$\widetilde{rpe}(x \stackrel{\$}{\leftarrow} d, \mathcal{E}) \leq \mathbb{E}_{(v_1, v_2) \sim M(-, -)}[\mathcal{E}\{v_1, v_2/x\langle 1 \rangle, x\langle 2 \rangle\}] \text{ SAMP}$$

$$f : \text{State} \times \text{State} \rightarrow (D \rightarrow D) \text{ bijection}$$

$$\widetilde{rpe}(x \stackrel{\$}{\leftarrow} U(D), \mathcal{E}) \leq \frac{1}{|D|} \sum_{v \in D} \mathcal{E}\{v, f(-, -)(v)/x\langle 1 \rangle, x\langle 2 \rangle\} \text{ UNIF}$$

$$\frac{[e\langle 1 \rangle \wedge e\langle 2 \rangle] \cdot \widetilde{rpe}(c, \mathcal{I}) + [\neg e\langle 1 \rangle \wedge \neg e\langle 2 \rangle] \cdot \mathcal{E} + [e\langle 1 \rangle \neq e\langle 2 \rangle] \cdot \infty \leq \mathcal{I}}{\widetilde{rpe}(\text{while } e \text{ do } c, \mathcal{E}) \leq \mathcal{I}} \text{ INV}$$

## Step 2: Bounding relational pre-expectations

### Lots of proof rules

$$\frac{\mathcal{E} \leq \mathcal{E}'}{\widetilde{rpe}(c, \mathcal{E}) \leq \widetilde{rpe}(c, \mathcal{E}')} \text{ MONO}$$

$$\frac{FV(\mathcal{E}') \cap MV(c) = \emptyset}{\widetilde{rpe}(c, \mathcal{E} + \mathcal{E}') \leq \widetilde{rpe}(c, \mathcal{E}) + \mathcal{E}'} \text{ CONST}$$

$$\frac{}{\widetilde{rpe}(c, \mathcal{E}) + \widetilde{rpe}(c, \mathcal{E}') \leq \widetilde{rpe}(c, \mathcal{E} + \mathcal{E}')} \text{ SUPADD}$$

$$\frac{f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \text{ linear, with } f(\infty) \triangleq \infty}{\widetilde{rpe}(c, f \circ \mathcal{E}) = f \circ \widetilde{rpe}(c, \mathcal{E})} \text{ SCALE}$$

$$\frac{M : \text{State} \times \text{State} \rightarrow \Gamma(\llbracket d \rrbracket, \llbracket d \rrbracket)}{\widetilde{rpe}(x \stackrel{\$}{\leftarrow} d, \mathcal{E}) \leq \mathbb{E}_{(v_1, v_2) \sim M(-, -)}[\mathcal{E}\{v_1, v_2/x\langle 1 \rangle, x\langle 2 \rangle\}]} \text{ SAMP}$$

$$\frac{f : \text{State} \times \text{State} \rightarrow (D \rightarrow D) \text{ bijection}}{\widetilde{rpe}(x \stackrel{\$}{\leftarrow} U(D), \mathcal{E}) \leq \frac{1}{|D|} \sum_{v \in D} \mathcal{E}\{v, f(-, -)(v)/x\langle 1 \rangle, x\langle 2 \rangle\}} \text{ UNIF}$$

$$\frac{[e\langle 1 \rangle \wedge e\langle 2 \rangle] \cdot \widetilde{rpe}(c, \mathcal{I}) + [\neg e\langle 1 \rangle \wedge \neg e\langle 2 \rangle] \cdot \mathcal{E} + [e\langle 1 \rangle \neq e\langle 2 \rangle] \cdot \infty \leq \mathcal{I}}{\widetilde{rpe}(\text{while } e \text{ do } c, \mathcal{E}) \leq \mathcal{I}} \text{ INV}$$

## Step 3: Proving the soundness theorem

Key construction: Kantorovich metric  $Kant(d)$

- ▶ Lifts distance  $d$  on memories to distance  $Kant(d)$  on distributions
- ▶ Varying  $d$  leads to different distances between distributions

## Step 3: Proving the soundness theorem

Key construction: Kantorovich metric  $Kant(d)$

- ▶ Lifts distance  $d$  on memories to distance  $Kant(d)$  on distributions
- ▶ Varying  $d$  leads to different distances between distributions

Main Theorem

$$Kant(d)(c(in), c(in')) \leq rpe(c, d)(in, in')$$

## Step 3: Proving the soundness theorem

Key construction: Kantorovich metric  $Kant(d)$

- ▶ Lifts distance  $d$  on memories to distance  $Kant(d)$  on distributions
- ▶ Varying  $d$  leads to different distances between distributions

Main Theorem

$$Kant(d)(c(in), c(in')) \leq rpe(c, d)(in, in')$$

Combine with upper-bound on  $rpe$  to verify sensitivity property:

$$Kant(d)(c(in), c(in')) \leq rpe(c, d)(in, in') \leq dist(in, in')$$

# Task: Estimating the value of a policy $\pi$

## Example: TD(0) algorithm

**TD0**( $V$ )

$n \leftarrow 0$ ;

**while**  $n < N$  **do**

$i \leftarrow 0$ ;

**while**  $i < |\mathcal{S}|$  **do**

$a \xleftarrow{\$} \pi(i)$ ;  $r \xleftarrow{\$} \mathcal{R}(i, a)$ ;  $j \xleftarrow{\$} \mathcal{P}(i, a)$ ;

$W[i] \leftarrow (1 - \alpha) \cdot V[i] + \alpha \cdot (r + \gamma \cdot V[j])$ ;

$i \leftarrow i + 1$

$V \leftarrow W$ ;  $n \leftarrow n + 1$ ;

## Input

- ▶ Initial guess  $V$ : value of each state

## Output

- ▶ Estimated value of each state
- ▶ Final estimate is randomized

## Verifying Convergence for TD(o)

Use proof rules to verify upper-bound on  $rpe$ :

$$rpe(TD(0), dist(V, V')) \leq (1 - \alpha + \alpha \cdot \gamma)^N \cdot dist(V, V')$$

## Verifying Convergence for TD(o)

Use proof rules to verify upper-bound on  $rpe$ :

$$rpe(TD(0), dist(V, V')) \leq (1 - \alpha + \alpha \cdot \gamma)^N \cdot dist(V, V')$$

Combine with soundness theorem:

$$Kant(dist)(TD(0)(V), TD(0)(V')) \leq (1 - \alpha + \alpha \cdot \gamma)^N \cdot dist(V, V')$$

## Verifying Convergence for TD(o)

Use proof rules to verify upper-bound on  $rpe$ :

$$rpe(TD(0), dist(V, V')) \leq (1 - \alpha + \alpha \cdot \gamma)^N \cdot dist(V, V')$$

Combine with soundness theorem:

$$Kant(dist)(TD(0)(V), TD(0)(V')) \leq (1 - \alpha + \alpha \cdot \gamma)^N \cdot dist(V, V')$$

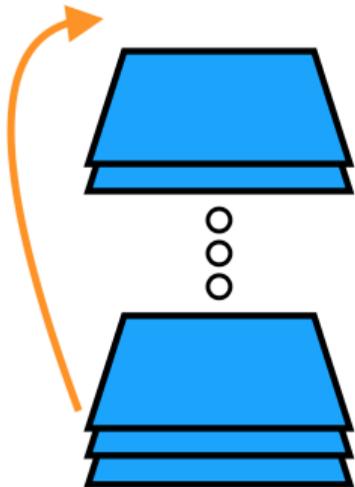
**Verified convergence for TD(o)!**

More Examples:

Algorithms for Shuffling Cards

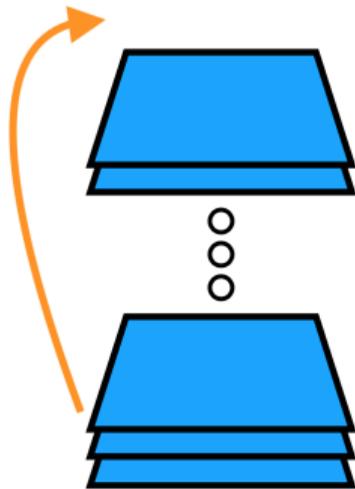
# Three simple models of card shuffling

Random-to-top

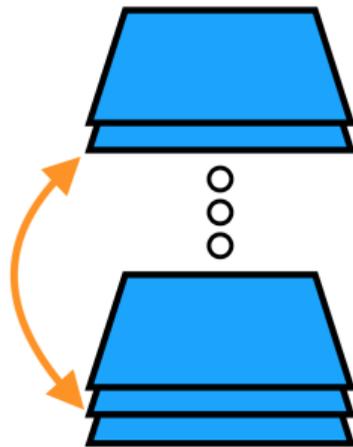


# Three simple models of card shuffling

Random-to-top

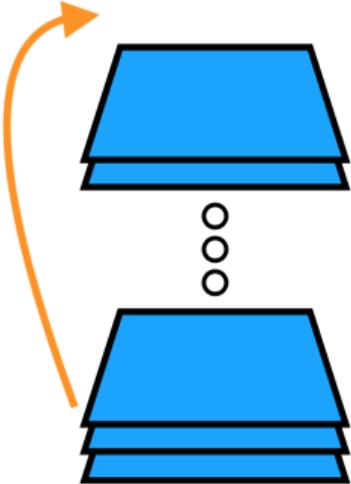


Random swap

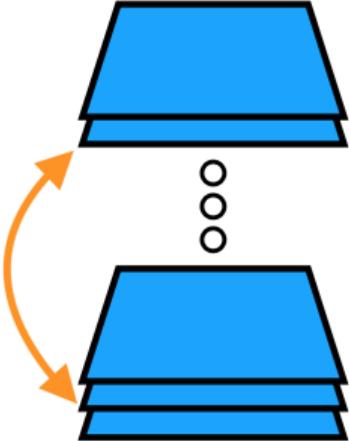


# Three simple models of card shuffling

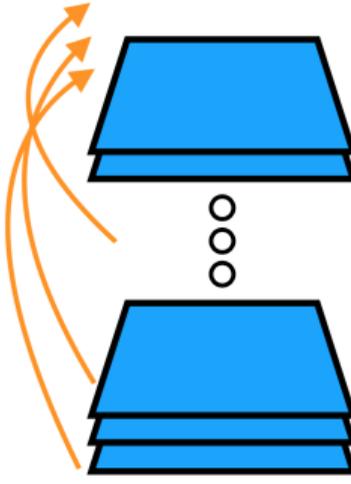
Random-to-top



Random swap

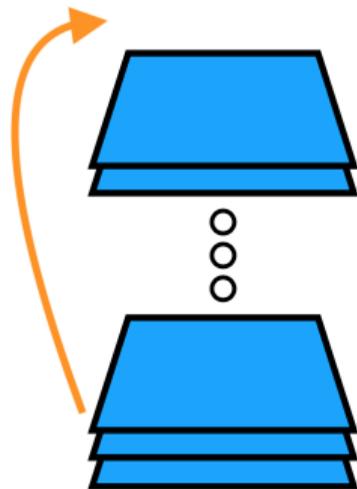


Riffle

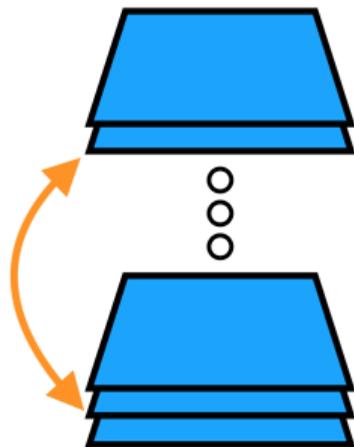


# Three simple models of card shuffling

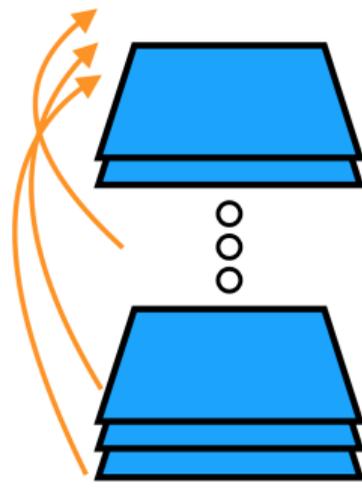
Random-to-top



Random swap



Riffle



Q: How **well mixed** are the cards after repeating  $K$  times?

## Verify different convergence rates

For a deck of  $N$  cards,  $K$  shuffling steps, and any two decks  $d_1, d_2$ :

$$TV(\llbracket \mathbf{rTop} \rrbracket(d_1, N, K), \llbracket \mathbf{rTop} \rrbracket(d_2, N, K)) \leq N \left( \frac{N-1}{N} \right)^K$$

$$TV(\llbracket \mathbf{rTrans} \rrbracket(d_1, N, K), \llbracket \mathbf{rTrans} \rrbracket(d_2, N, K)) \leq N \left( 1 - \frac{1}{N^2} \right)^K$$

$$TV(\llbracket \mathbf{riffle} \rrbracket(d_1, N, K), \llbracket \mathbf{riffle} \rrbracket(d_2, N, K)) \leq N^2 \left( \frac{1}{2} \right)^K$$

# Wrapping Up

# Plenty more in the paper!

## Verification details for each example

- ▶ Surprisingly familiar: loop invariants, push back through assignments, ...

## Connections between *rpe* and relational Hoare logics

- ▶ Embed core version of relational Hoare logic  $\mathbb{E}p\text{RHL}$  into *rpe*

## Other applications besides convergence

- ▶ Proving uniformity, lower bounds on distances, ...

# In summary

## Our work

- ▶ **Target:** sensitivity properties for probabilistic programs
- ▶ **Develop:** approach using relational pre-expectation transformers
- ▶ **Verify:** convergence for algorithms from ML, RL, probability theory

## Open questions

- ▶ How to prove sharper, more precise bounds on distances?
- ▶ How to automate the verification process?

# A Pre-Expectation Calculus for Probabilistic Sensitivity

Alejandro Aguirre, Gilles Barthe,  
Justin Hsu\*, Benjamin Kaminski,  
Joost-Pieter Katoen, Christoph Matheja

Originally appeared at POPL 2021