Abstraction in Reinforcement Learning

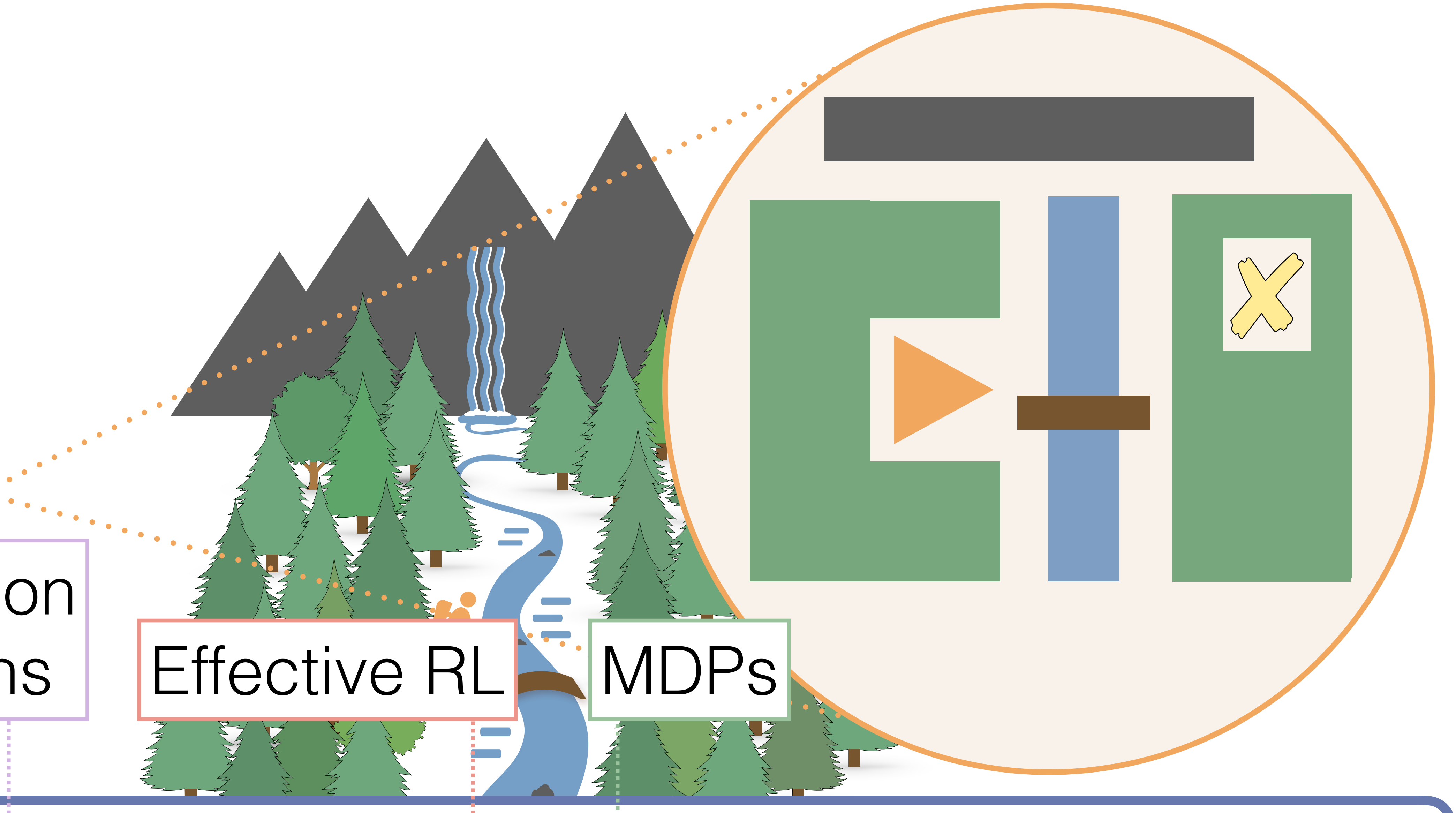
David Abel

Purdue
July 14, 2021

Dissertation: david-abel.github.io/thesis.pdf

Contact: dmabel@deepmind.com

Abstraction



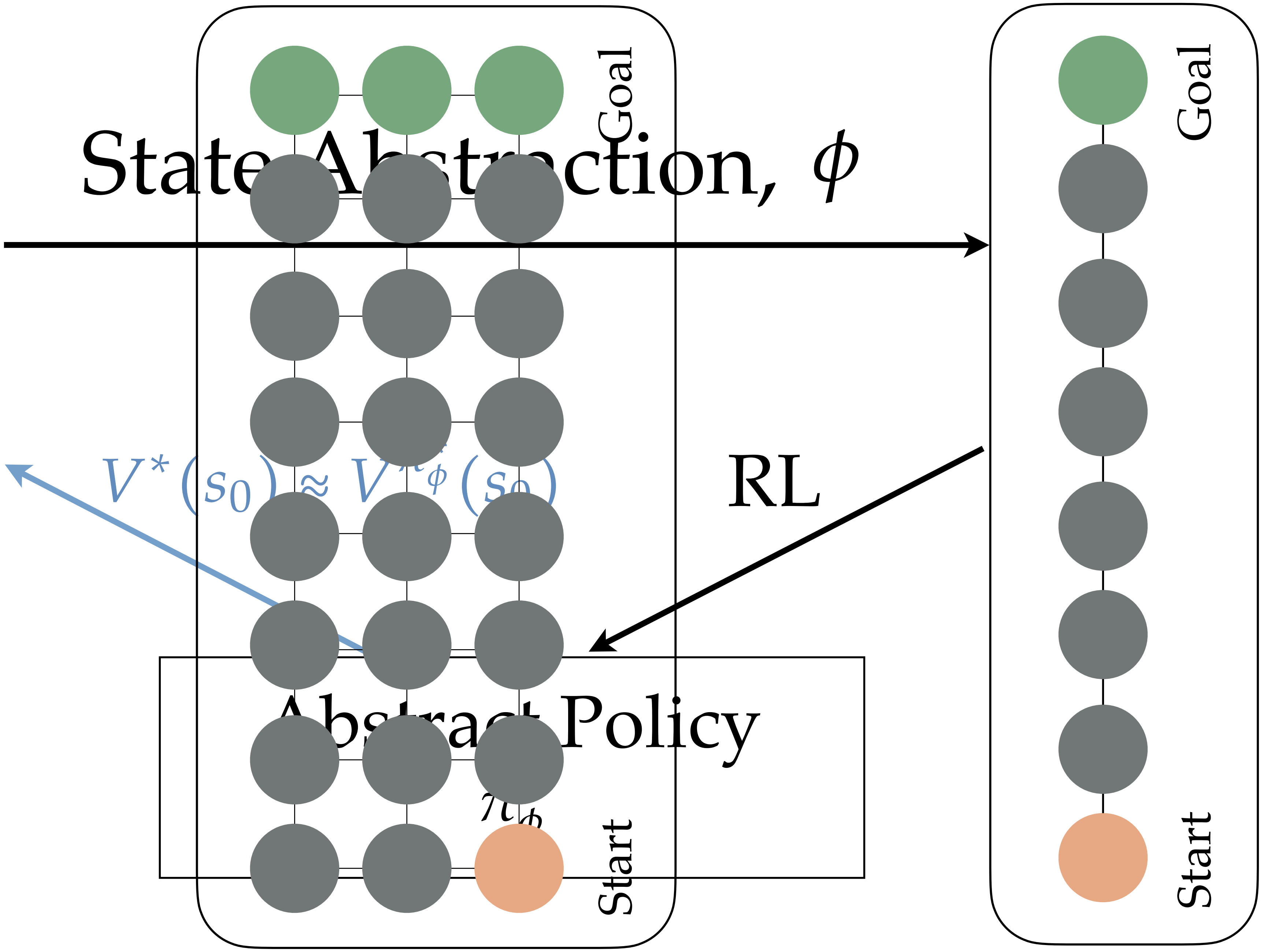
State & Action Abstractions

Effective RL

MDPs

Question: *How do intelligent agents come up with the right abstract understanding of the worlds they inhabit?*

State Abstraction



State Abstraction, ϕ

RL

$V^*(s_0) \approx V^{\pi_\phi}(s_0)$

Abstract Policy

π_ϕ

State Abstraction

Definition. A state abstraction is a function $\phi : \mathcal{S} \rightarrow \mathcal{S}_\phi$ that maps every ground state to an abstract state.

[Fox ’73]

[Whitt ’78]

[Singh et al. ’95]

[Dean, Givan ’97]

[Dieterrich ’00]

[Andre, Russell ’02]

[Ravindran, Barto ’03, ’04]

[Jong, Stone ’05]

[Ferns et al., ’04, ’06]

[Li et. al ‘06]

[Whiteson et al.’07]

[Castro, Precup ’09]

[Mugan, Kuipers ’12]

[Ortner et al. ’07, ’14, ’19]

[Hutter ’14, ’16, ‘19]

[Jiang et al., ‘’14, 15]

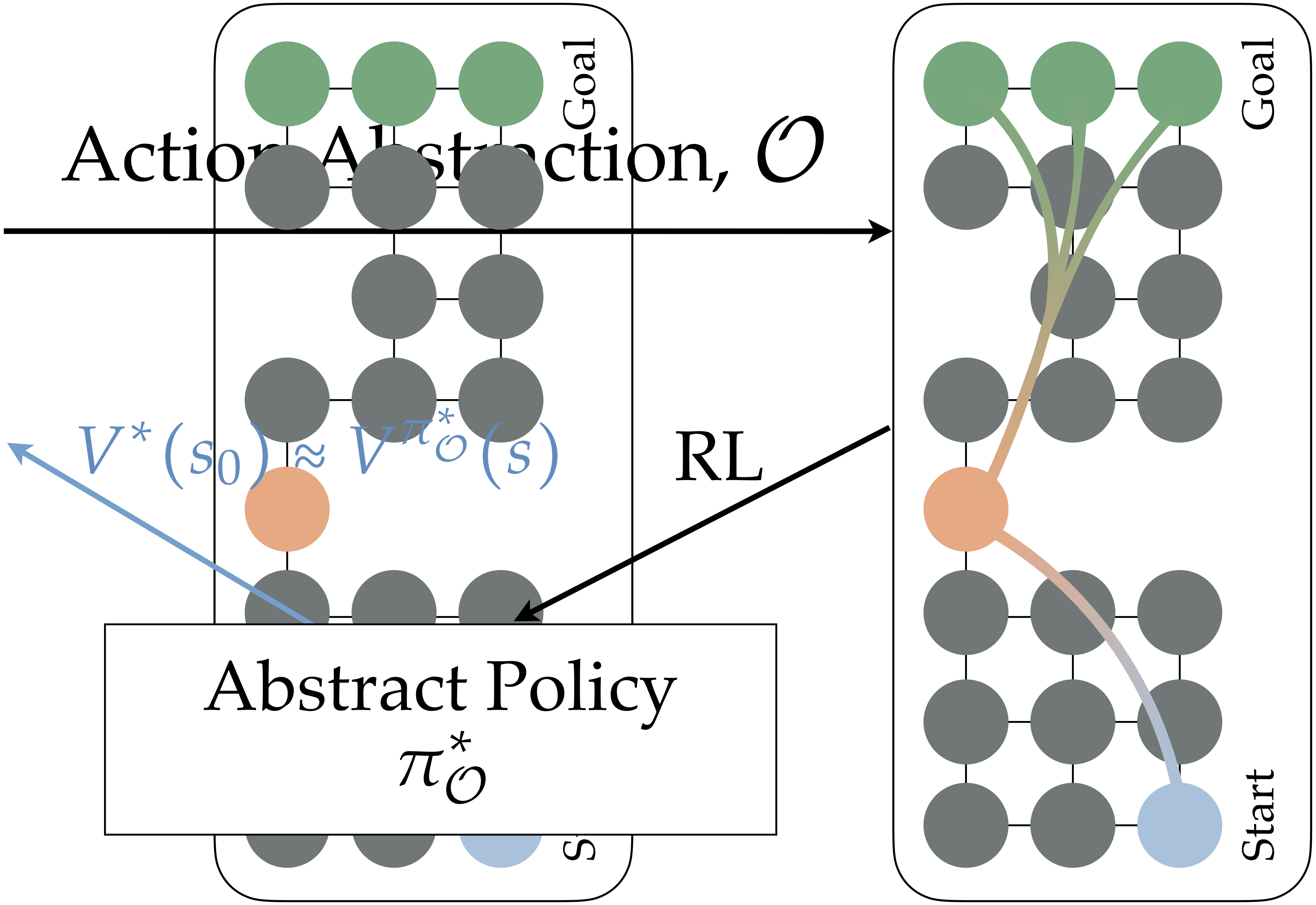
[Akrour et al., ’18]

[Menashe, Stone ’18]

[Taïga et al. ’18]

[Hostetler et al. ’14, ’15, ’17]

Action Abstraction



Action Abstraction, $\mathcal{O}$

RL

Abstract Policy
$\pi^*_{\mathcal{O}}$

$V^*(s_0) \approx V^{\pi^*_{\mathcal{O}}}(s)$

Action Abstraction

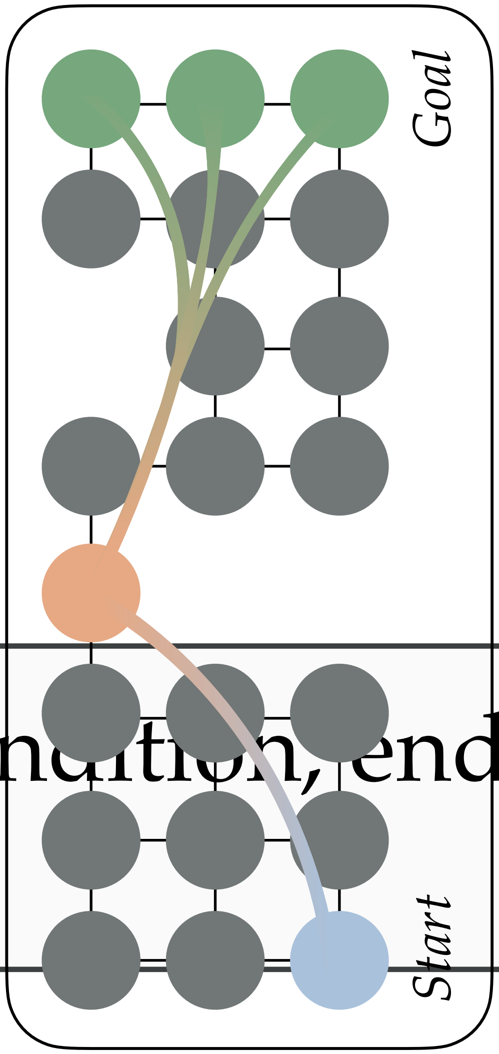
Example:

$o_1 = ($ 🔵, 🟠, $\pi_1)$

[Sutton, Precup, Singh 1999]

$o_2 = ($ 🟠, 🟢, $\pi_2)$

Definition (Option): A start condition, end condition, and a policy.



Action Abstraction

Definition (Action Abstraction): An action abstraction replaces the primitive actions with the option set $\mathcal{O}$.

[McGovern et. al. ‘97]

[Sutton, Precup, Singh ‘99]

[Simsek, Barto, ‘05, ’08]

[Jong, Hester, Stone ‘08]

[Brunskill, Li ’14]

[Ciosek, Silver ’15]

[Konidaris et al. ’06, ’07, ’09, ’10, ’18]

[Bacon et al. ’17, ’18]

[Fruit et al. ’17, ’17]

[Machado et al. ’17]

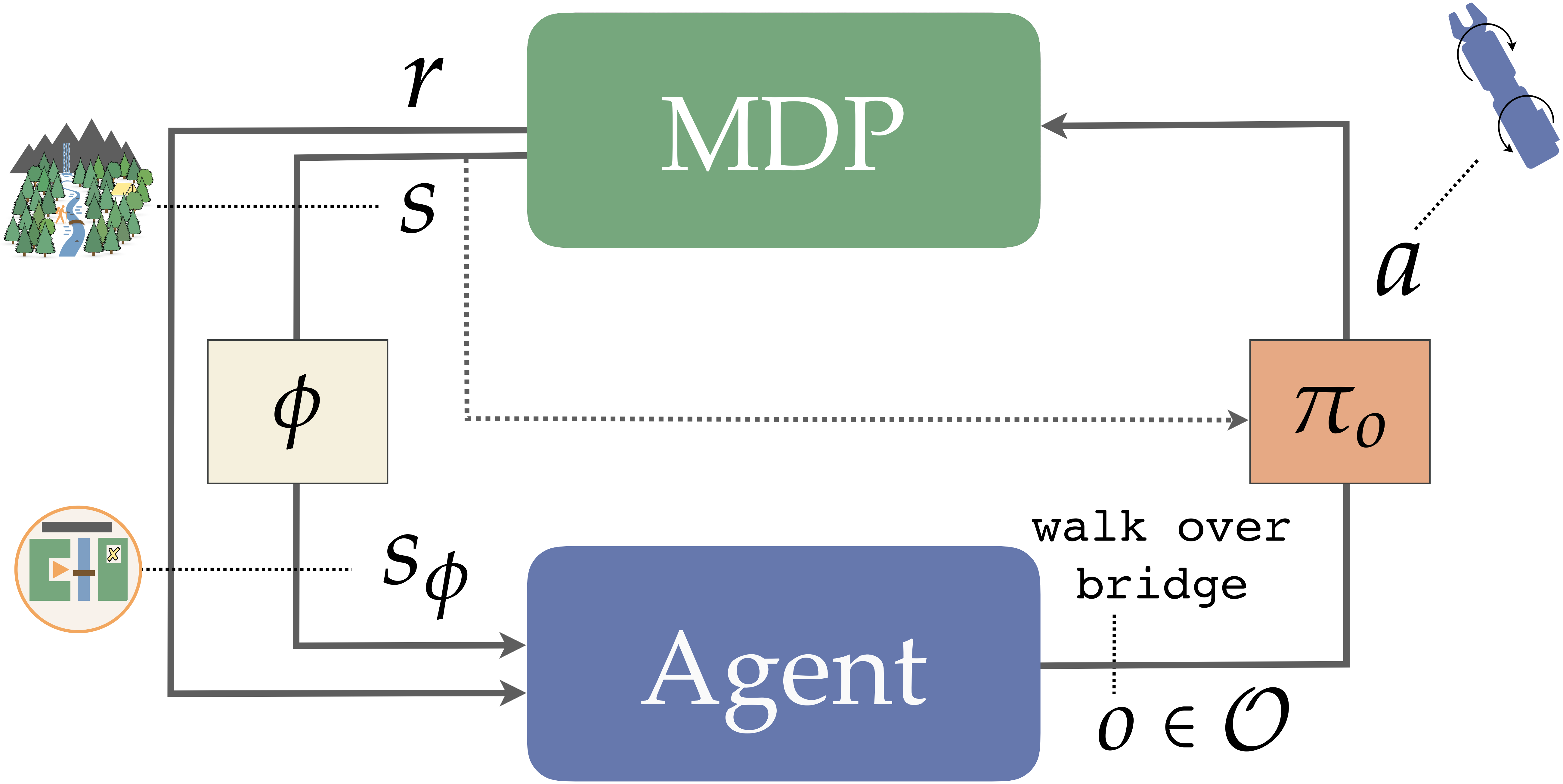
[Harutyunyan et al. ’18]

[Eysenbach et al. ‘18]

[Majeed & Hutter ’19]

…and more!

Abstraction in RL



r

MDP

s

a

ϕ

π_o

s_ϕ

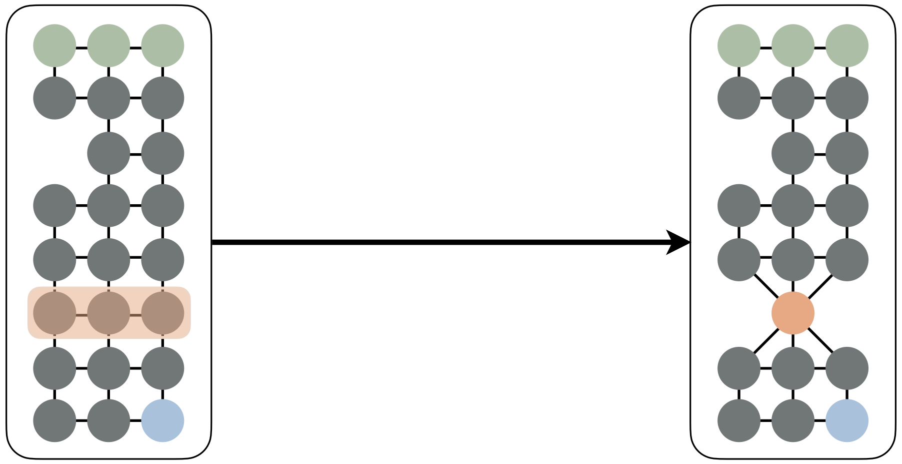
walk over bridge

Agent

$o \in \mathcal{O}$

Part 1

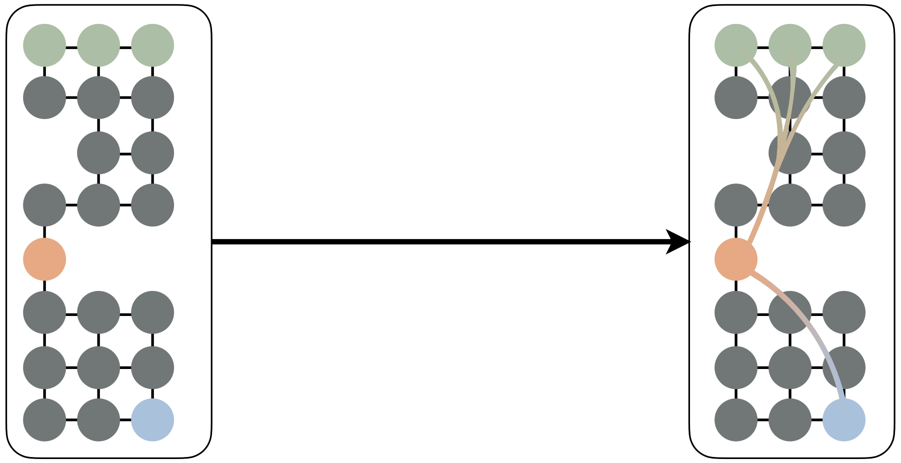
STATE ABSTRACTION



1. Approximate State Abstraction
   ICML 2016
2. State Abstraction In Lifelong RL
   ICML 2018
3. **State Abstraction As Compression**
   AAAI 2019

Part 2

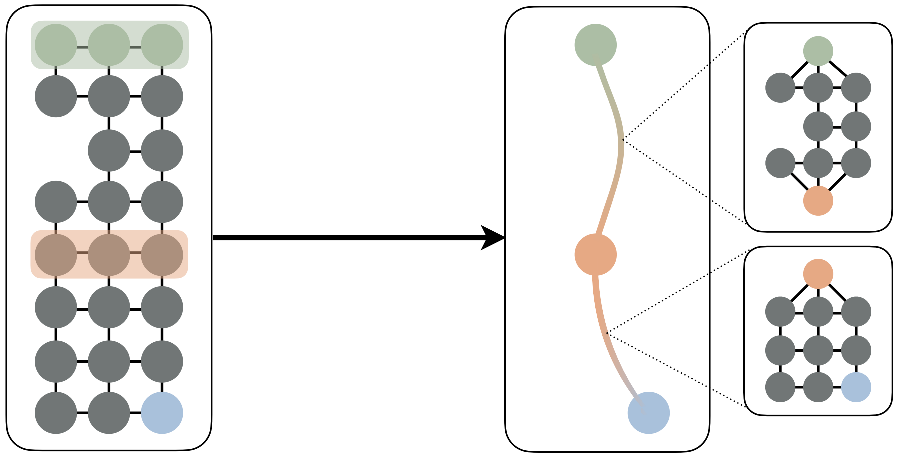
ACTION ABSTRACTION



4. **Options for Planning**
   ICML 2019
5. Options for Exploration
   ICML 2019
6. A New Option Model
   IJCAI 2019

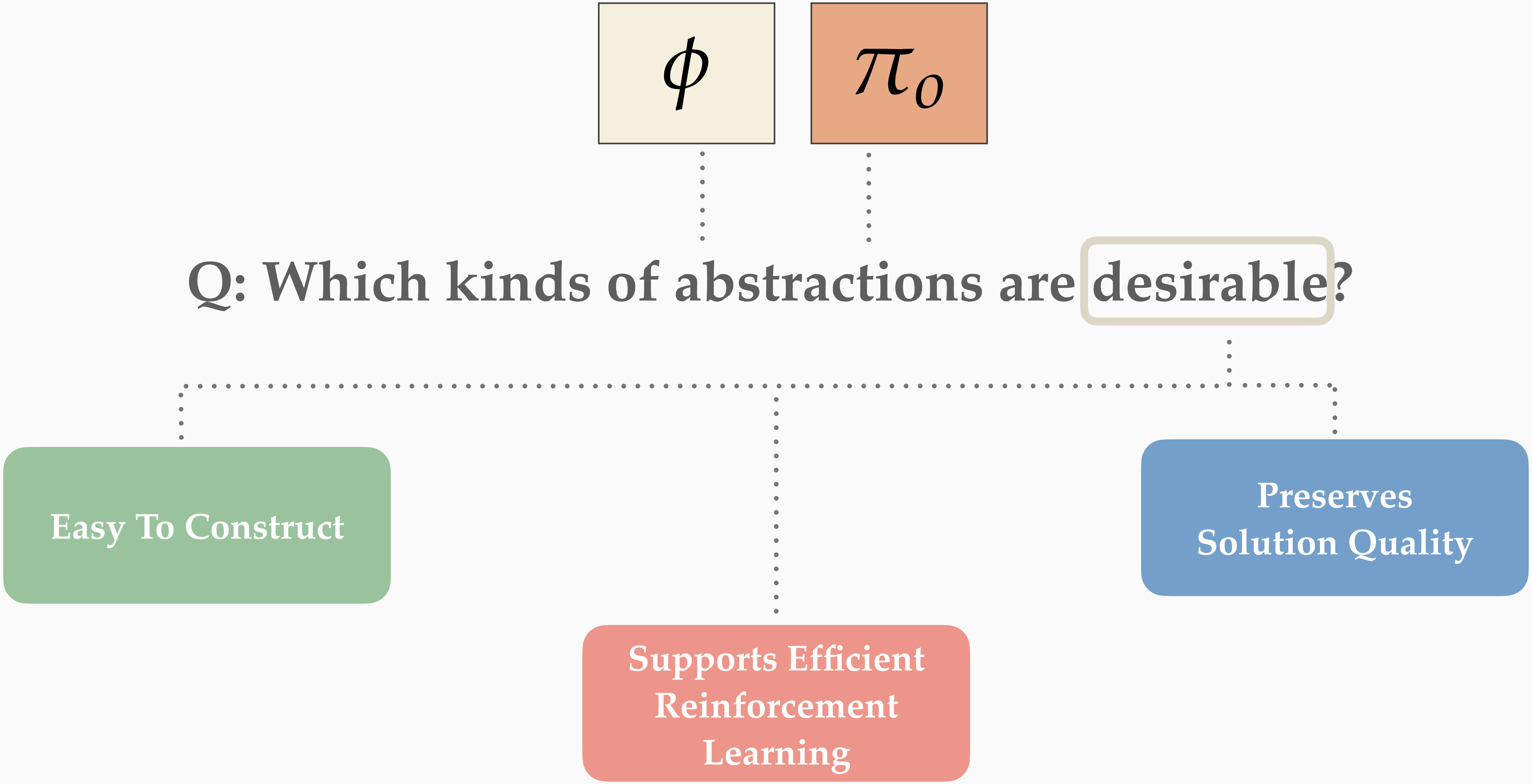
Part 3

STATE-ACTION ABSTRACTION



7. **Value-Preserving Hierarchies**
   AISTATS 2020

Desirable Abstractions



ϕ

π_o

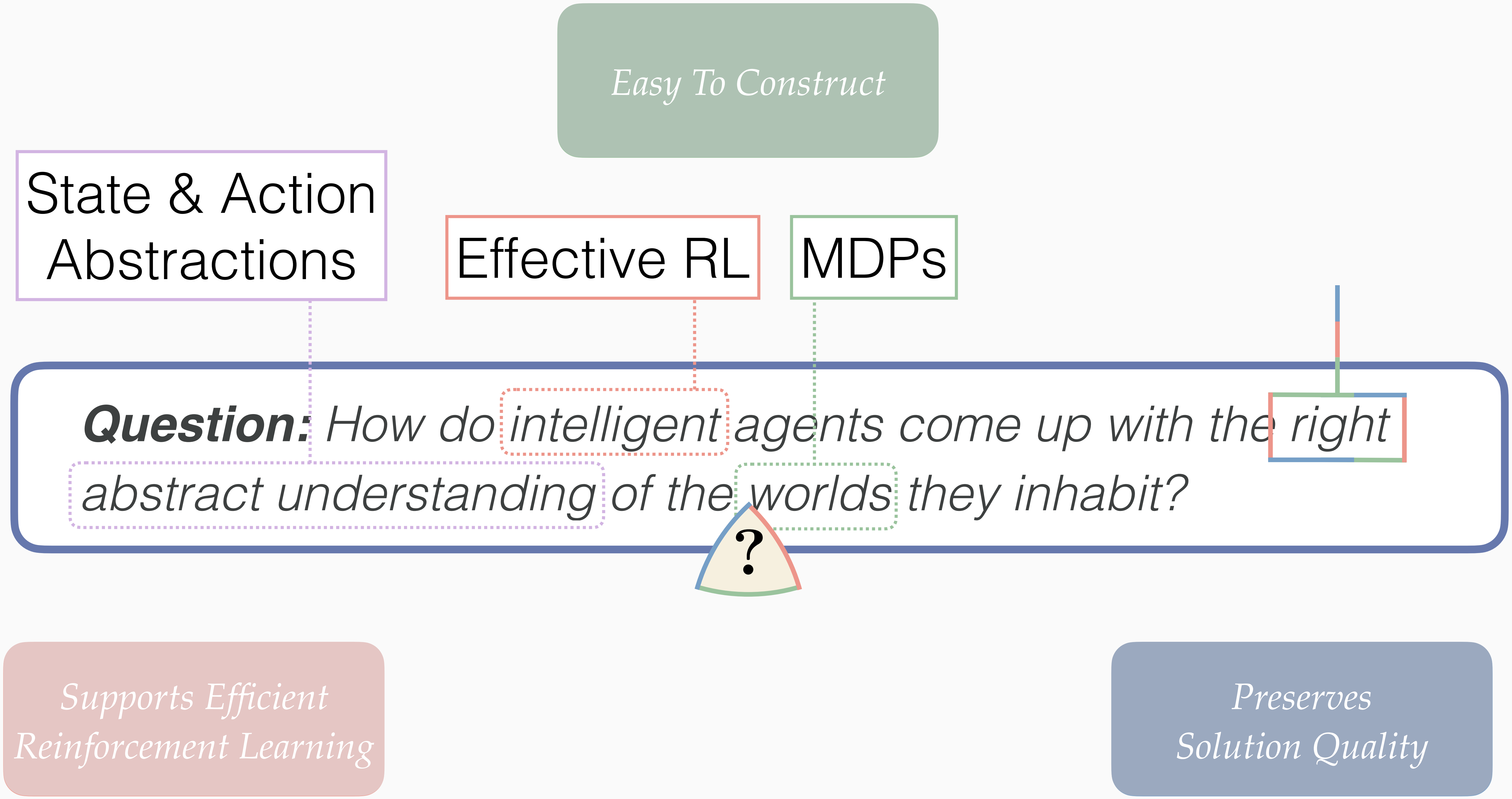
Q: Which kinds of abstractions are desirable?

Easy To Construct

Supports Efficient Reinforcement Learning

Preserves Solution Quality

Abstraction Desiderata



Easy To Construct

State & Action Abstractions

Effective RL

MDPs

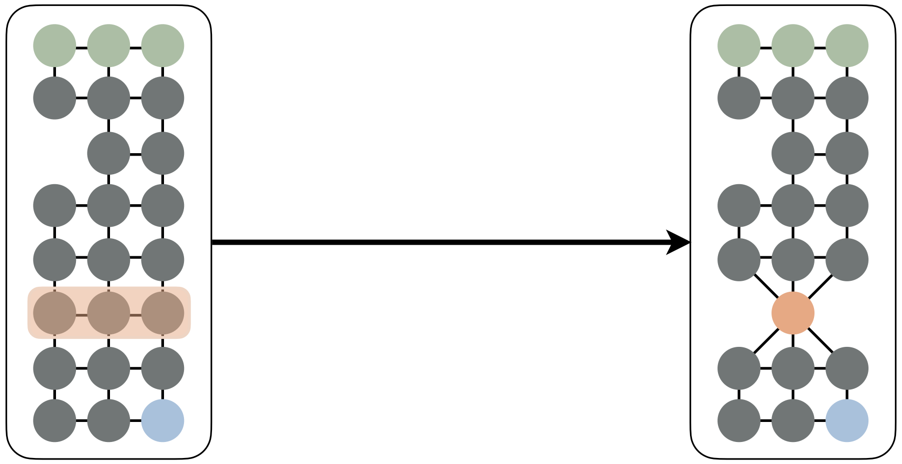
Question: *How do intelligent agents come up with the right abstract understanding of the worlds they inhabit?*

Supports Efficient Reinforcement Learning

Preserves Solution Quality

Part 1

STATE ABSTRACTION



1. Approximate State Abstraction
   ICML 2016
2. State Abstraction In Lifelong RL
   ICML 2018
3. **State Abstraction As Compression**
   AAAI 2019

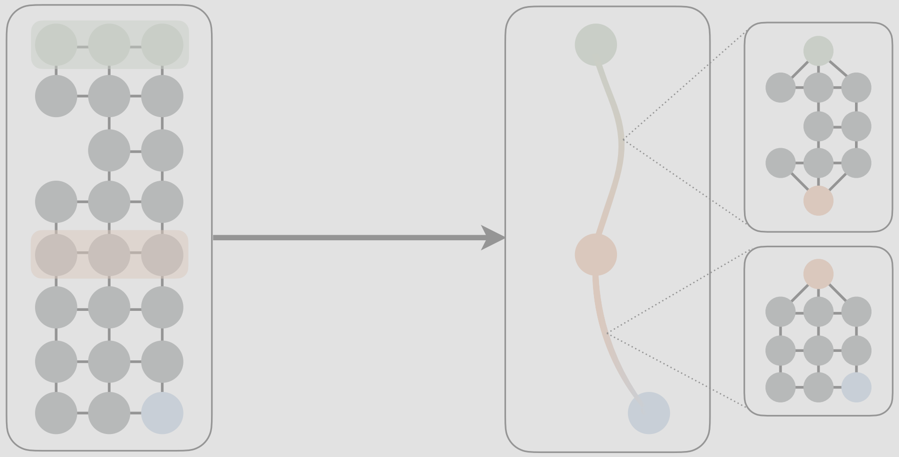
Part 2

ACTION ABSTRACTION

4. **Options for Planning**
   ICML 2019
5. Options for Exploration
   ICML 2019
6. A New Option Model
   IJCAI 2019

Part 3

STATE-ACTION ABSTRACTION

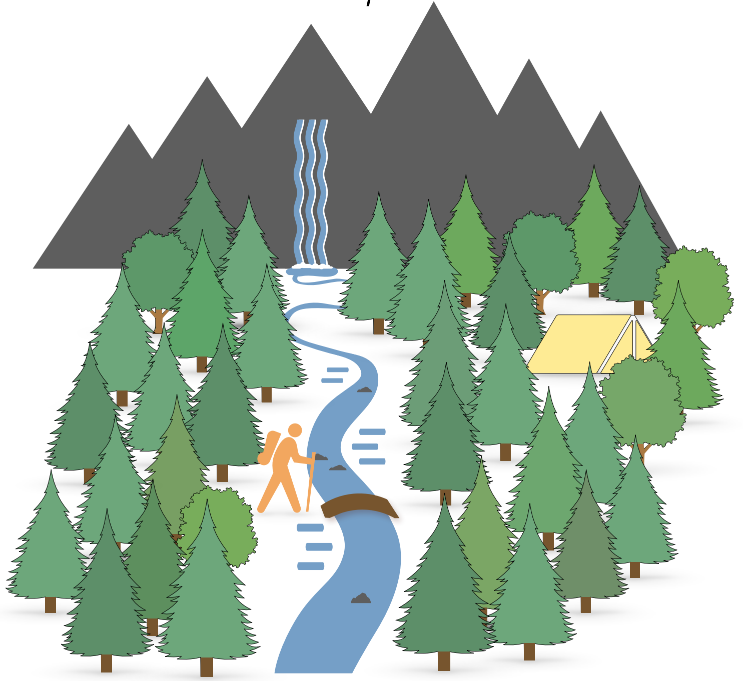


7. **Value-Preserving Hierarchies**
   AISTATS 2020

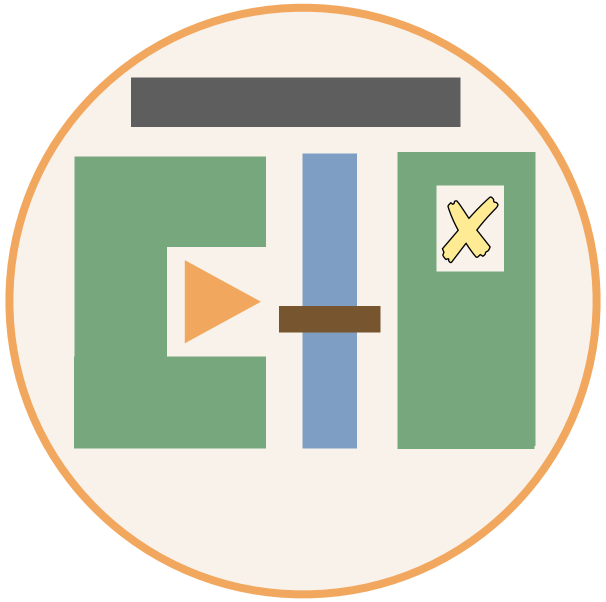
State Abstraction as Compression

*[**A**AAJWL, AAAI 2019]*

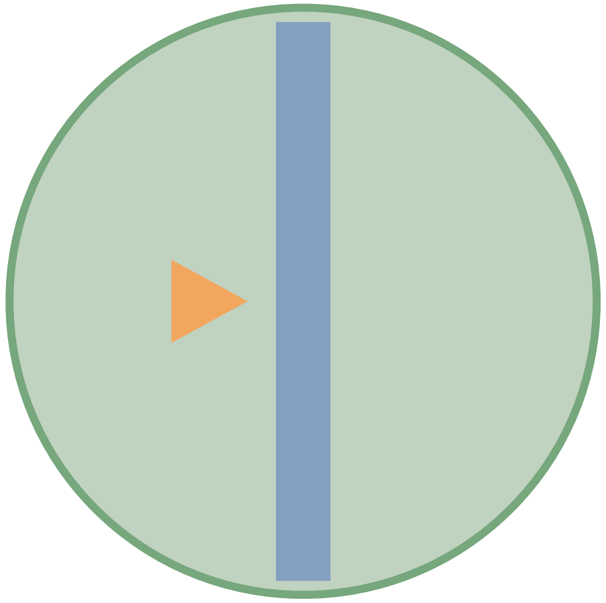
High Value
No Compression

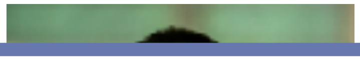
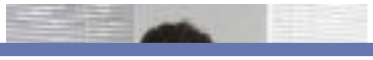


Some Value
Some Compression



No Value
High Compression



***Question:** How can we construct state abstractions that trade-off between compression and representational quality?*

Dilip
Arumugam

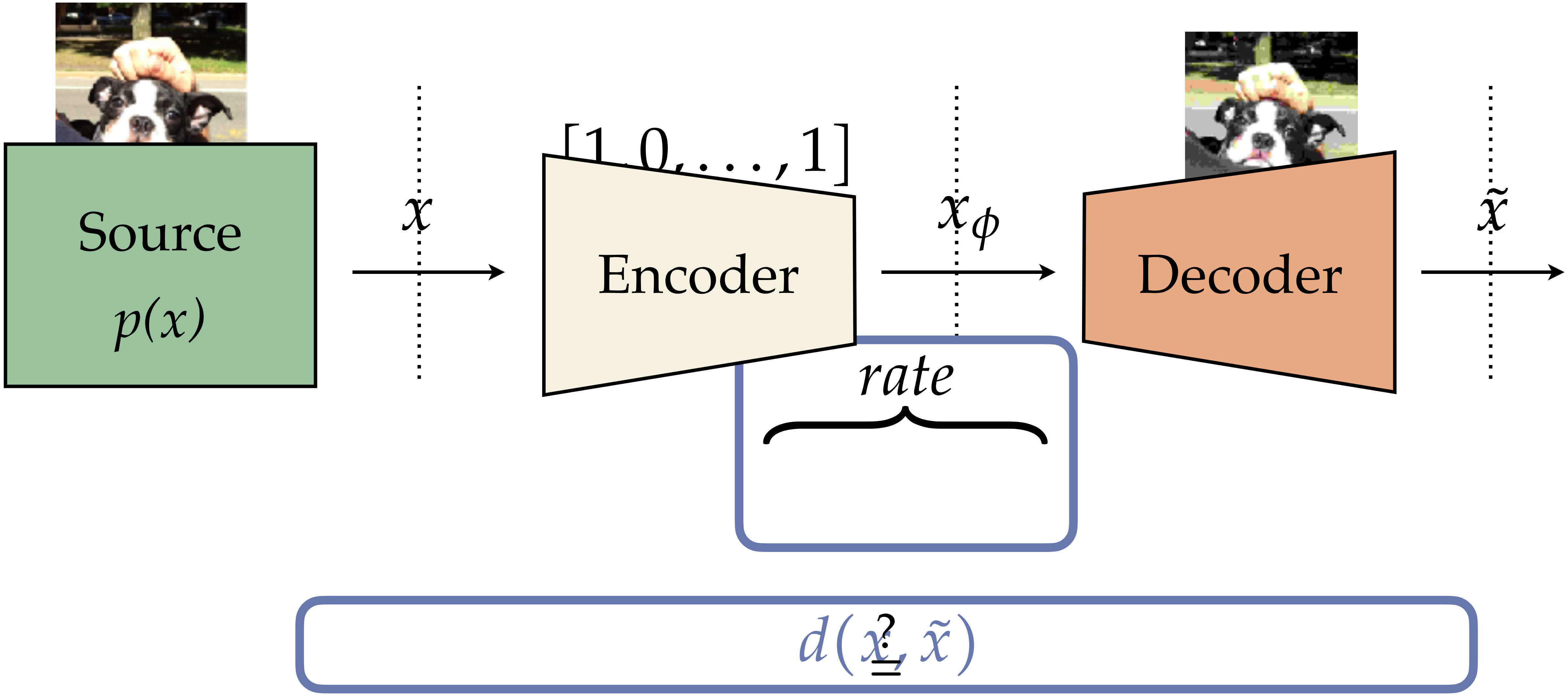
Kavosh
Asadi

Yuu
Jinnai

Lawson
L.S. Wong

Michael L.
Littman

Rate-Distortion Theory



Source

$p(x)$

x

Encoder

$[1,0,\ldots,1]$

x_ϕ

Decoder

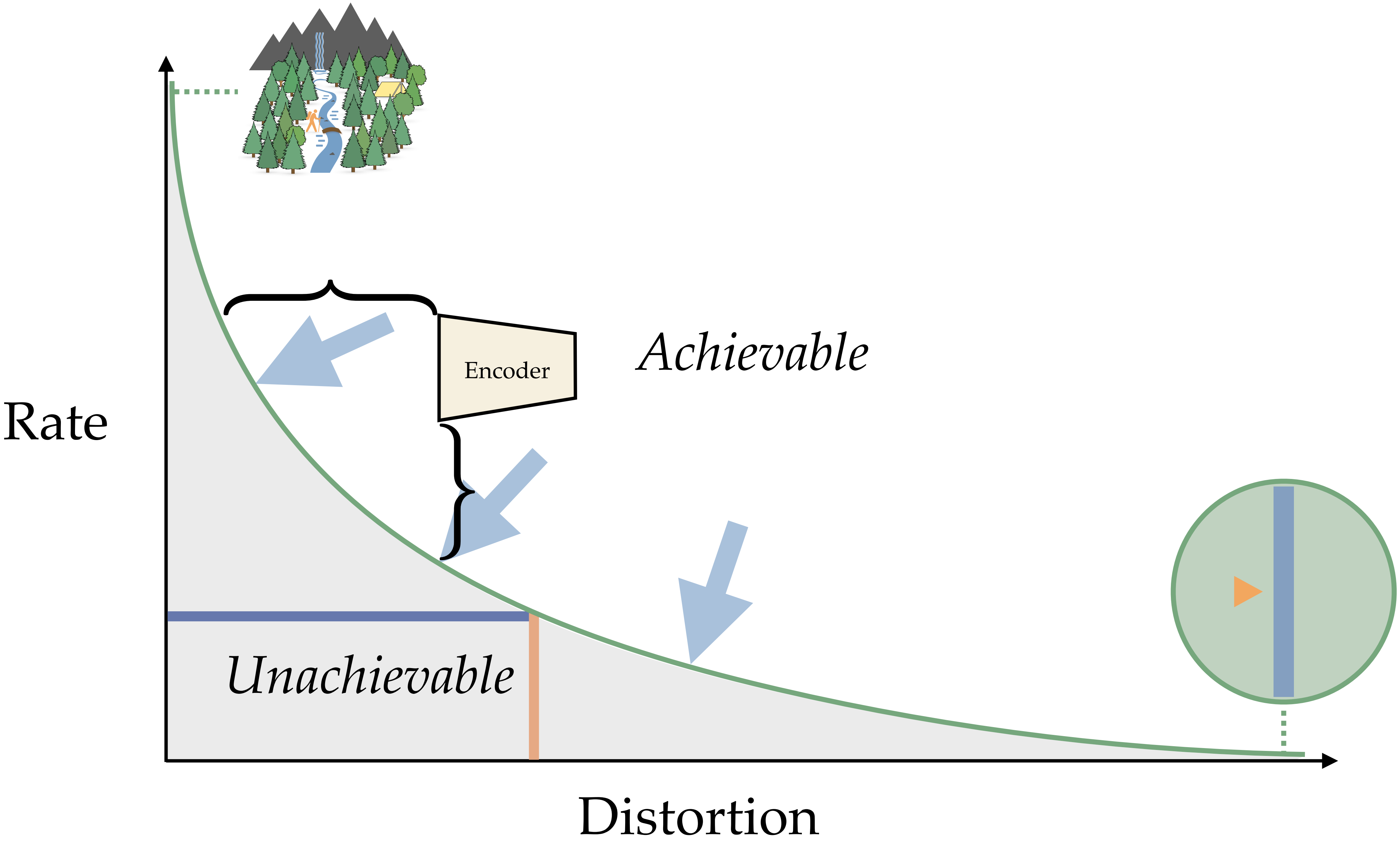
$\tilde{x}$

$\overbrace{}^{rate}$

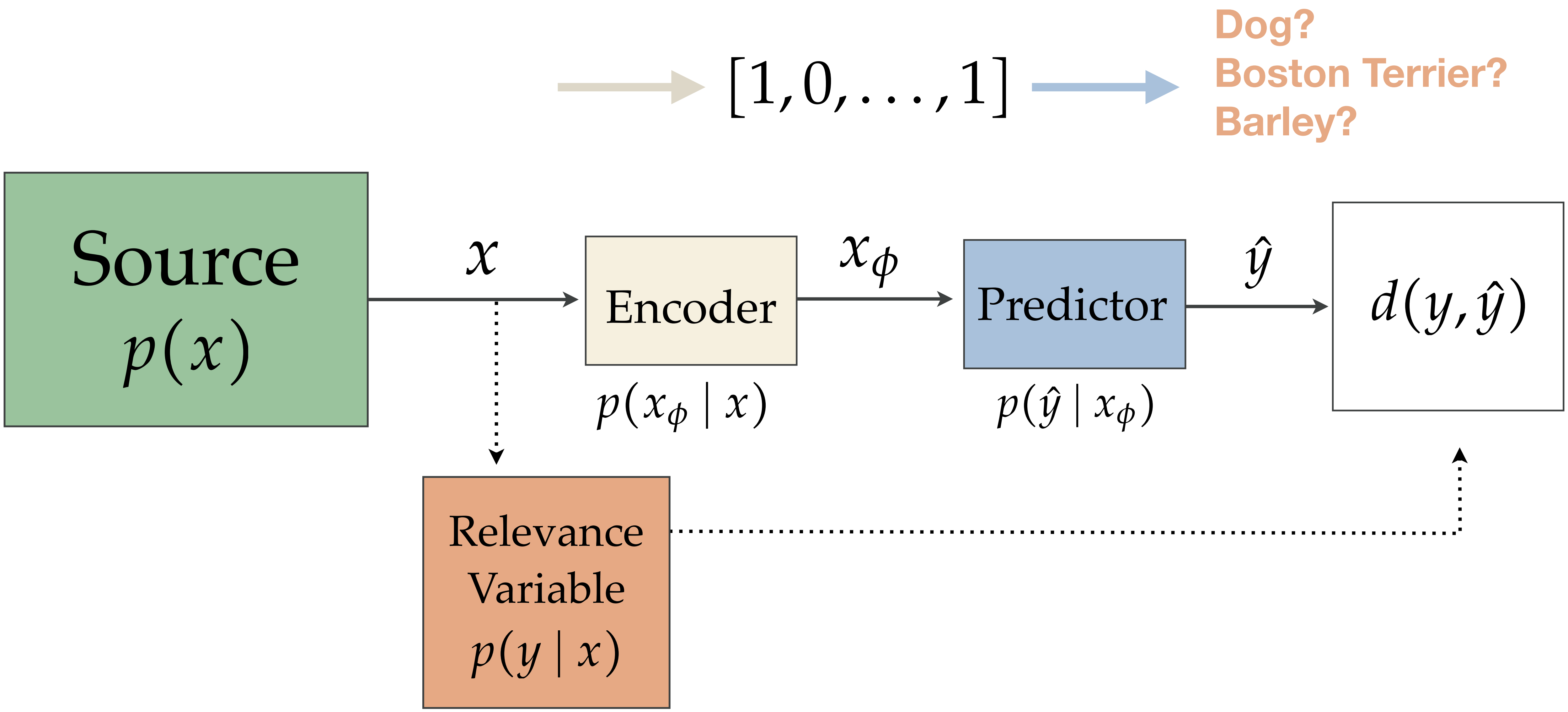
$d(x \overset{?}{=}, \tilde{x})$

[Shannon '48, Berger '03]

Rate-Distortion Theory



Information Bottleneck



$[1, 0, \ldots, 1]$

Dog?
Boston Terrier?
Barley?

Source
$p(x)$

x

Encoder
$p(x_\phi \mid x)$

x_ϕ

Predictor
$p(\hat{y} \mid x_\phi)$

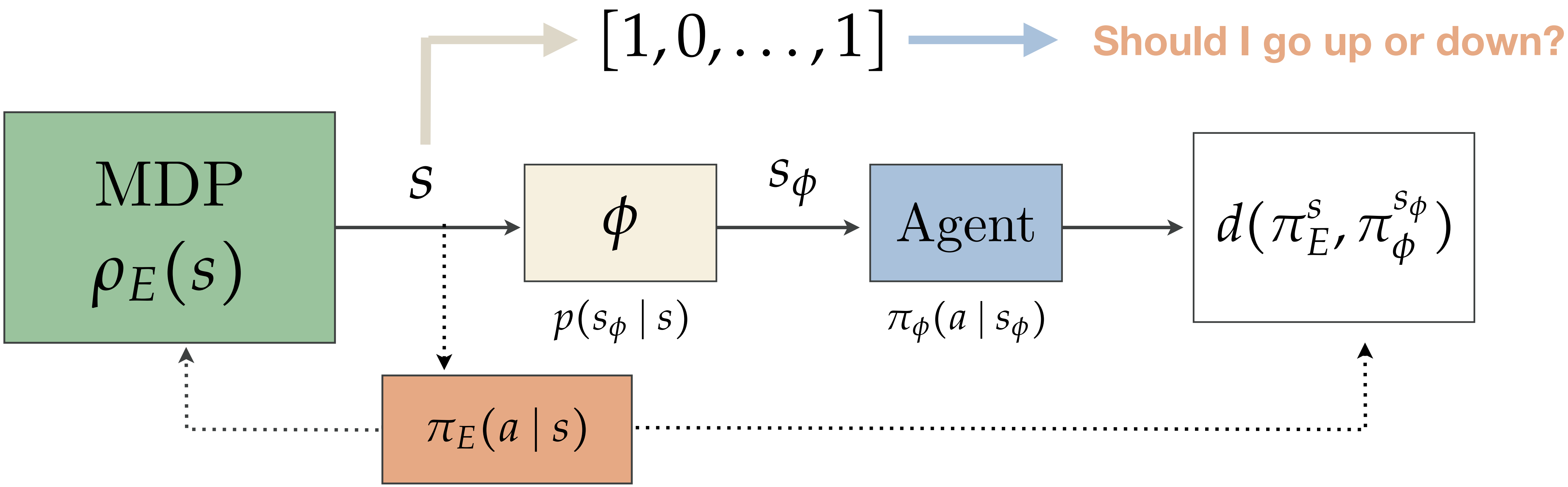
$\hat{y}$

$d(y, \hat{y})$

Relevance
Variable
$p(y \mid x)$

[Tishby, Pereira, Bialek ’99]

State Abstraction as Compression



$$\min_{\phi}\left(|\mathcal{S}_\phi| + \boxed{\beta}\, \mathbb{E}_{\rho_E(s)}\left[\underbrace{V^{\pi_E}(s) - V^{\pi_\phi^*}(\phi(s))}_{\textit{Value Loss}}\right]\right)$$

Our Objective

State Abstraction as Compression

Theorem.

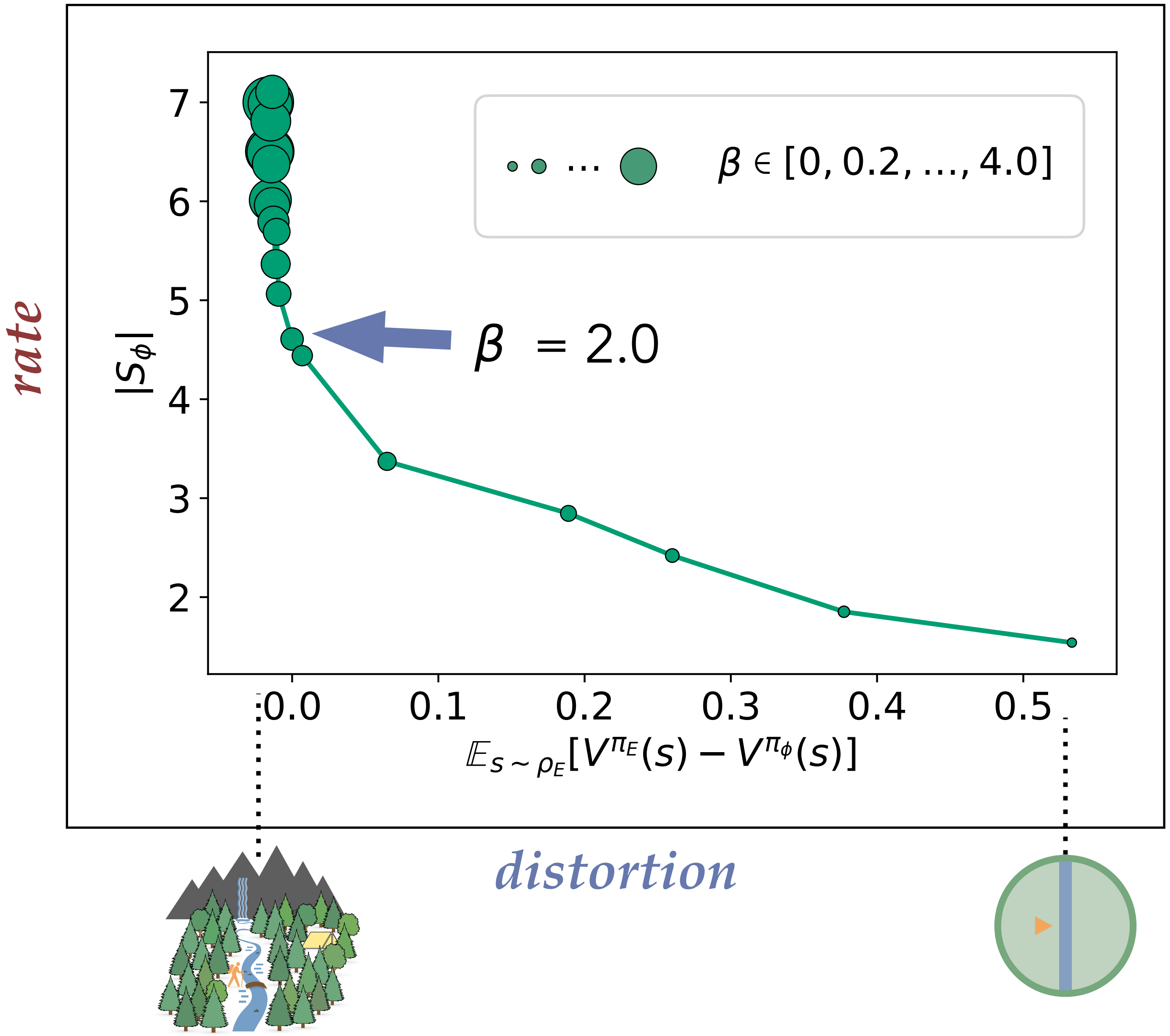
$$\min_{\phi} \left(|\mathcal{S}_{\phi}| + \beta \mathop{\mathbb{E}}_{\rho_E(s)} \left[V^{\pi_E}(s) - V^{\pi^*_{\phi}}(\phi(s)) \right] \right)$$

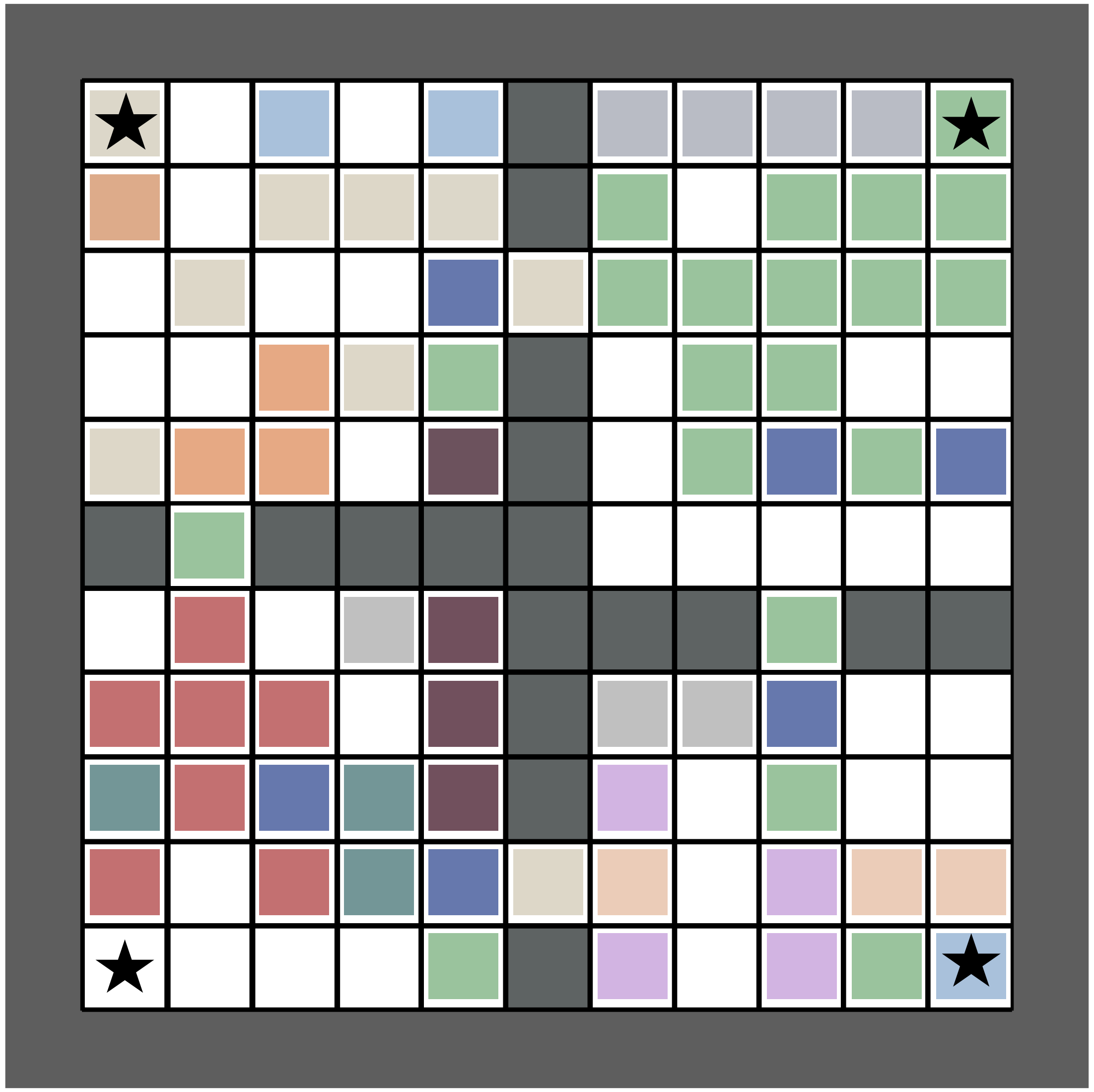
Our Objective $\leq$ *DIB Objective*

[Strouse & Schwab, '17]

$$\min_{\phi} \left(\frac{H(\rho_{\phi})}{\delta \log \frac{1}{\delta}} + 2\text{VM}\textsc{ax}\beta \mathop{\mathbb{E}}_{\rho_E(s)} \left[D_{\text{KL}}(\pi_E(s) \,\|\, \pi^*_{\phi}(\phi(s))) \right] \right)$$

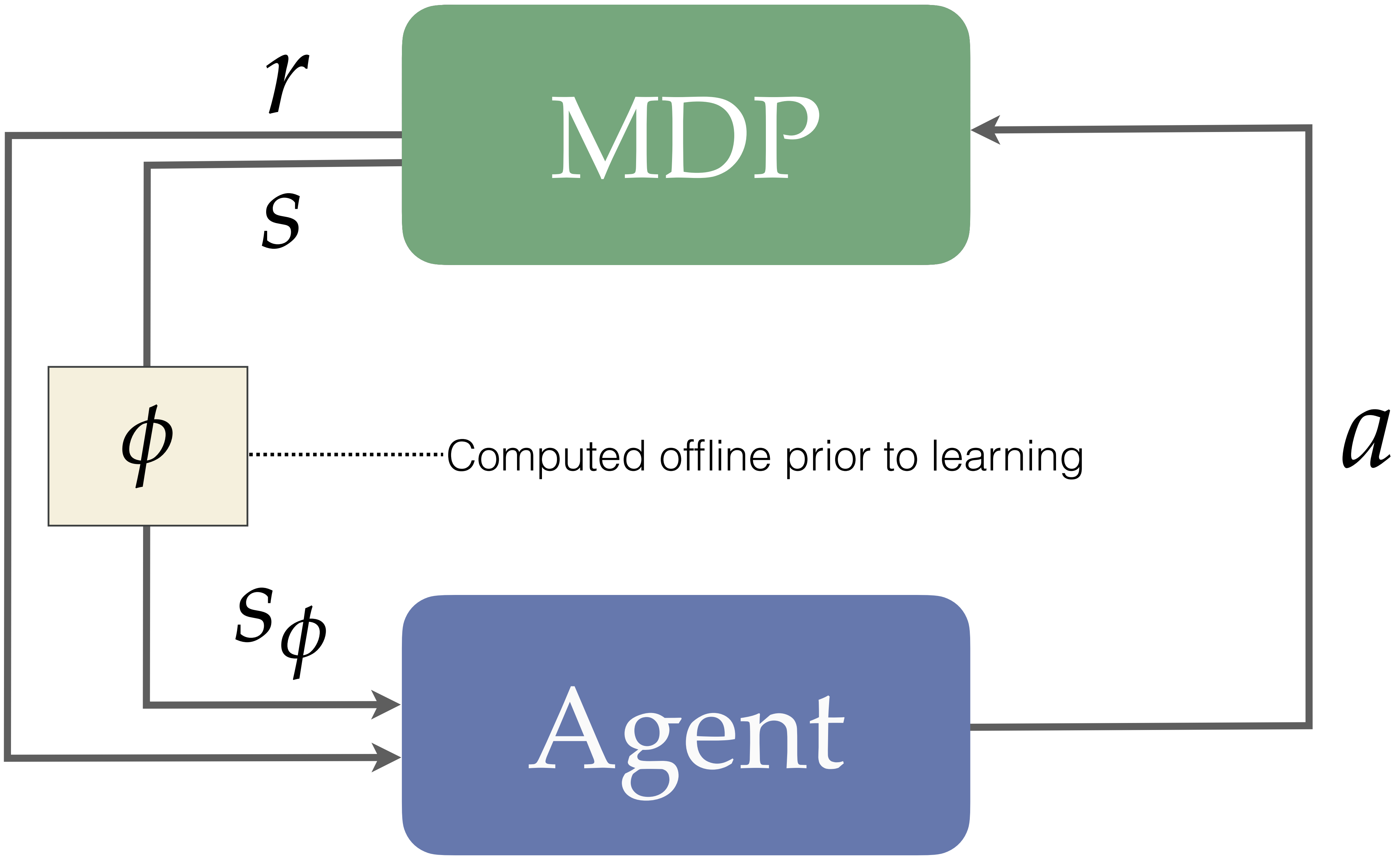
State Abstraction as Compression



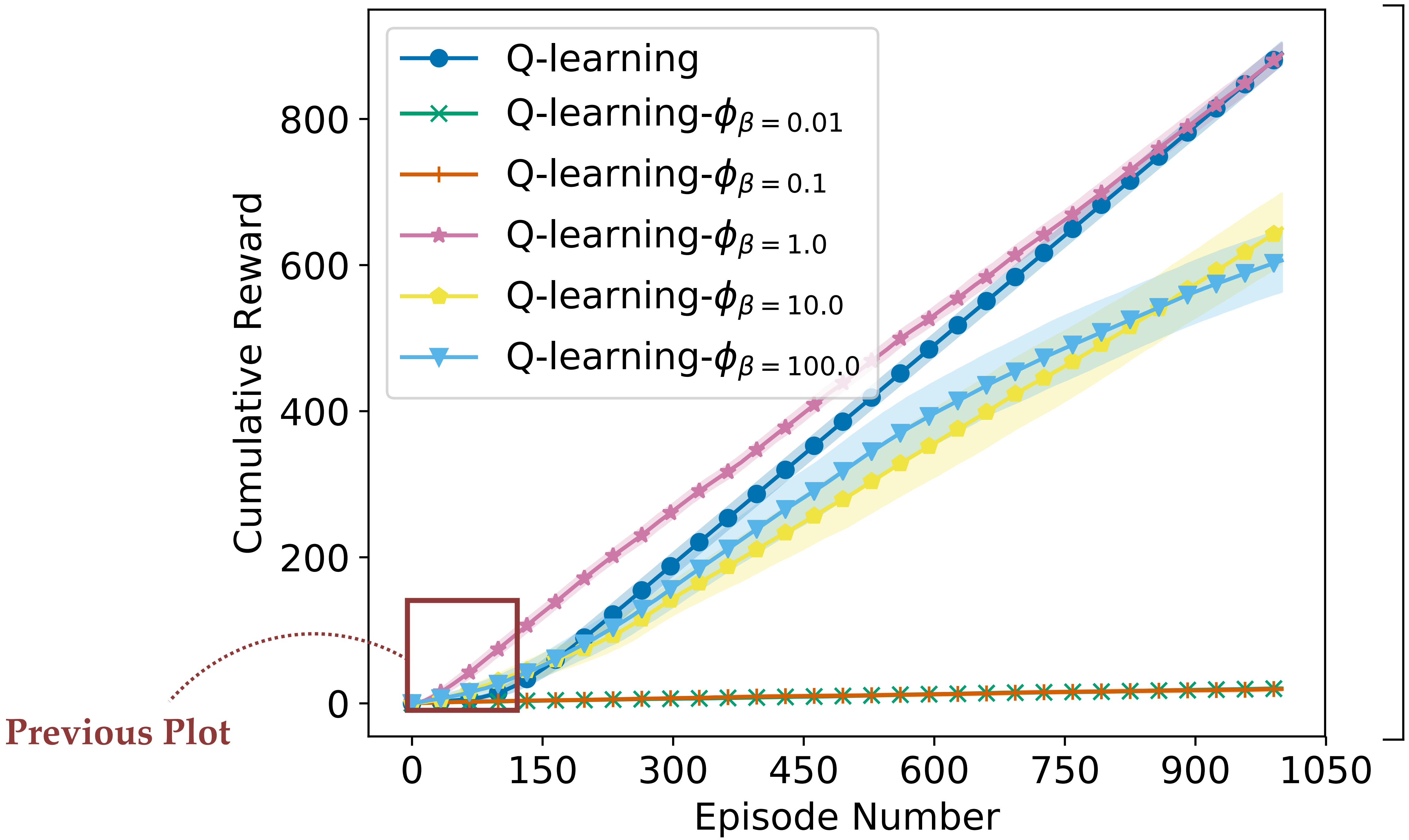


Multitask Abstraction

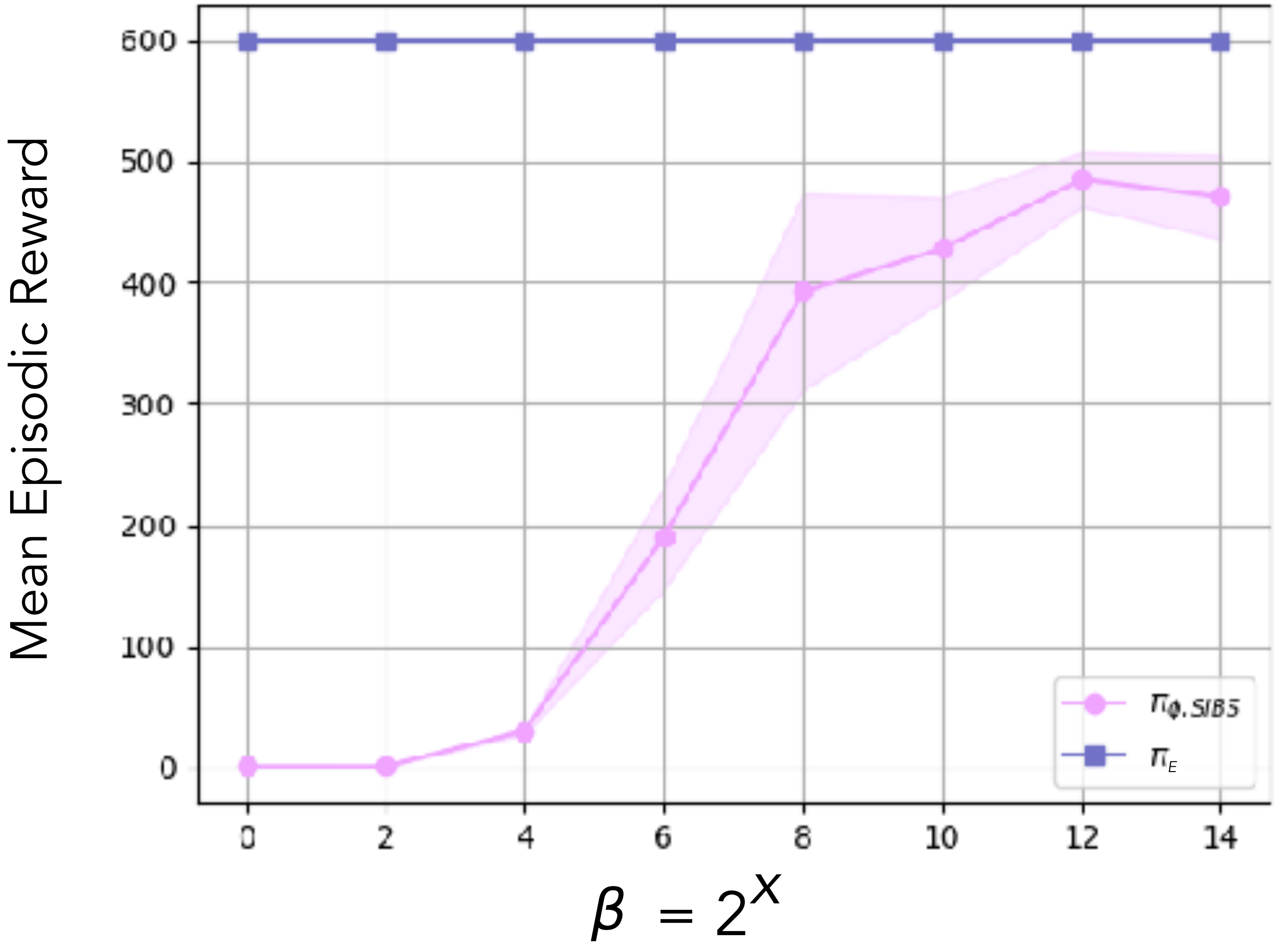
Learning Experiments

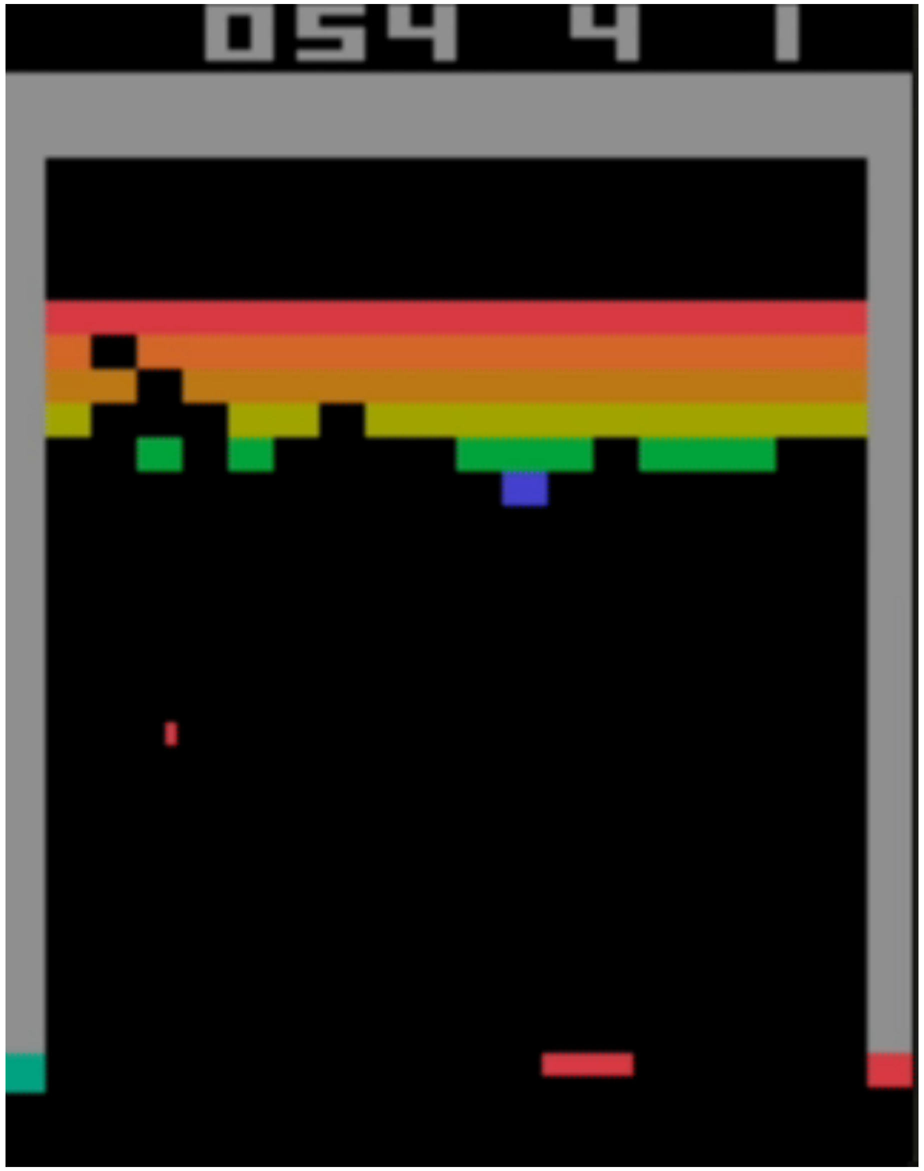


Experiments: Four Rooms



Experiments: Breakout





Extension: Continuous State

Theorem. *For any $\delta \in (0,1)$, n the size of the training data set, $\Delta \in \mathbb{R}$ training loss, and ρ a fixed distribution on states used to train $\tilde{\phi} \in \Phi$, with probability at least $1-\delta$:*

$$\mathop{\mathbb{E}}_{s\sim\rho}\left[\left\|\left(\pi^*(\cdot\,|\,s)-\pi_{\tilde{\phi}}(\cdot\,|\,s)\right)\right\|_1\right] \le \boxed{\frac{\Delta}{2}} + \boxed{2\sqrt{2}Rad(\Phi)} + \boxed{\sqrt{\frac{2\ln\frac{1}{\delta}}{n}}}$$

[Barlett, Mendelson '02]

training error

hypothesis class richness

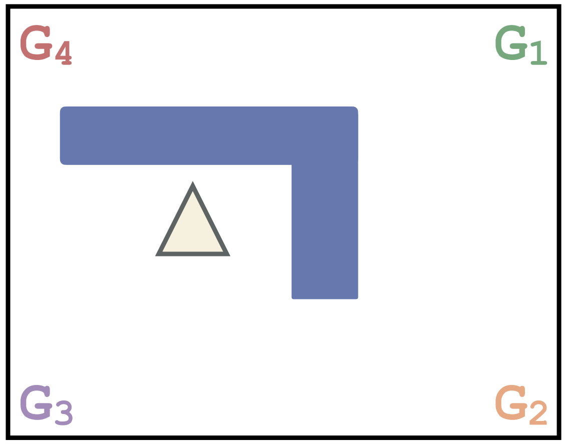
size of training data

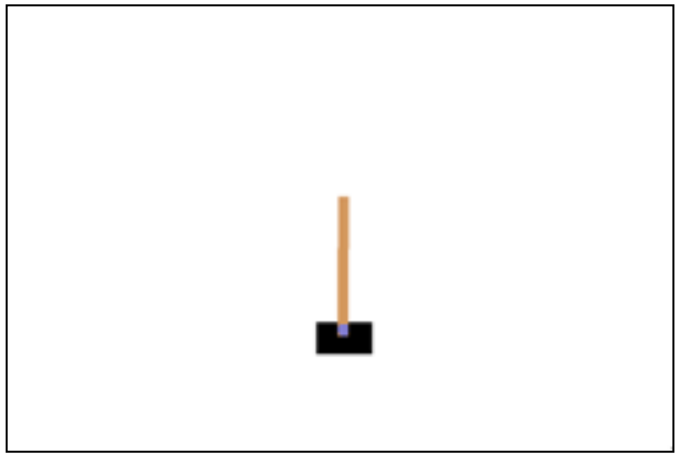
led project →

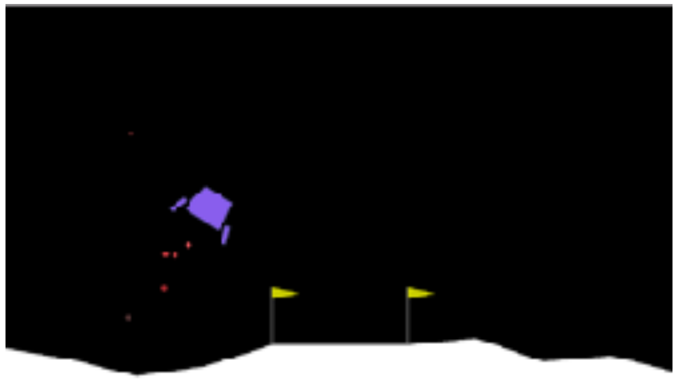
Kavosh Asadi

Michael L. Littman

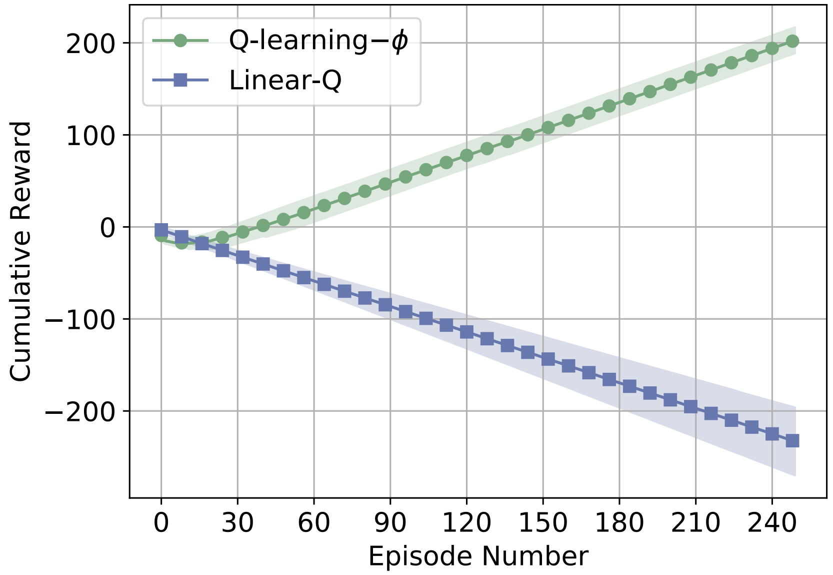
Extension: Continuous State



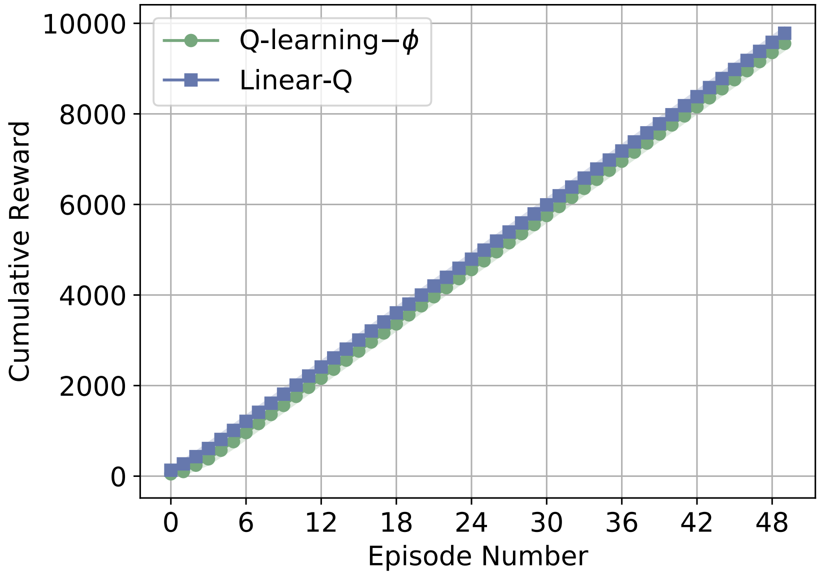




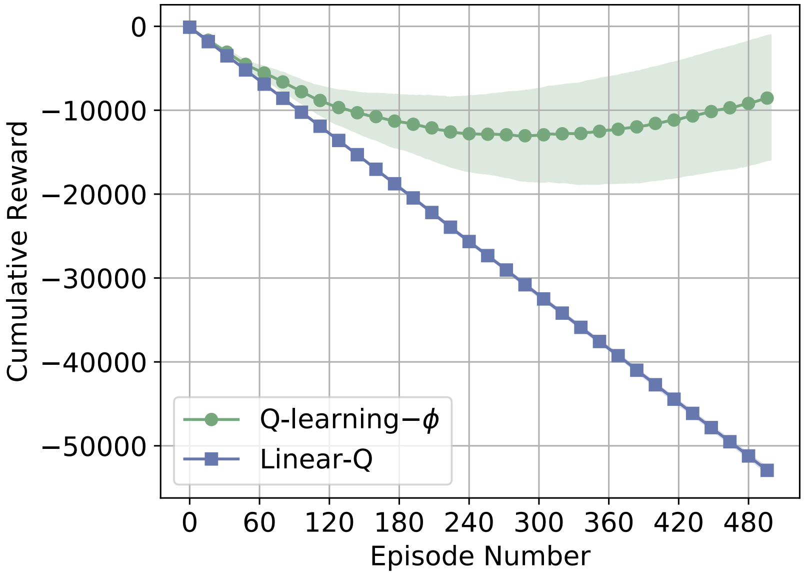
Cumulative Reward: Puddle H 1.0 W 1.0



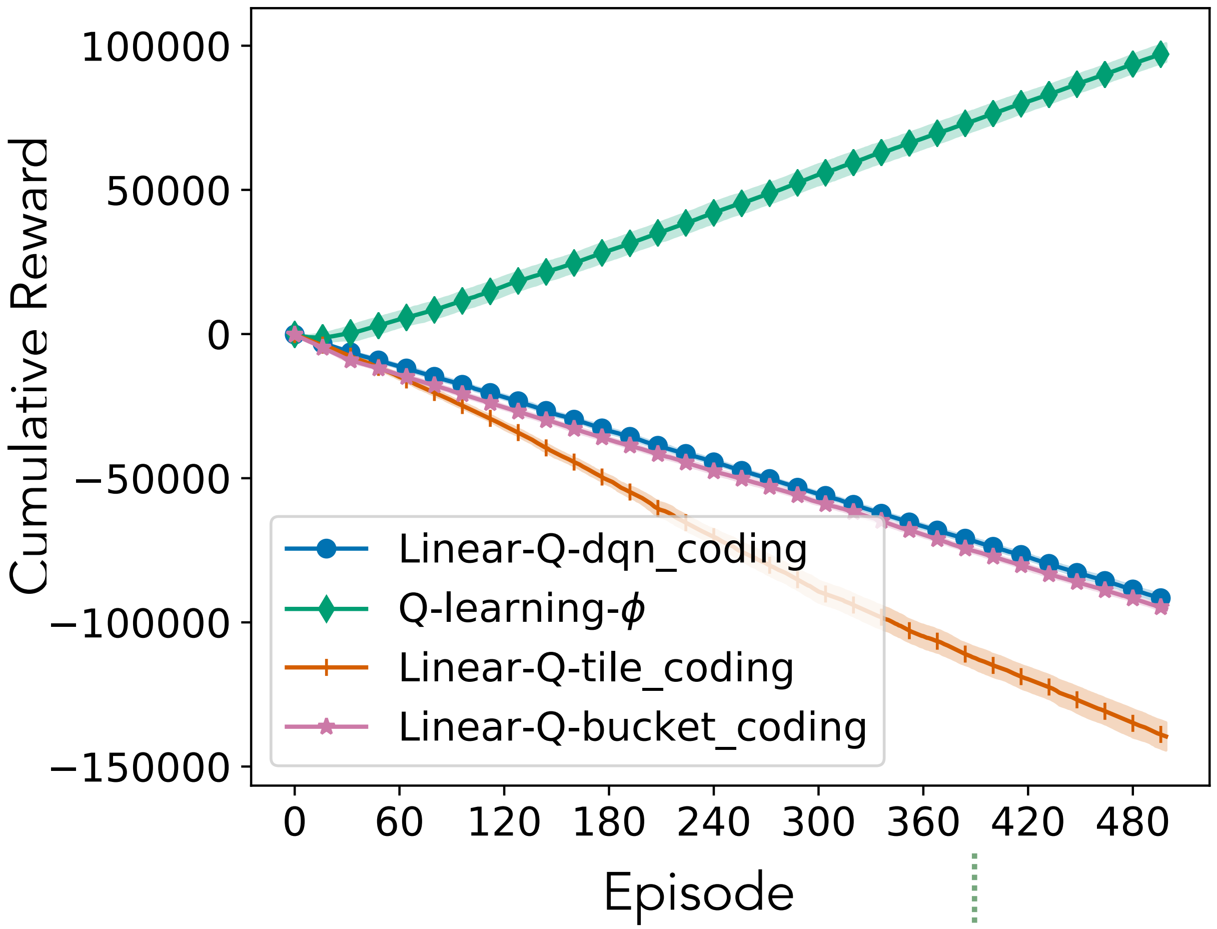
Cumulative Reward: Lifelong Cart Pendulum



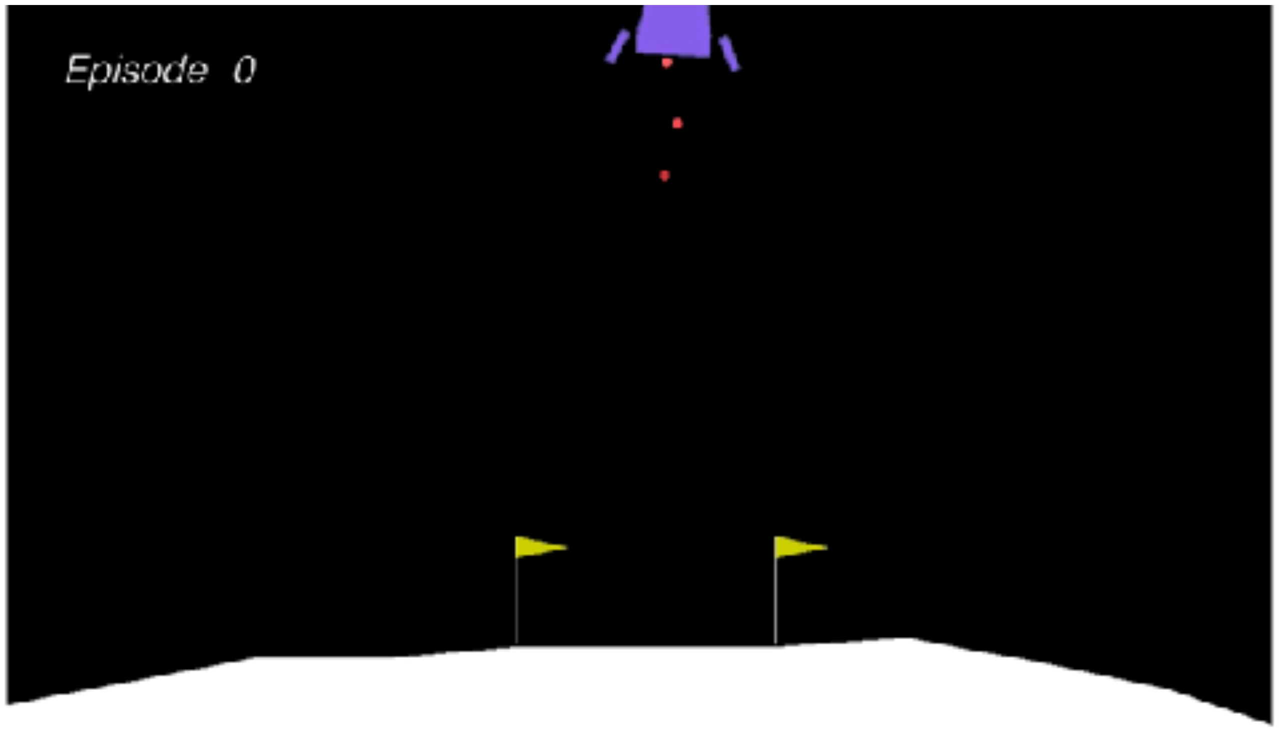
Cumulative Reward: Gym LunarNoShaping V0



Experiments: Lunar Lander



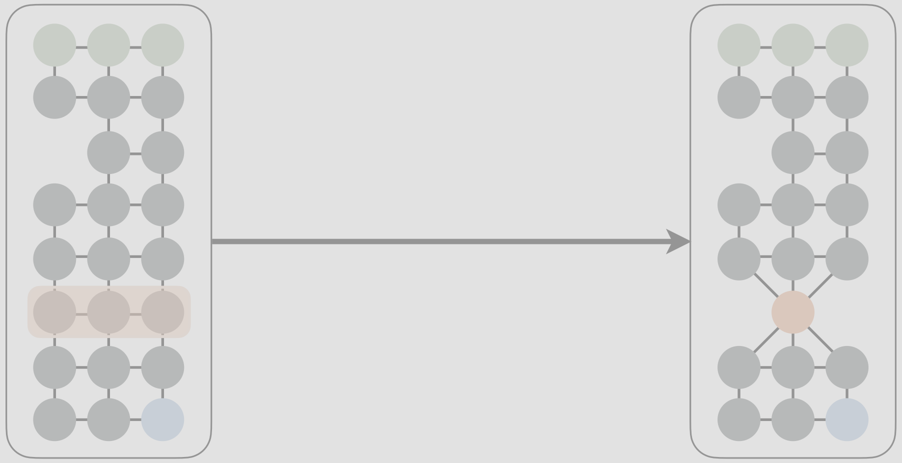
[Sutton '96]



Tabular Q-Learning with ϕ

Part 1

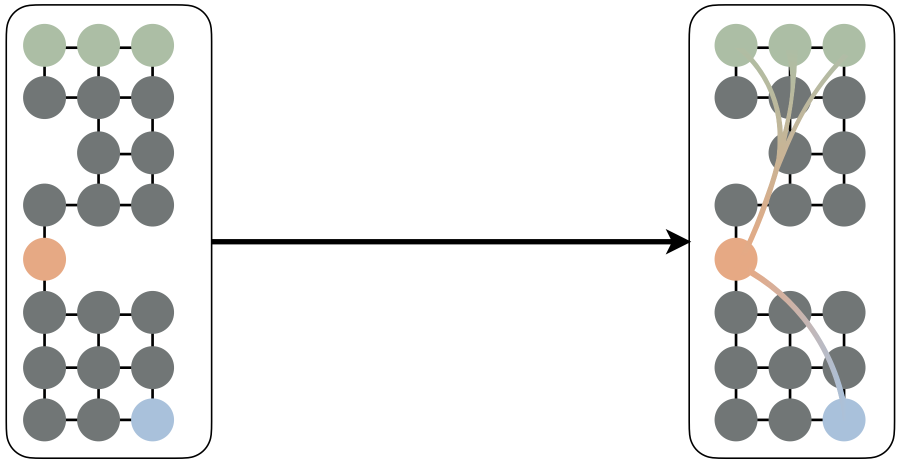
STATE ABSTRACTION



1. Approximate State Abstraction
   ICML 2016
2. State Abstraction In Lifelong RL
   ICML 2018
3. **State Abstraction As Compression**
   AAAI 2019

Part 2

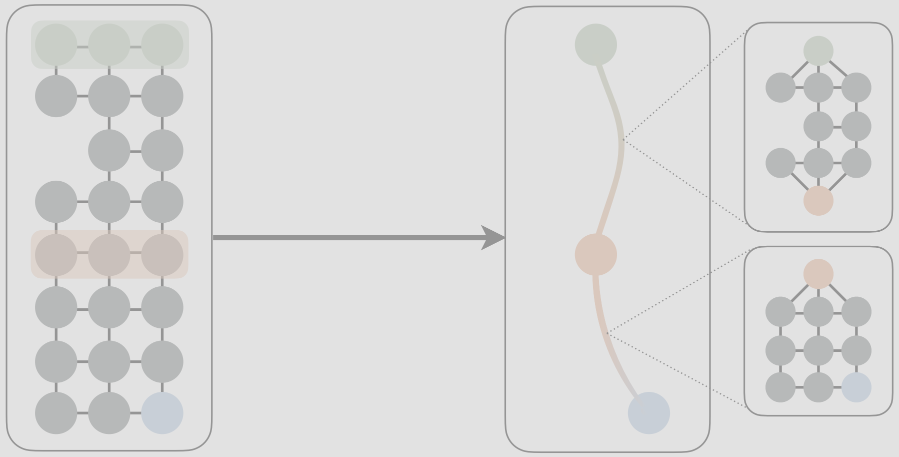
ACTION ABSTRACTION



4. **Options for Planning**
   ICML 2019
5. Options for Exploration
   ICML 2019
6. A New Option Model
   IJCAI 2019

Part 3

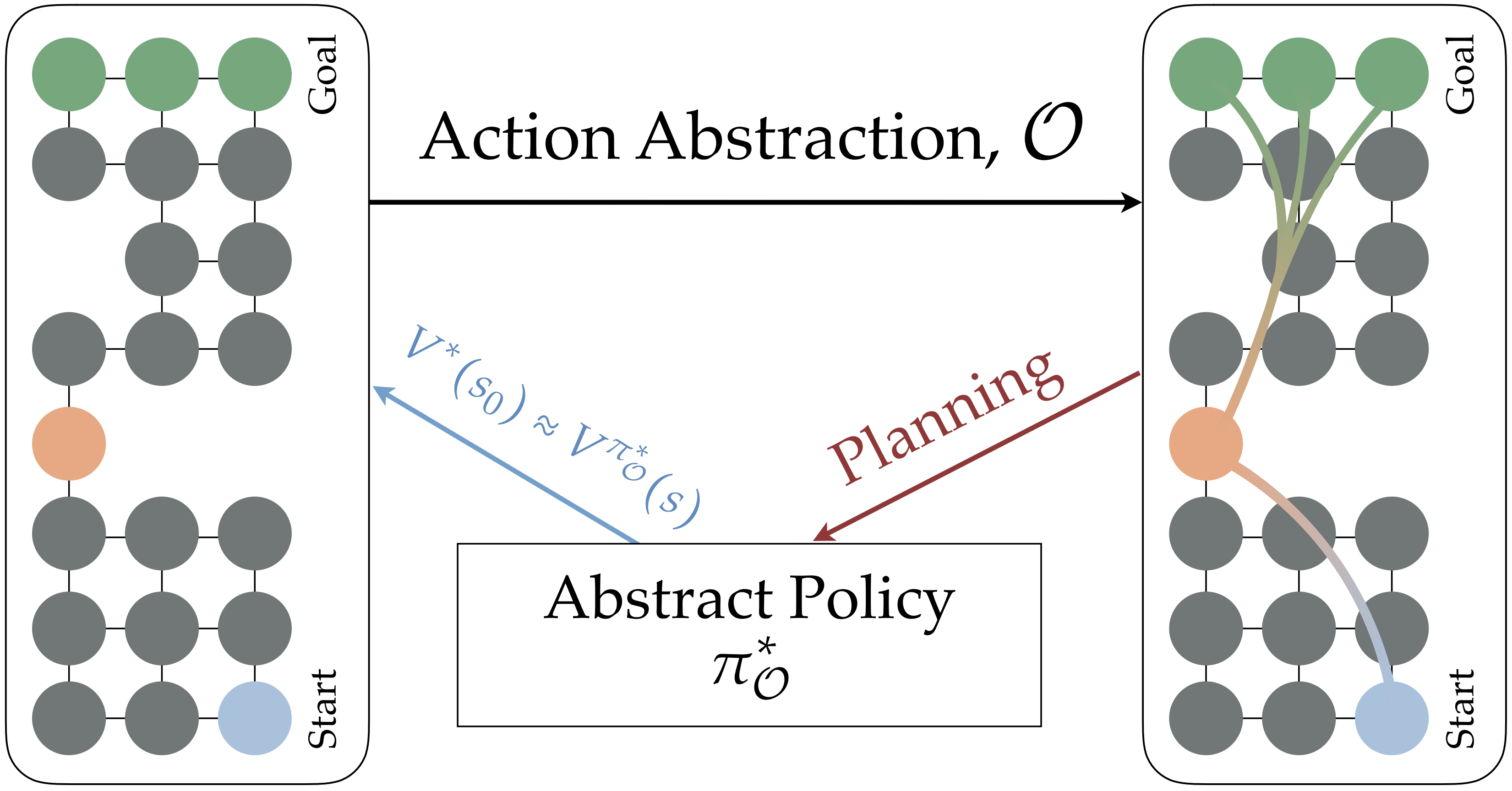
STATE-ACTION ABSTRACTION

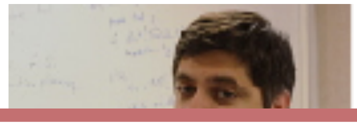


7. **Value-Preserving Hierarchies**
   AISTATS 2020

Options for Planning

*[J**A**HLK, ICML 2019]*


Goal
Start
Action Abstraction, $\mathcal{O}$
$V^*(s_0) \approx V^{\pi^*_{\mathcal{O}}}(s)$
Planning
Abstract Policy
$\pi^*_{\mathcal{O}}$
Goal
Start



***Question:** How can we find the set of options that make planning as fast as possible?*

project → Jinnai Hershkowitz Littman Konidaris

Options for Planning

[J**A**HLK, ICML 2019]

Theorem. Finding the set of options that minimizes planning time is:

1) NP-hard in general.

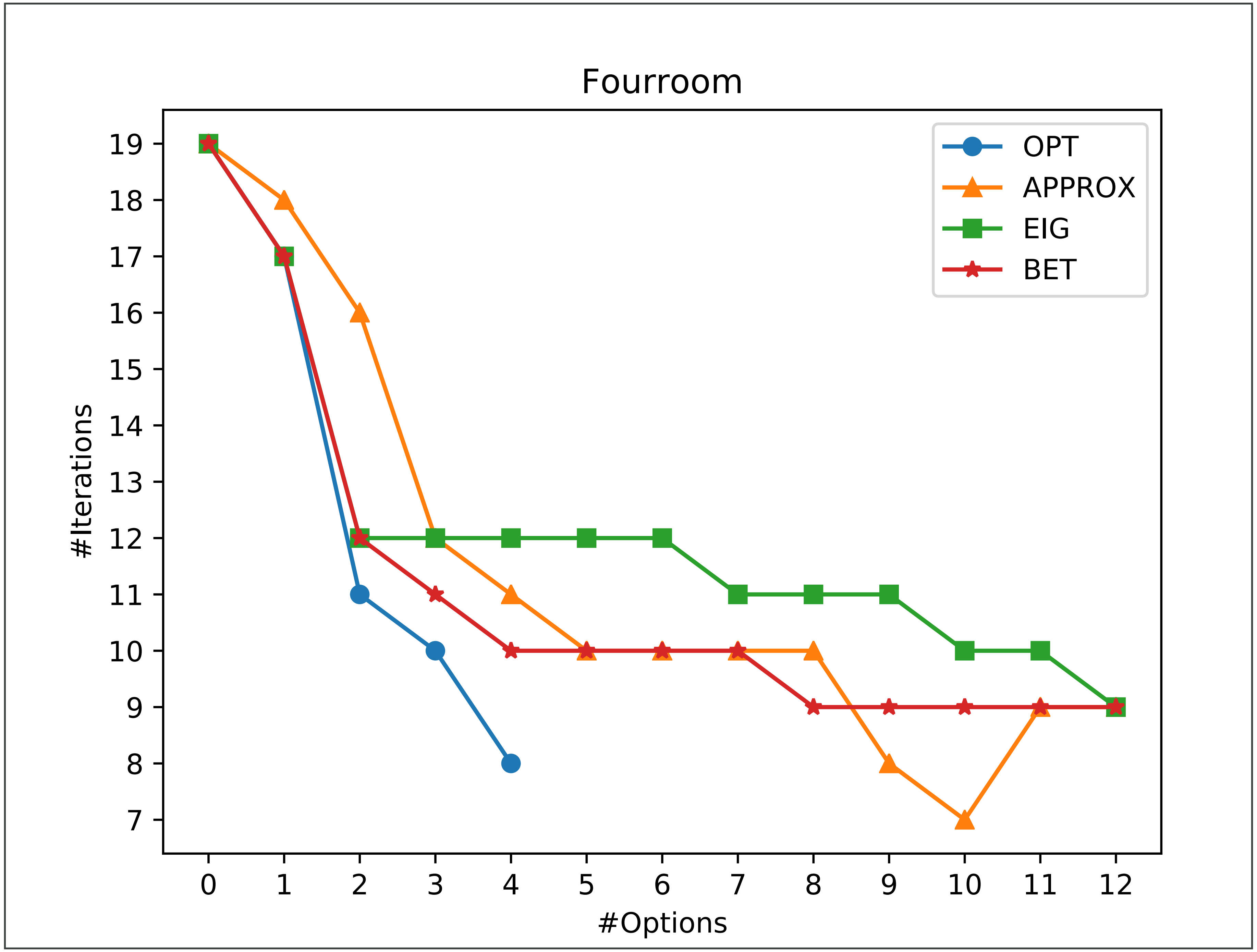
2) $2^{\log^{1-\varepsilon} n}$ -hard to approximate.[1]

[1] Unless NP $\subseteq$ DTIME($n^{\text{poly} \log n}$) [Dinitz et al. 2012]

Question: *How can we find the set of options that make planning as fast as possible?*

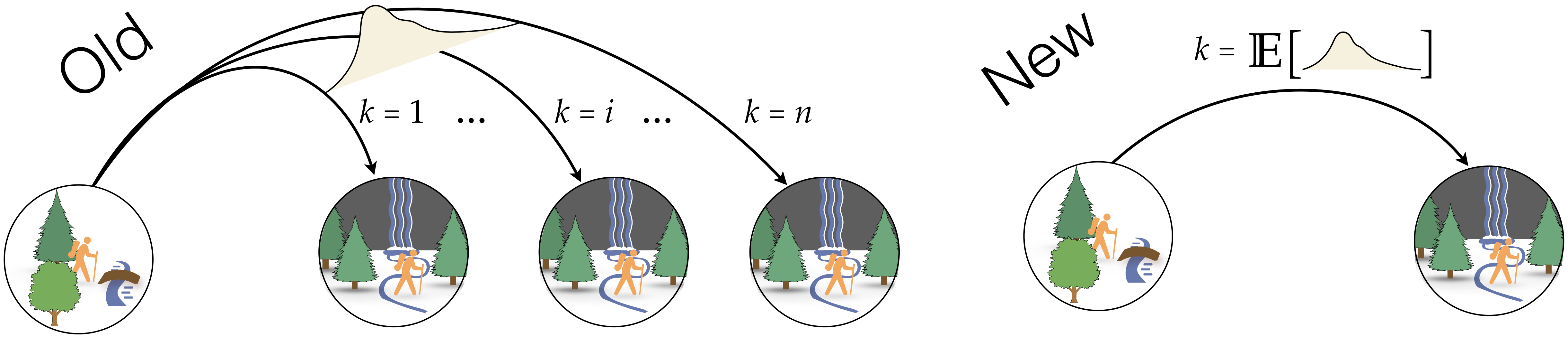
Empirical Evaluation

*[J**A**HLK, ICML 2019]*



A New Option Model

*[**A**WdL, IJCAI 2019]*



jointly led project →

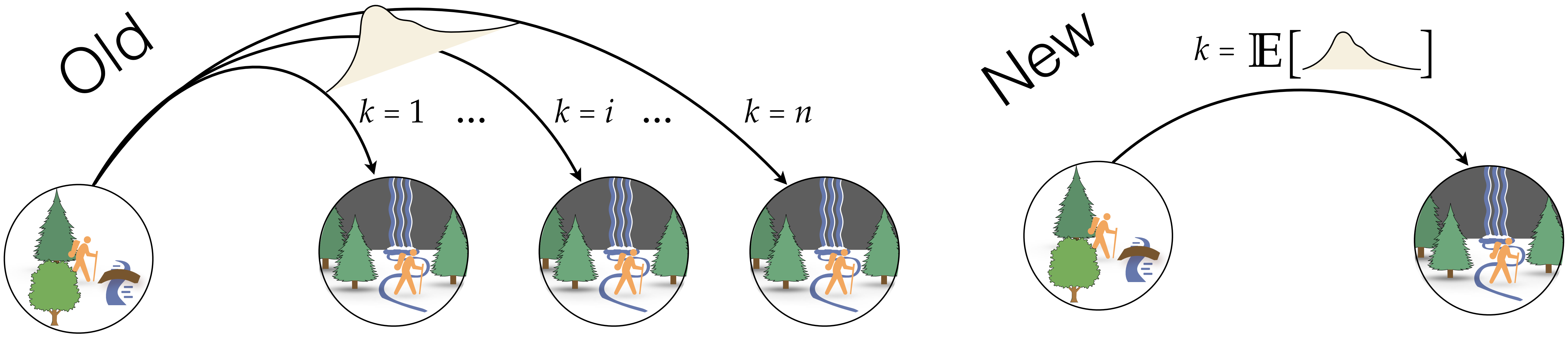
John Winder

Marie desJardins

Michael L. Littman

A New Option Model

*[**A**WdL, IJCAI 2019]*



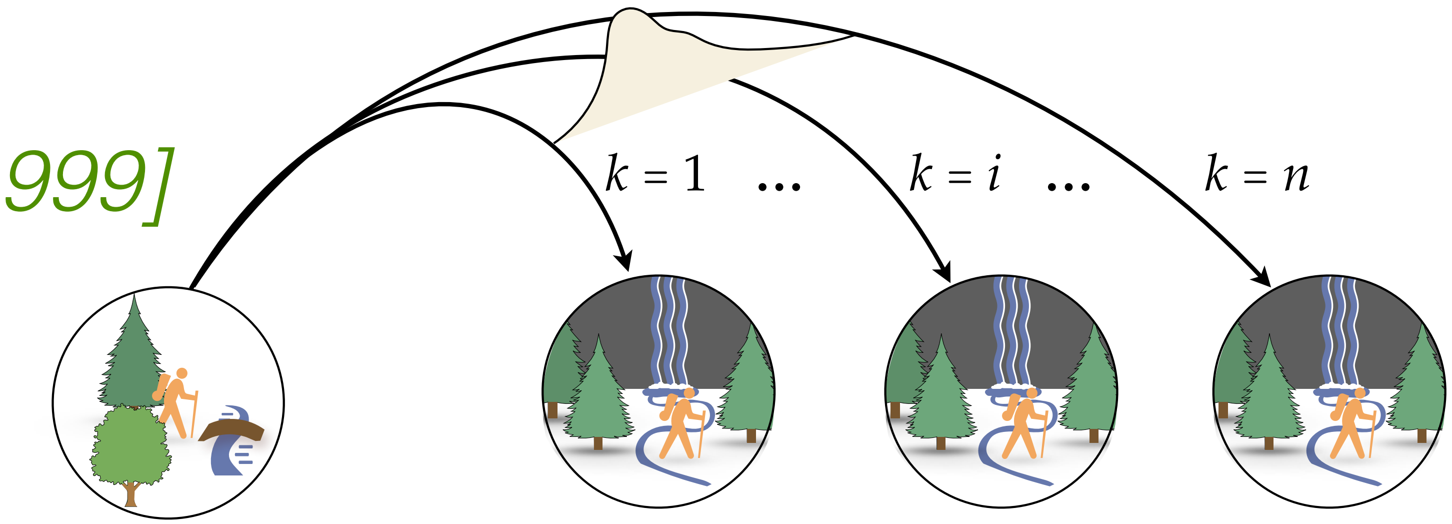
Question: *How can we efficiently estimate the transition and reward models of options?*

A New Option Model

[**A**WdL, IJCAI 2019]

Multi-Time Model

[Sutton, Precup, Singh 1999]



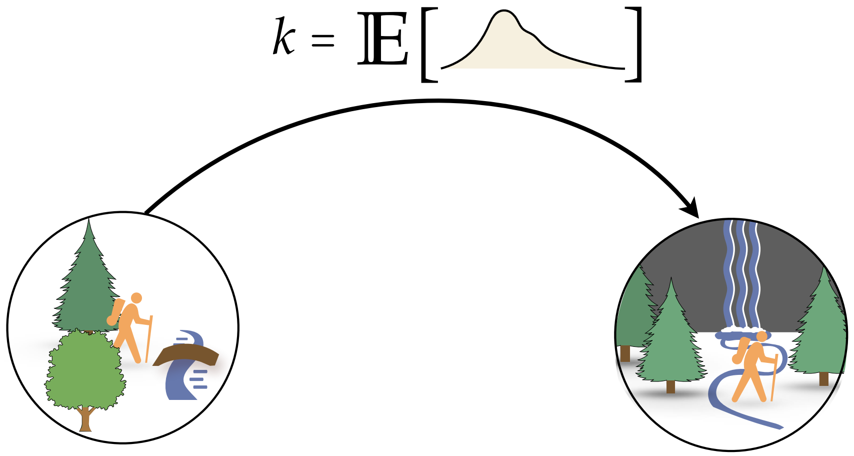
$$T_\gamma(s' \mid s, o) := \sum_{k=0}^{\infty} \gamma^k \beta(s_k) \mathbb{P}(s_k = s' \mid s, o)$$

$$R_\gamma(s, o) := \mathop{\mathbb{E}}_{k, s_{1\ldots k}} \left[r_1 + \gamma r_2 \ldots + \gamma^{k-1} r_k \,\middle|\, s, o \right]$$

A New Option Model

[**A**WdL, IJCAI 2019]

Expected Length Model



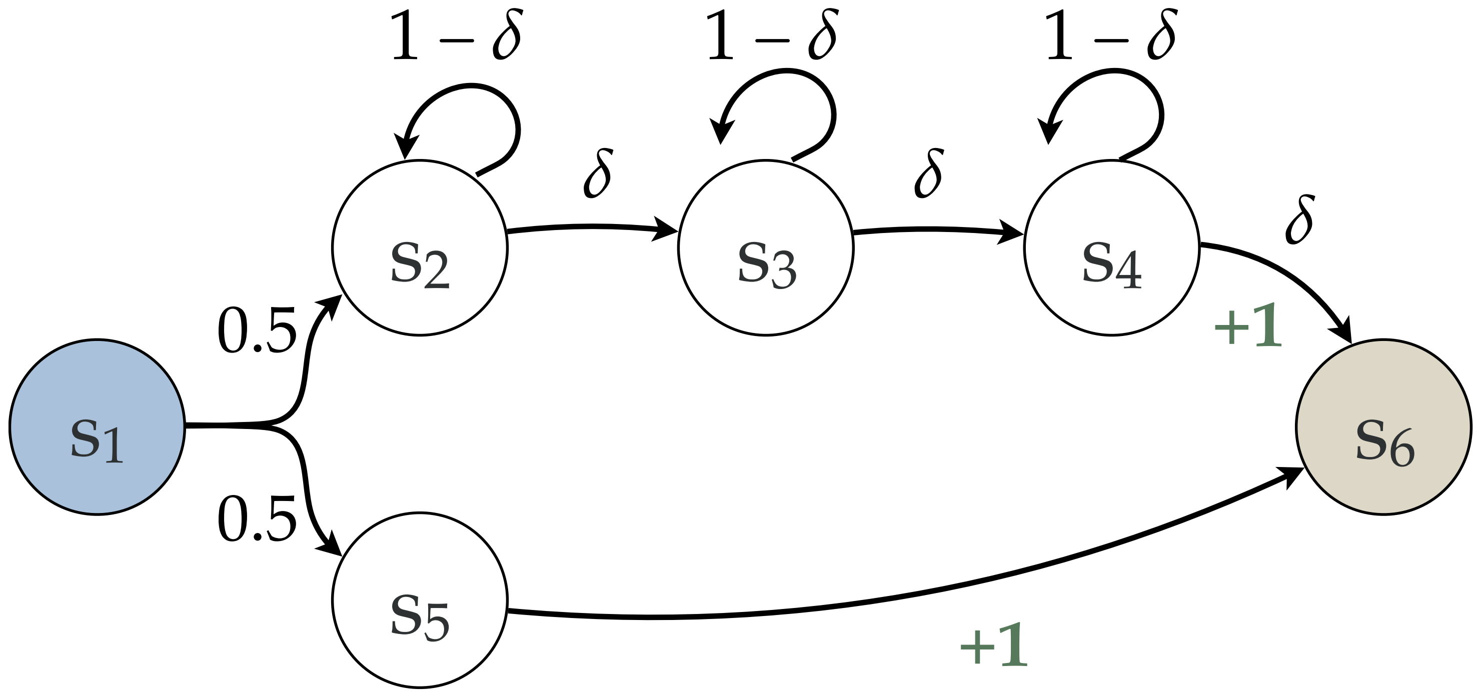
$$T_{\mu_k}(s' \mid s, o) := \gamma^{\mu_k} p(s' \mid s, o),$$

$$R_{\mu_k}(s, o, s') := \gamma^{\mu_k} \mathbb{E}\left[r_1 + r_2 \ldots + r_{\mu_k} \mid s, o\right],$$

where $\mu_k = \mathbb{E}[k \mid s, o]$.

A New Option Model

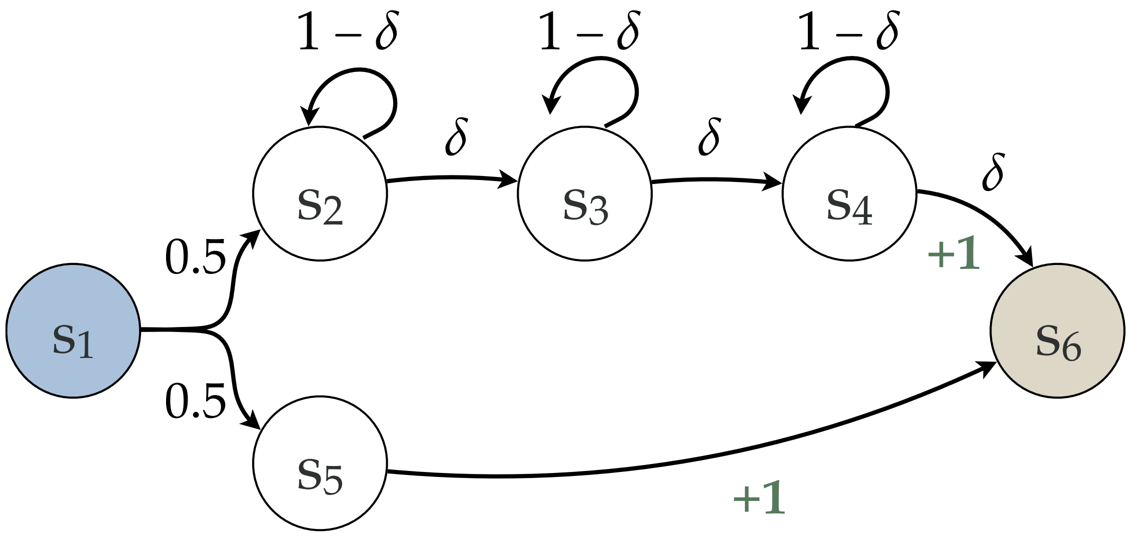
[***A***WdL, IJCAI 2019]

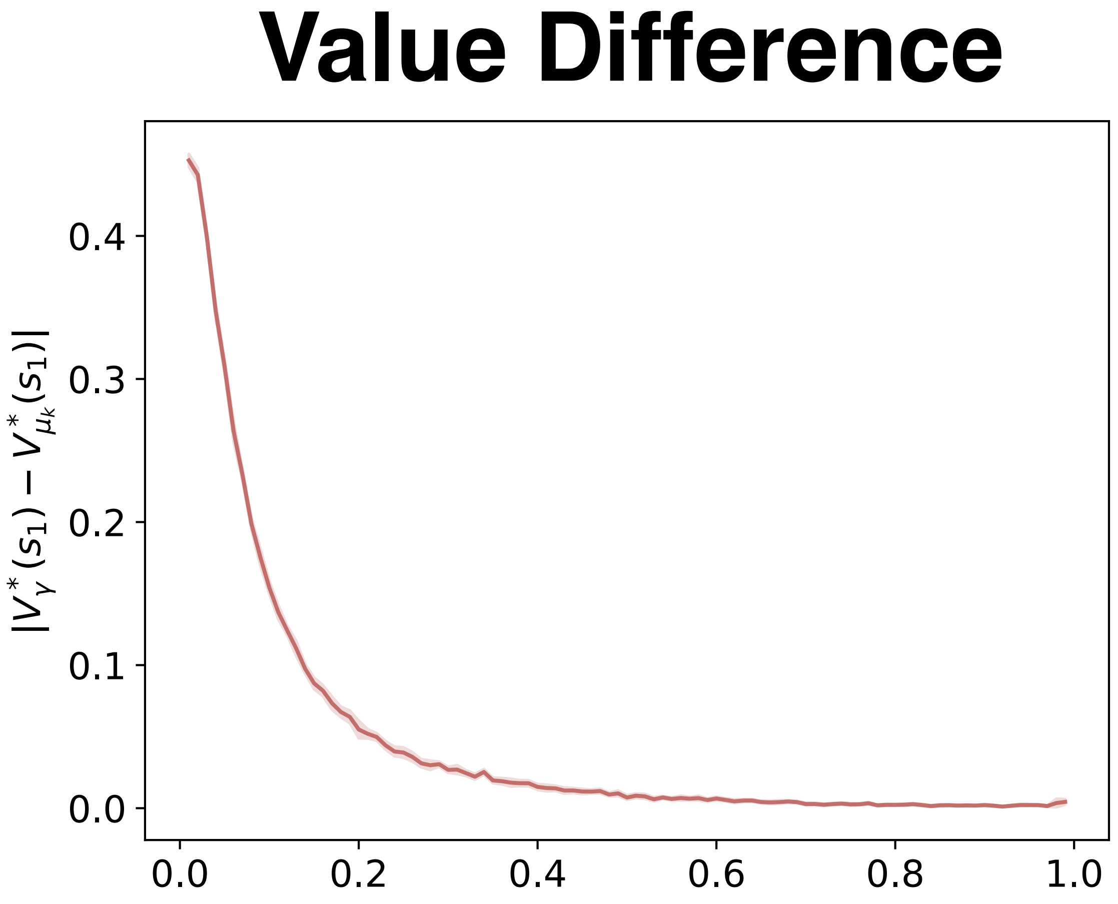


$$\mathbb{P}(s_6, k \mid s, o)$$

A New Option Model

*[**A**WdL, IJCAI 2019]*





more stochastic δ ***less stochastic***

$$\mathbb{P}(s_6, k \mid s, o)$$

A New Option Model

[**A**WdL, IJCAI 2019]

Lemma. *There exists a $\tau \geq 1$ such that*

$$|T_\gamma(s' \mid s, o) - T_{\mu_k}(s' \mid s, o)| \leq \gamma^{\mu_{k,o} - \tau}(2\tau + 1)e^{-\beta_{\min}}.$$

Lemma. *In stochastic shortest path MDPs,*

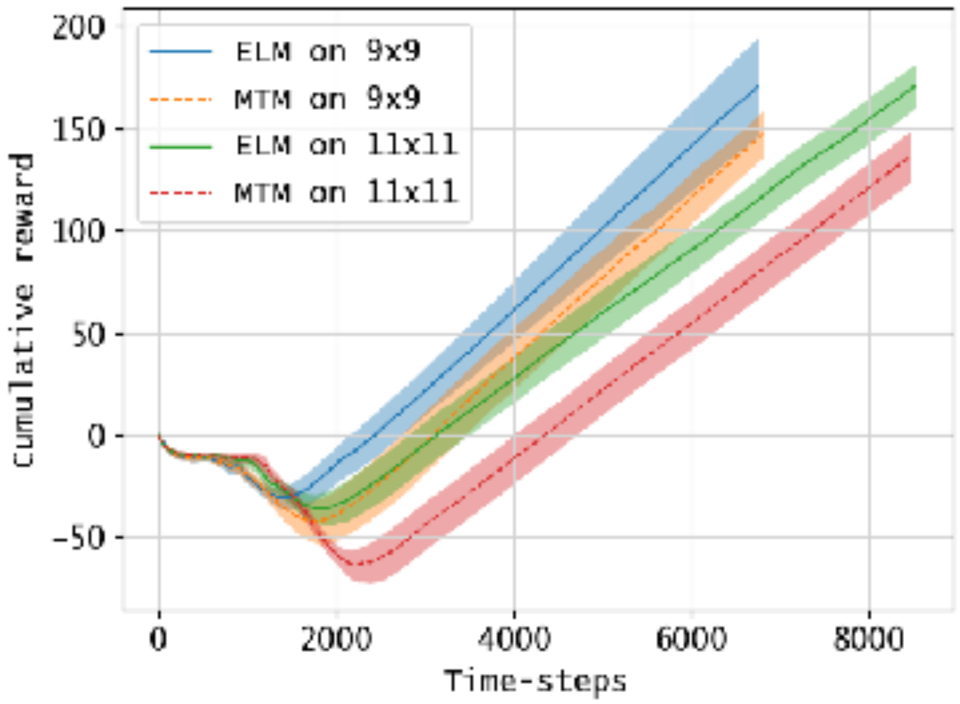
$$|R_\gamma(s, o) - R_{\mu_k}(s, o)| = |T_\gamma(s_g \mid s, o) - T_{\mu_k}(s_g \mid s, o)|.$$

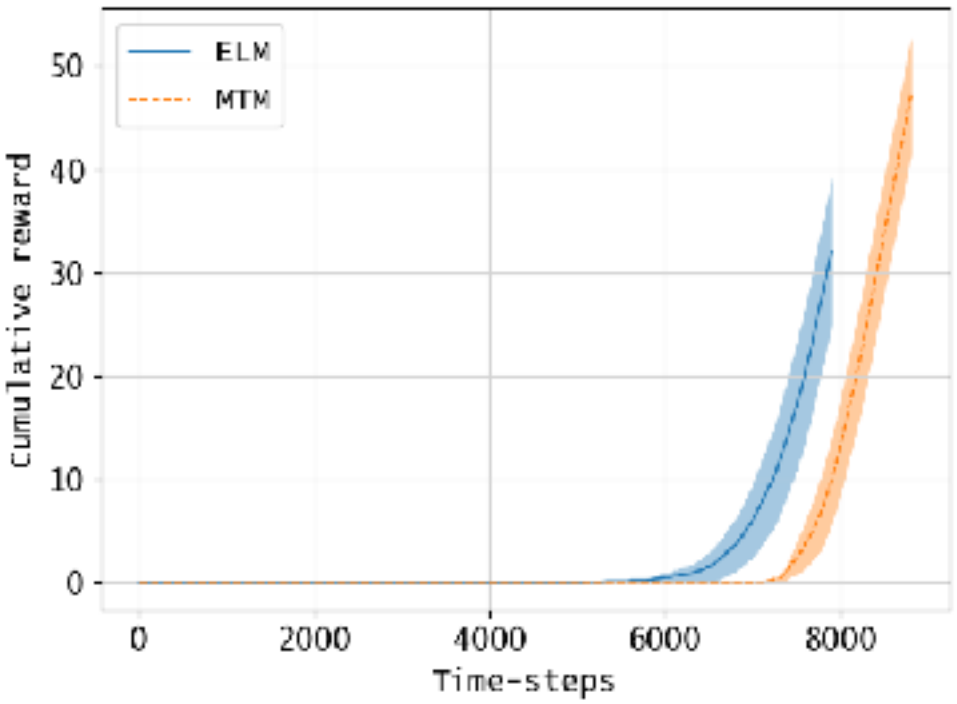
Theorem. *In stochastic shortest path MDPs,*

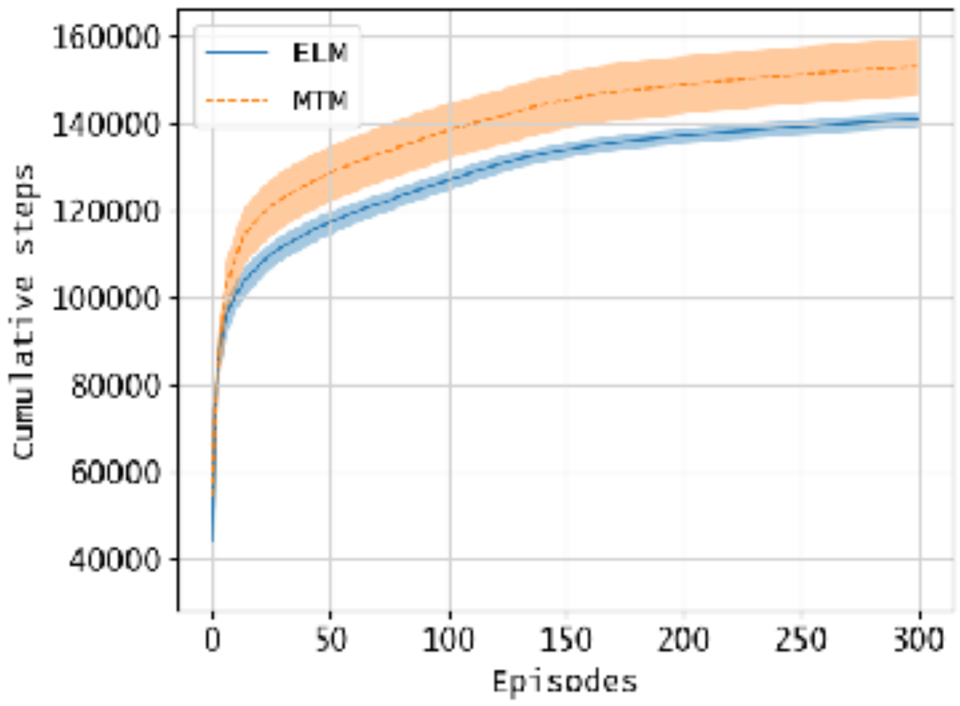
$$|V_\gamma^{\pi_o}(s) - V_{\mu_k}^{\pi_o}(s)| \leq \frac{\varepsilon(1 - \gamma^{\mu_k}) + \gamma^{\mu_k}\frac{\varepsilon}{2}\text{RMAX}}{(1 - \gamma^{\mu_k})(1 - \gamma^{\mu_k} + \frac{\varepsilon}{2}\gamma^{\mu_k})}.$$

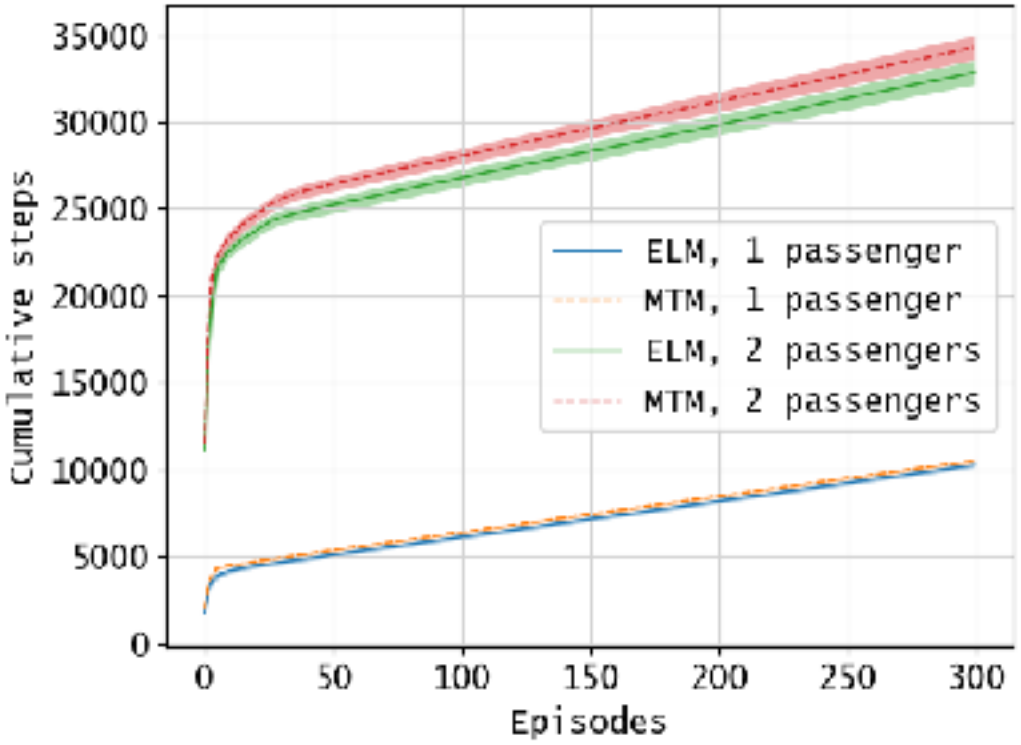
A New Option Model

*[**A**WdL, IJCAI 2019]*



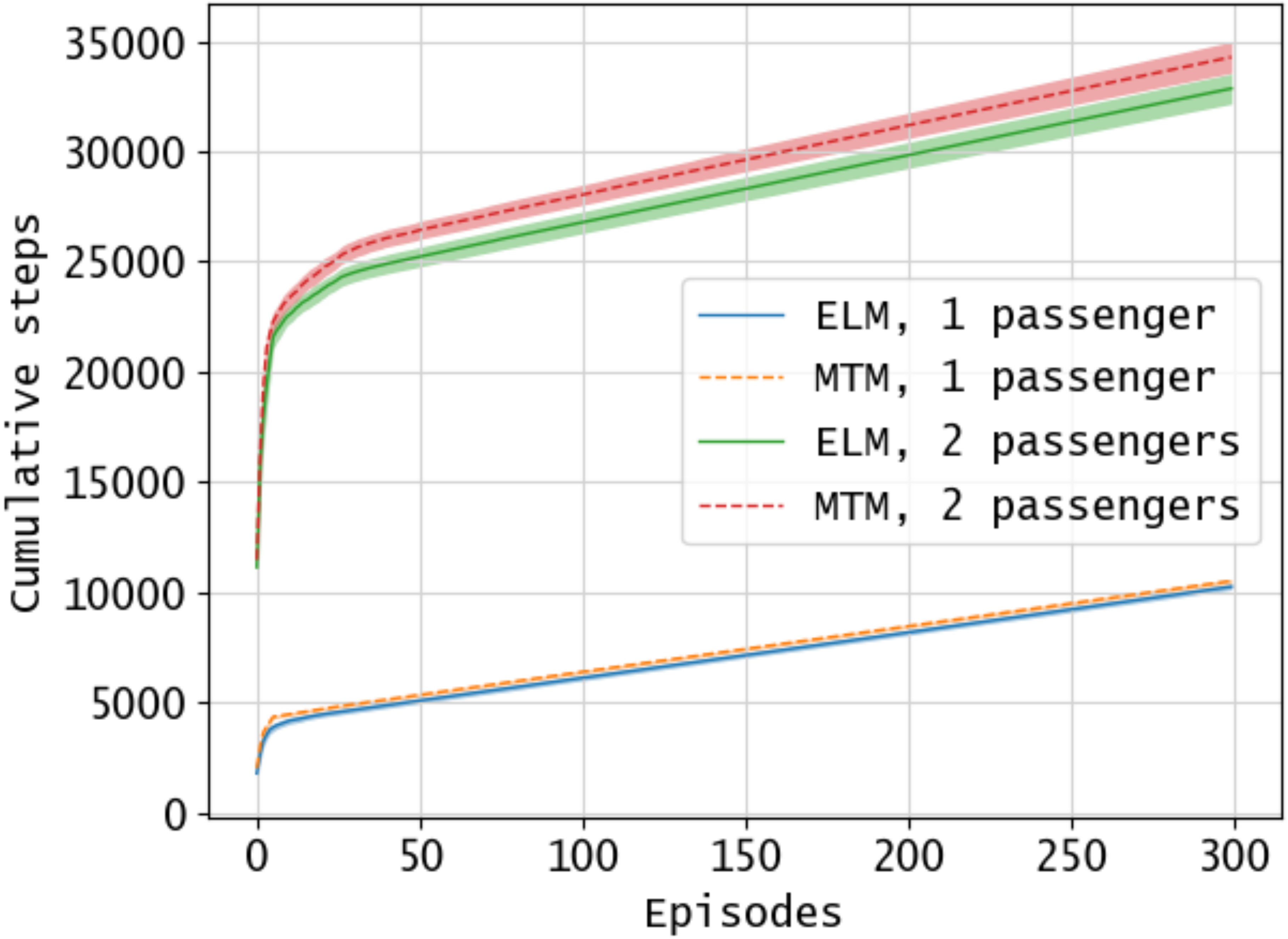






A New Option Model

[***A***WdL, IJCAI 2019]

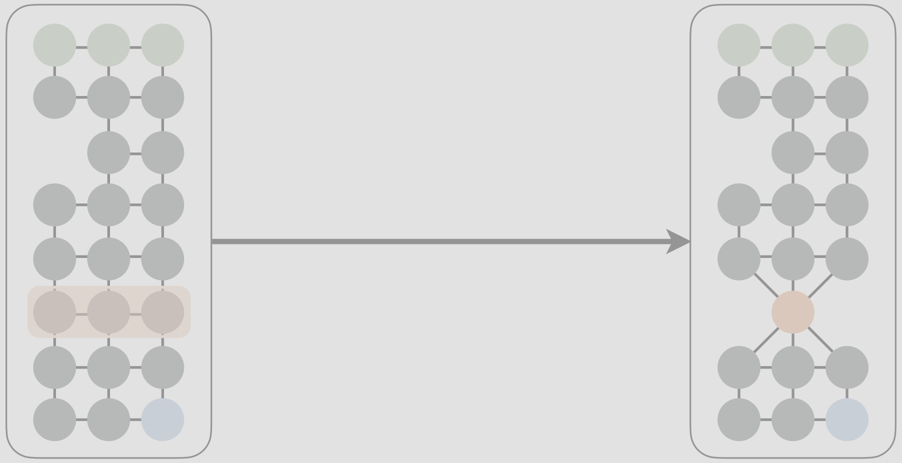


Taxi Domain

[Dietterich '00]

Part 1

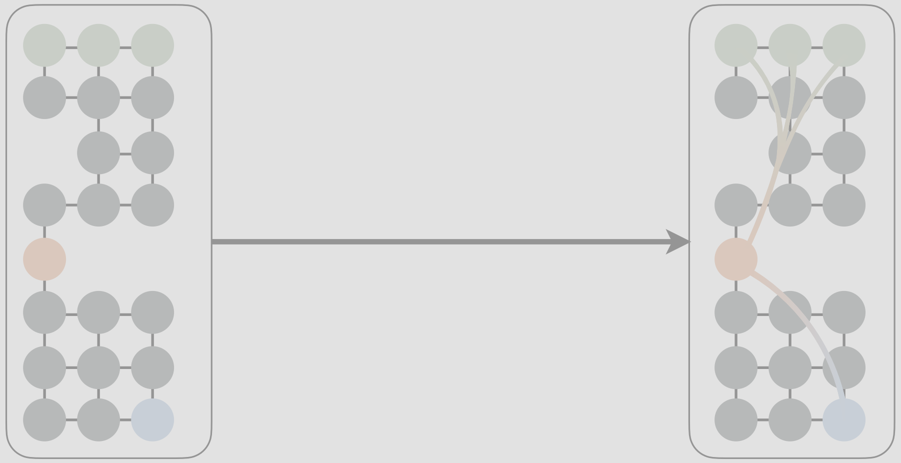
STATE ABSTRACTION



1. Approximate State Abstraction
ICML 2016
2. State Abstraction In Lifelong RL
ICML 2018
3. **State Abstraction As Compression**
AAAI 2019

Part 2

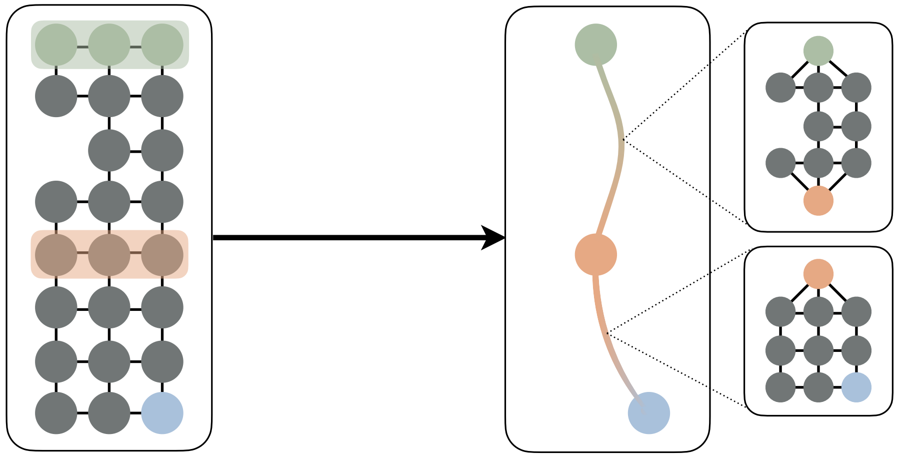
ACTION ABSTRACTION



4. **Options for Planning**
ICML 2019
5. A New Option Model
IJCAI 2019
6. Options for Exploration
ICML 2019

Part 3

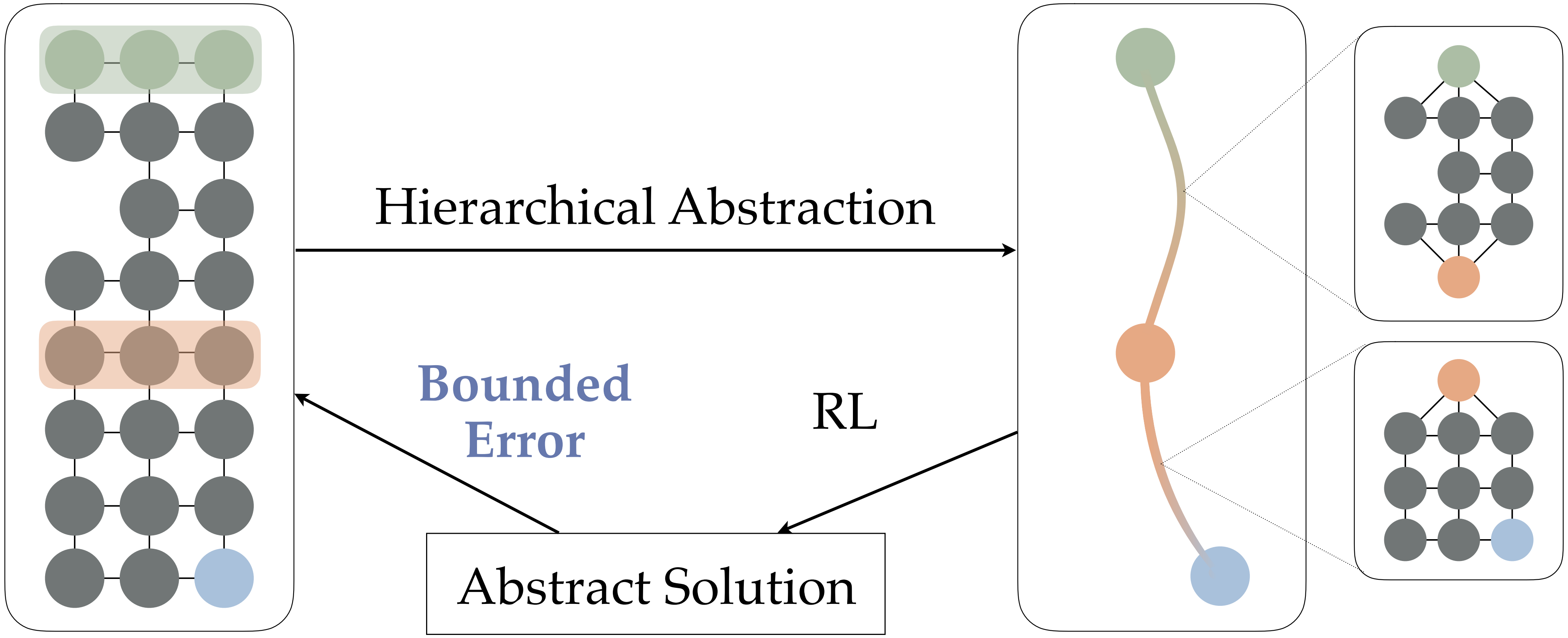
STATE-ACTION ABSTRACTION



7. **Value-Preserving Hierarchies**
AISTATS 2020

Value-Preserving Hierarchies

*[**A**UKAPL, AISTATS 2020]*



Hierarchical Abstraction

Bounded Error

RL

Abstract Solution

Nathan Umbanhower

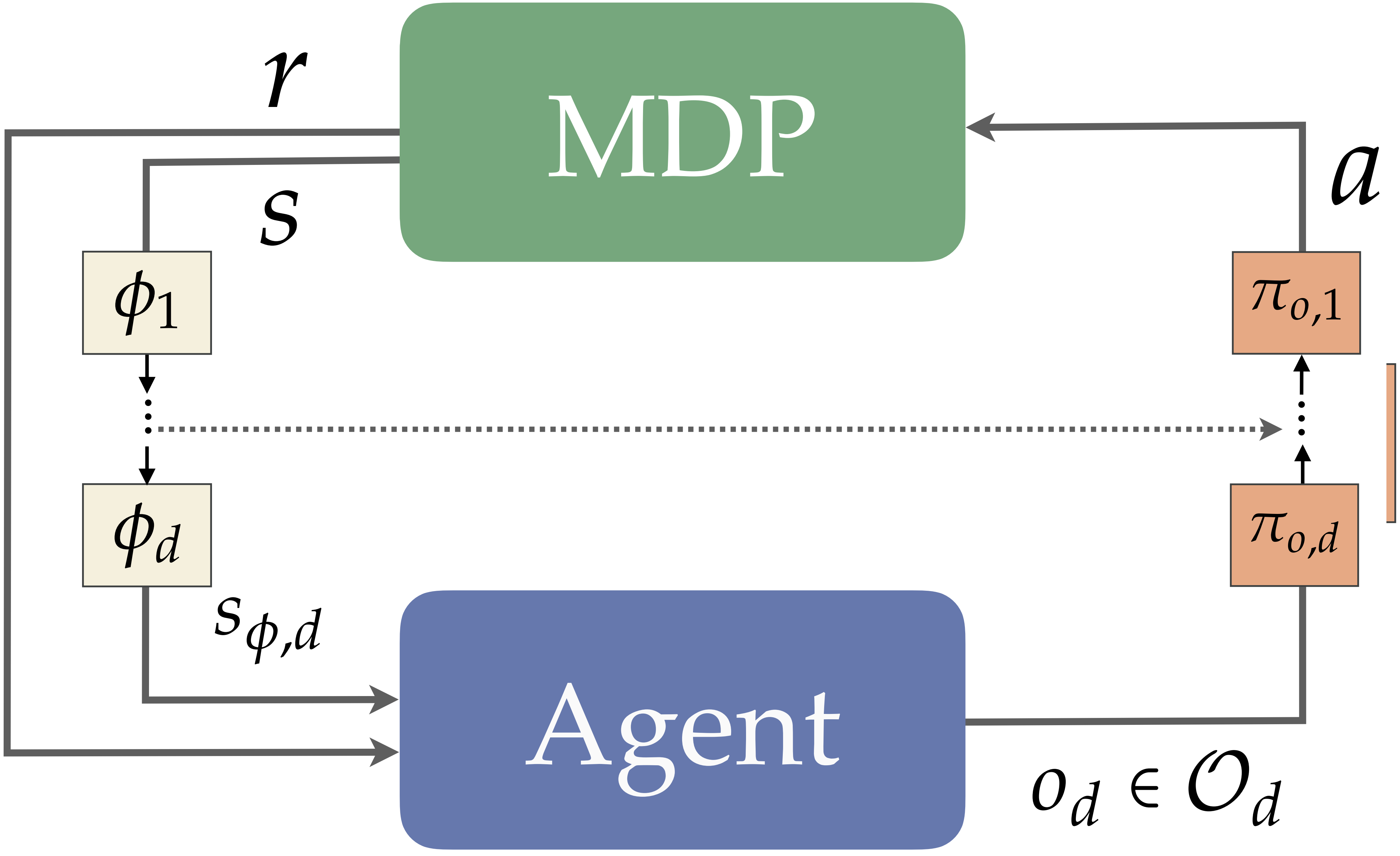
Khimya Khetarpal

Dilip Arumugam

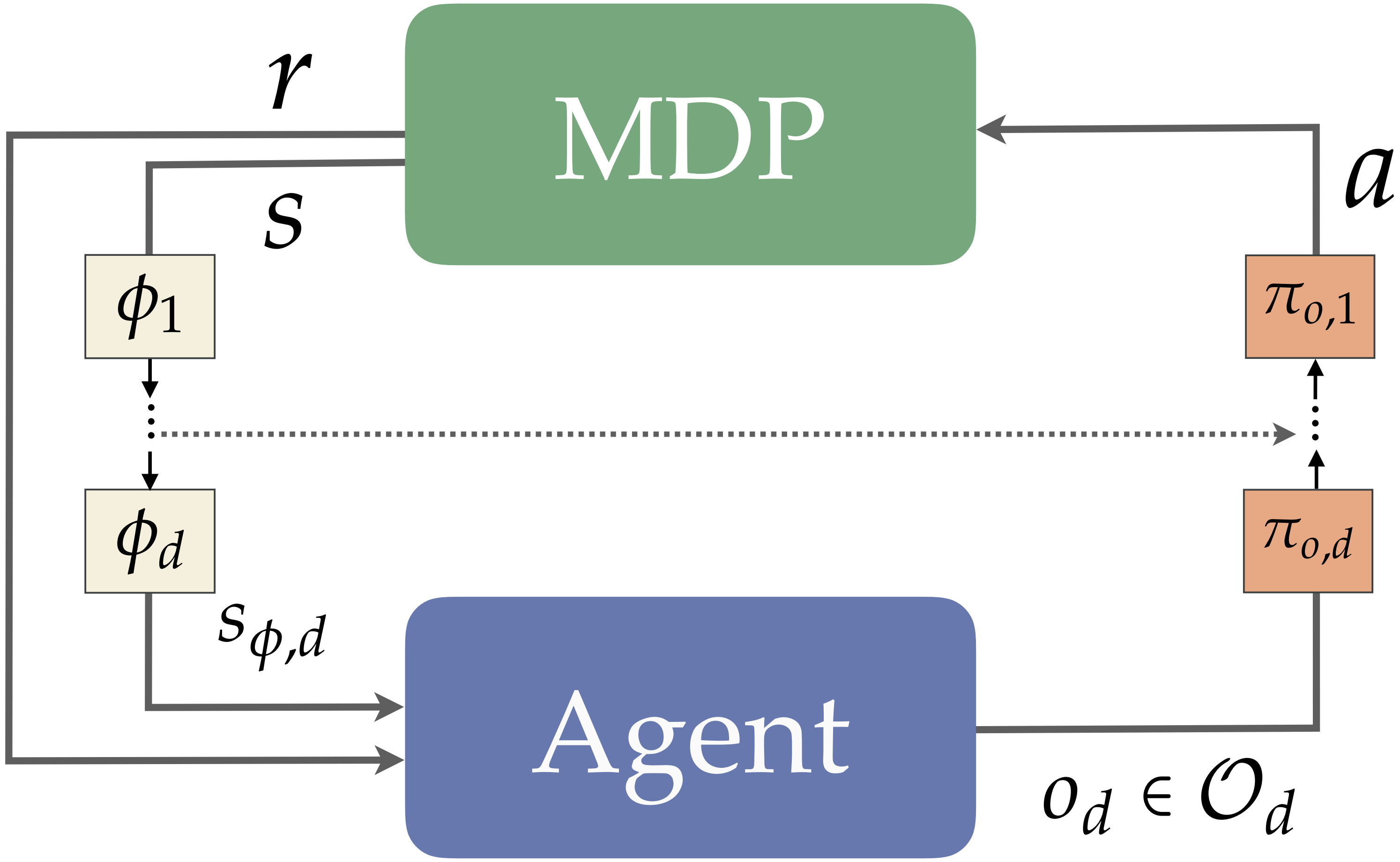
Doina Precup

Michael L. Littman

Hierarchical RL



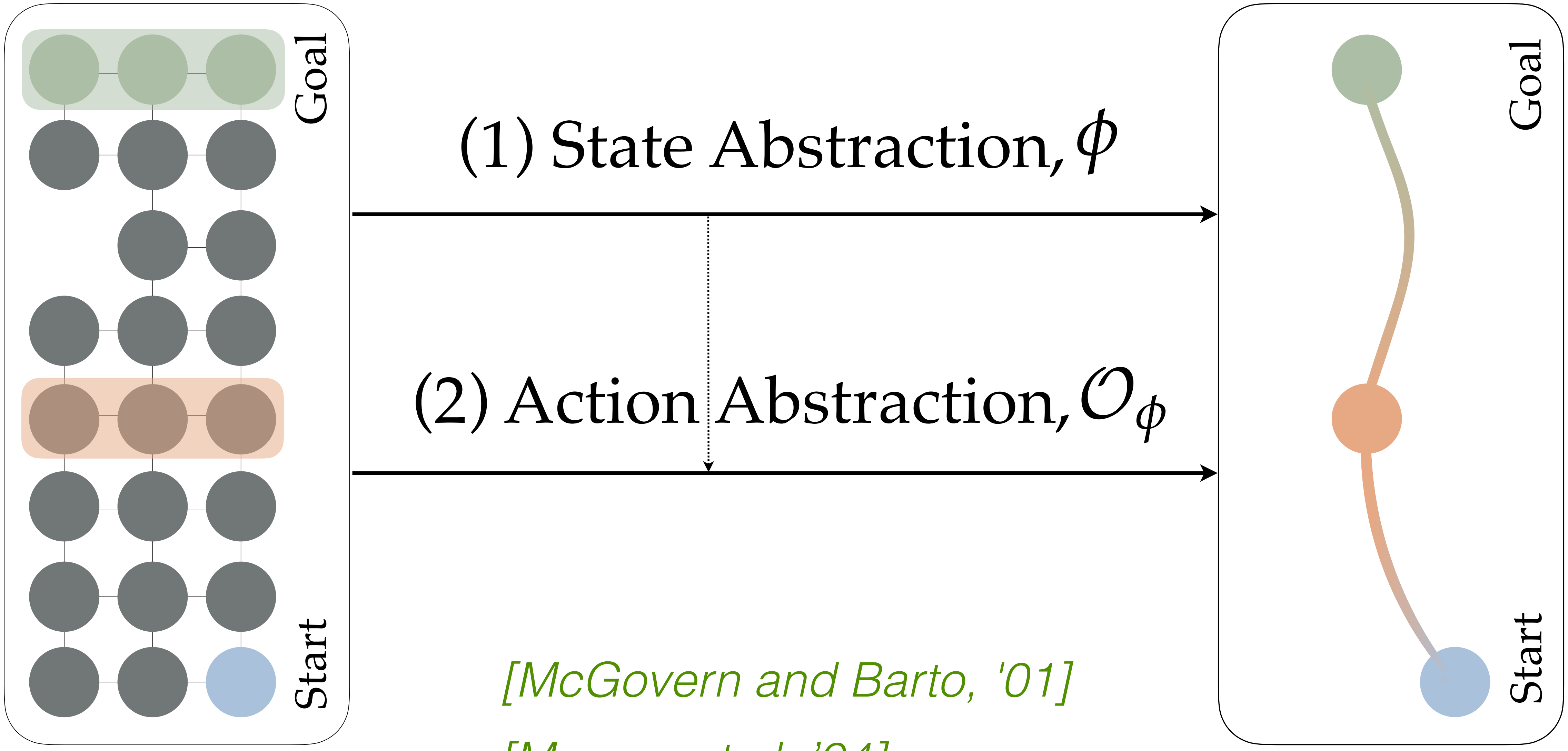
Value-Preserving Hierarchical RL



[Ravindran, Barto '03, '04]
[Majeed & Hutter '19]

Question: *Which combinations of state abstractions and options preserve representation of good behavior?*

ϕ-Relative Options



(1) State Abstraction, ϕ

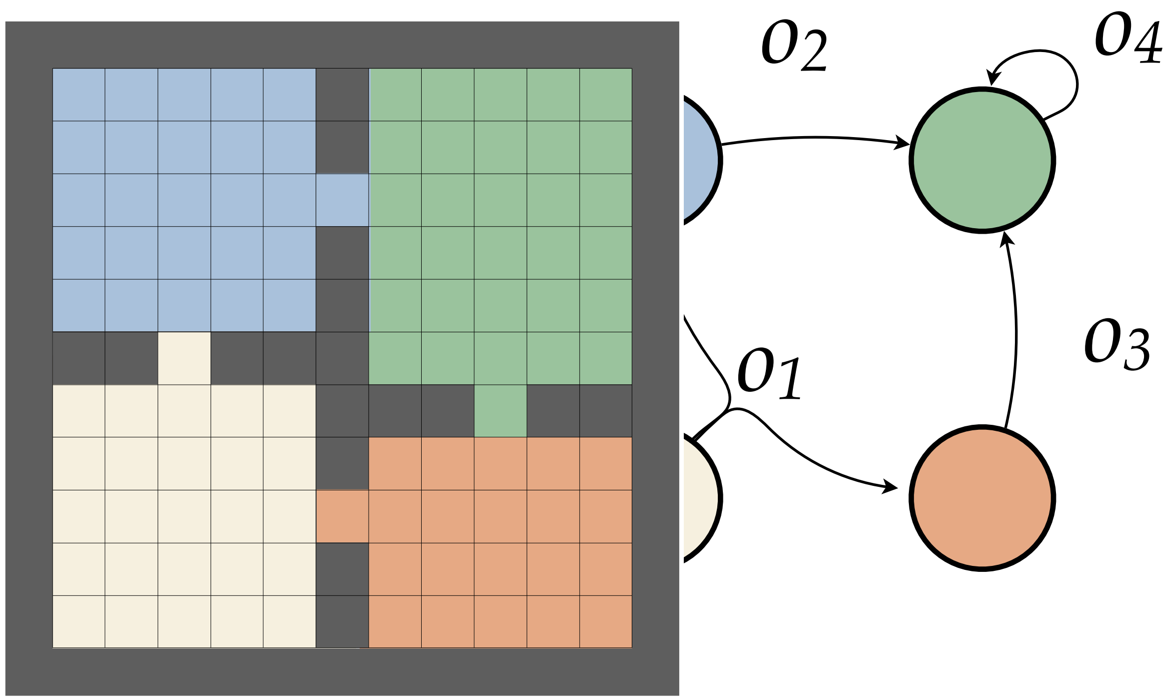
(2) Action Abstraction, $\mathcal{O}_{\phi}$

[McGovern and Barto, '01]

[Mannor et al. '04]

[Provost et al. '06]

ϕ-Relative Options



Given ϕ...

Options must respect the abstract state boundaries.

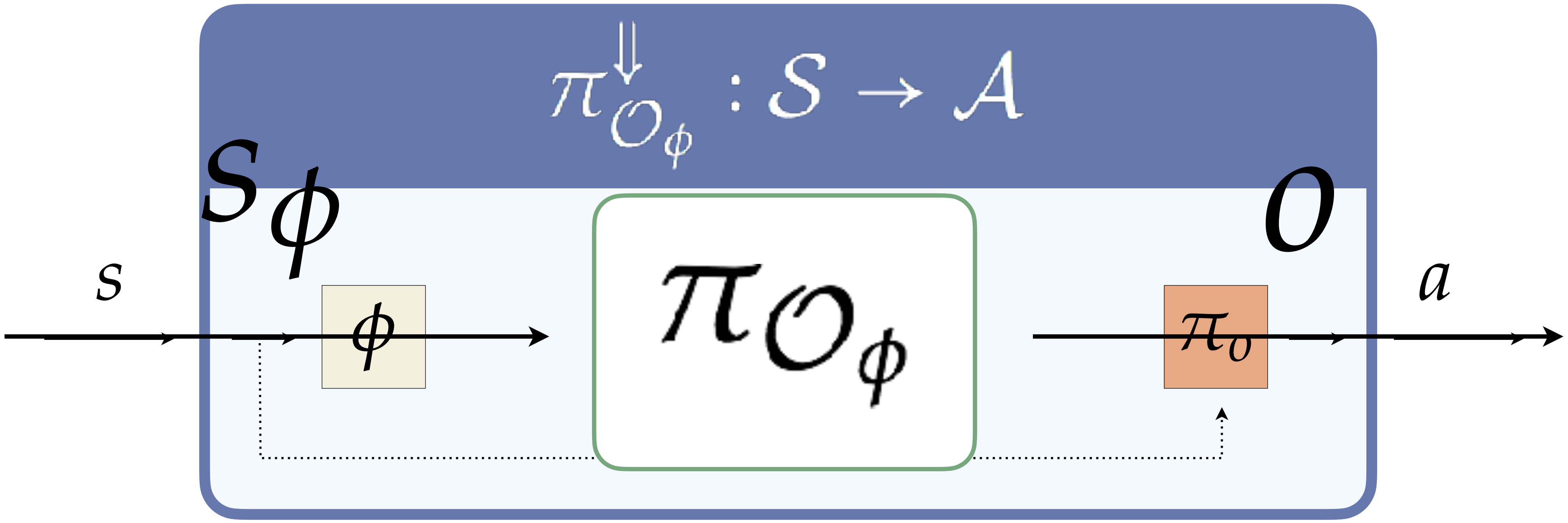
Definition. A set of options is said to be **ϕ-relative**, denoted $\mathcal{O}_\phi$, if:

1. Each $o \in \mathcal{O}_\phi$ initiates in some s_ϕ, terminates when $s \notin s_\phi$.
2. For each abstract state, there is at least one $o \in \mathcal{O}$ that initiates in that state.

ϕ-Relative Options

(1) State Abstraction, ϕ

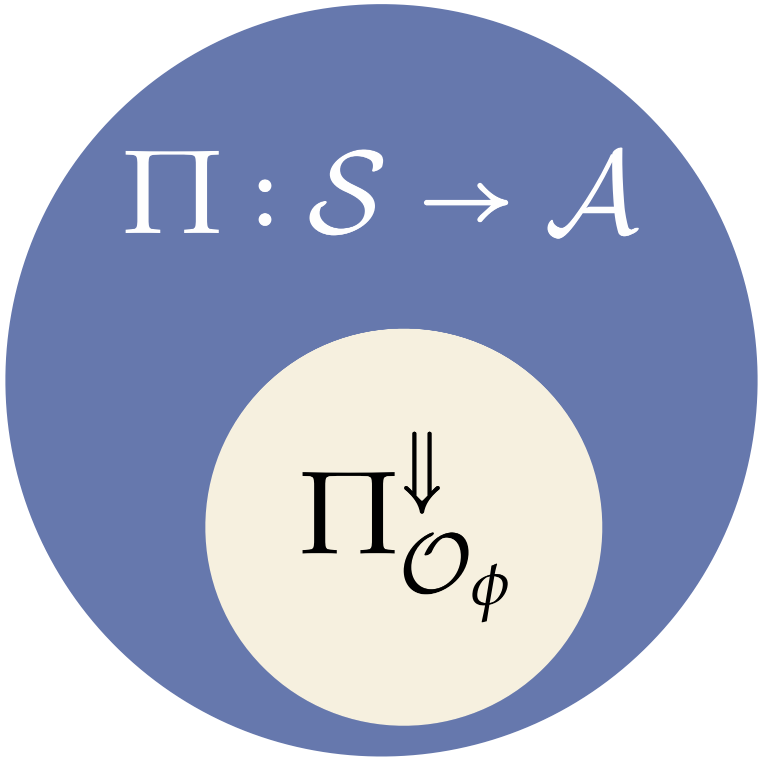
(2) Action Abstraction, $\mathcal{O}_\phi$


$\pi^{\Downarrow}_{\mathcal{O}_\phi} : \mathcal{S} \rightarrow \mathcal{A}$
s_ϕ
o
s
ϕ
$\pi_{\mathcal{O}_\phi}$
π_o
a

ϕ-Relative Options

(1) State Abstraction, ϕ

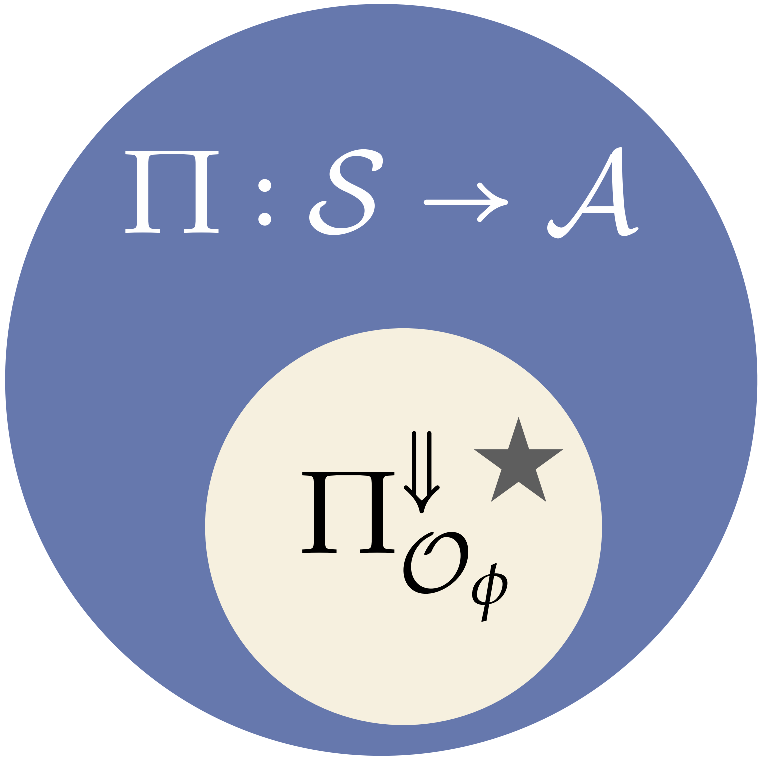
(2) Action Abstraction, $\mathcal{O}_\phi$


$\Pi : \mathcal{S} \rightarrow \mathcal{A}$
$\Pi^{\Downarrow}_{\mathcal{O}_\phi}$

ϕ-Relative Options

(1) State Abstraction, ϕ

(2) Action Abstraction, $\mathcal{O}_\phi$


$\Pi : \mathcal{S} \rightarrow \mathcal{A}$
$\Pi_{\mathcal{O}_\phi}^{\Downarrow}$
★

Question: *Which* $\phi, \mathcal{O}_\phi$ *pairs induce a policy class* $\Pi_{\mathcal{O}_\phi}^{\Downarrow}$ *such that the best abstract policy is still pretty good?*

Value-Preserving Abstractions

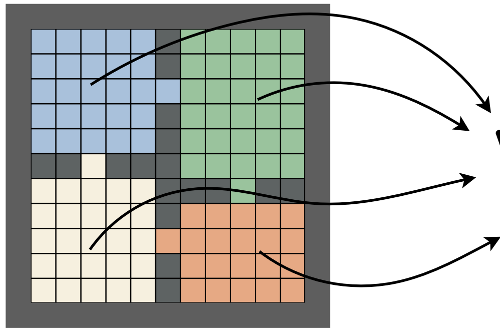
Theorem. There exist at least four classes of $\phi, \mathcal{O}_\phi$ with bounded value loss:

$$\min_{\pi^{\Downarrow}_{\mathcal{O}_\phi} \in \Pi^{\Downarrow}_{\mathcal{O}_\phi}} \max_{s \in \mathcal{S}} \left(V^*(s) - V^{\pi^{\Downarrow}_{\mathcal{O}_\phi}}(s) \right) \leq \eta_p,$$

where η_p varies depending on the class.

Question: *Which* $\phi, \mathcal{O}_\phi$ *pairs induce a policy class* $\Pi^{\Downarrow}_{\mathcal{O}_\phi}$ *such that the best abstract policy is still pretty good?*

ϕ-Relative Option Classes

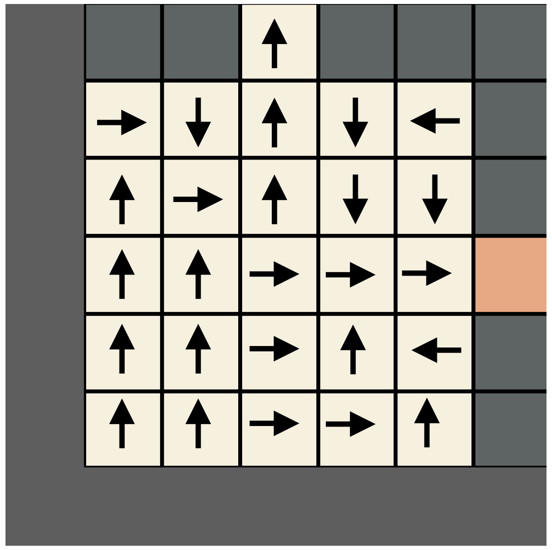


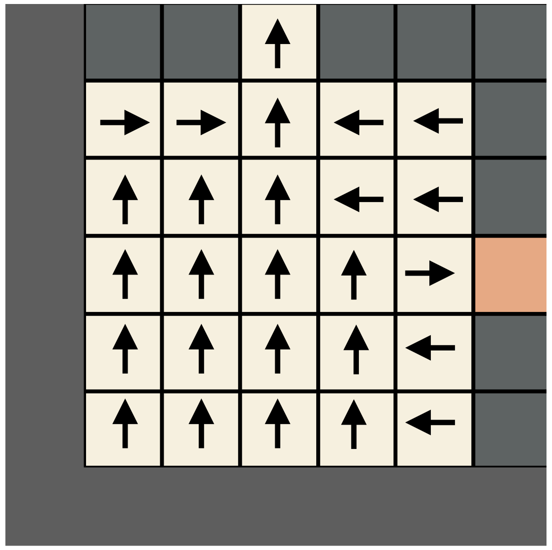
$$\forall_{s_\phi \in \mathcal{S}_\phi} \exists_{o \in \boxed{\Omega(s_\phi)}} :$$

All options that initiate in s_ϕ

Determines option class

$$o \boxed{\approx} o^*_{s_\phi}$$

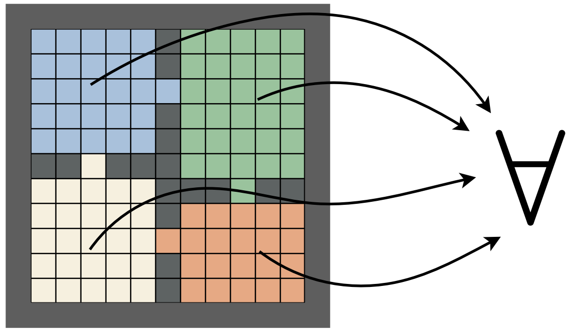




$$o^*_{s_\phi} = (\boxed{s \in s_\phi}, \boxed{s \notin s_\phi}, \boxed{\pi^*})$$

initiate ***terminate*** **policy**

ϕ-Relative Option Classes



$$\forall_{s_\phi \in \mathcal{S}_\phi} \exists_{o \in \boxed{\Omega(s_\phi)}} :$$

All options that initiate in s_ϕ

Expressive Q^* Options

$$\boxed{|Q^*(s_\phi, o) - Q^*(s_\phi, o^*_{s_\phi})| \le \varepsilon_Q}$$

Expressive Model Options

$$|R_\gamma(s_\phi, o^*) - R_\gamma(s_\phi, o)| \le \varepsilon_R$$

and

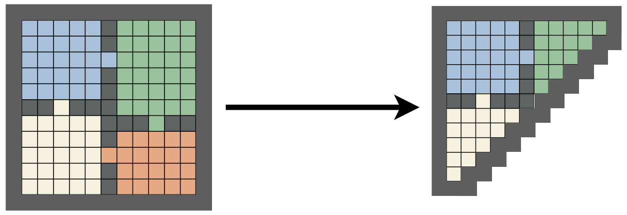
$$||T_\gamma(\cdot \mid s_\phi, o^*) - T_\gamma(\cdot \mid s_\phi, o)||_2 \le \varepsilon_T$$

Expressive k-Step Options

$$\boxed{\max_{s \in s_\phi, s' \in \mathcal{S}} |\mathbb{P}(s', k \mid s, o^*_{s_\phi}) - \mathbb{P}(s', k \mid s, o)| \le \tau}$$

[Nachum et al. 2019]

Homomorphism Options



[Ravindran and Barto '02, '03, '04]

Value-Preserving Abstractions

Theorem. $$\min_{\pi^{\Downarrow}_{\mathcal{O}_\phi} \in \Pi^{\Downarrow}_{\mathcal{O}_\phi}} \max_{s \in \mathcal{S}} \left(V^*(s) - V^{\pi^{\Downarrow}_{\mathcal{O}_\phi}}(s) \right) \leq \boxed{\eta_p},$$

Expressive Q^* Options

$$\boxed{\frac{\varepsilon_Q}{1-\gamma}}$$

Expressive Model Options

$$\boxed{\frac{\varepsilon_R + |\mathcal{S}|\varepsilon_T \text{R}\textsc{Max}}{(1-\gamma)^2}}$$

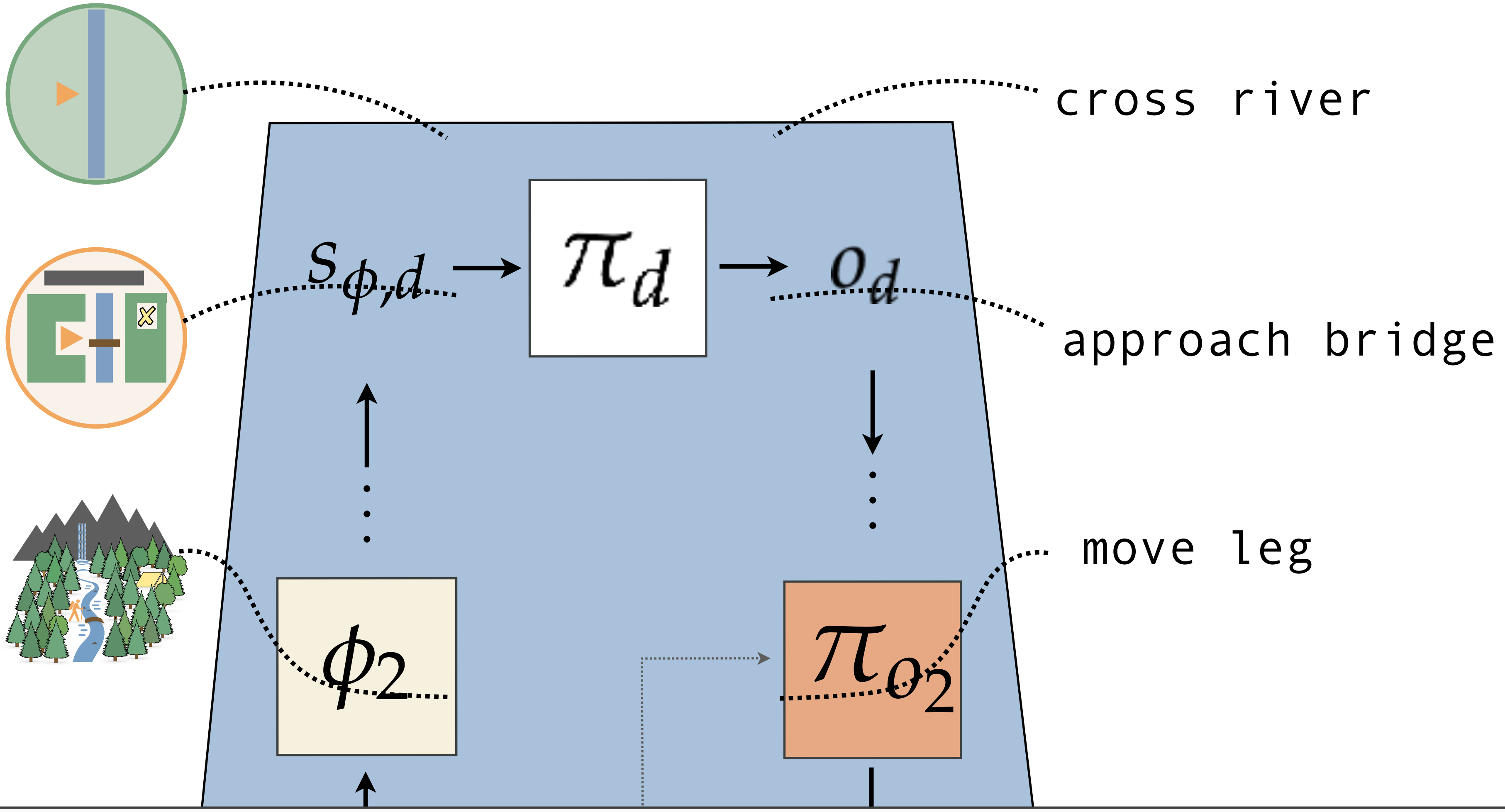
Expressive k-Step Options

$$\boxed{\frac{\tau\gamma|\mathcal{S}|}{(1-\gamma)^2}}$$

Homomorphism Options

$$\boxed{\frac{2}{1-\gamma}\left(\varepsilon_r + \frac{\gamma \text{R}\textsc{Max}}{1-\gamma}\frac{\varepsilon_p}{2}\right)}$$

Value-Preserving Hierarchies

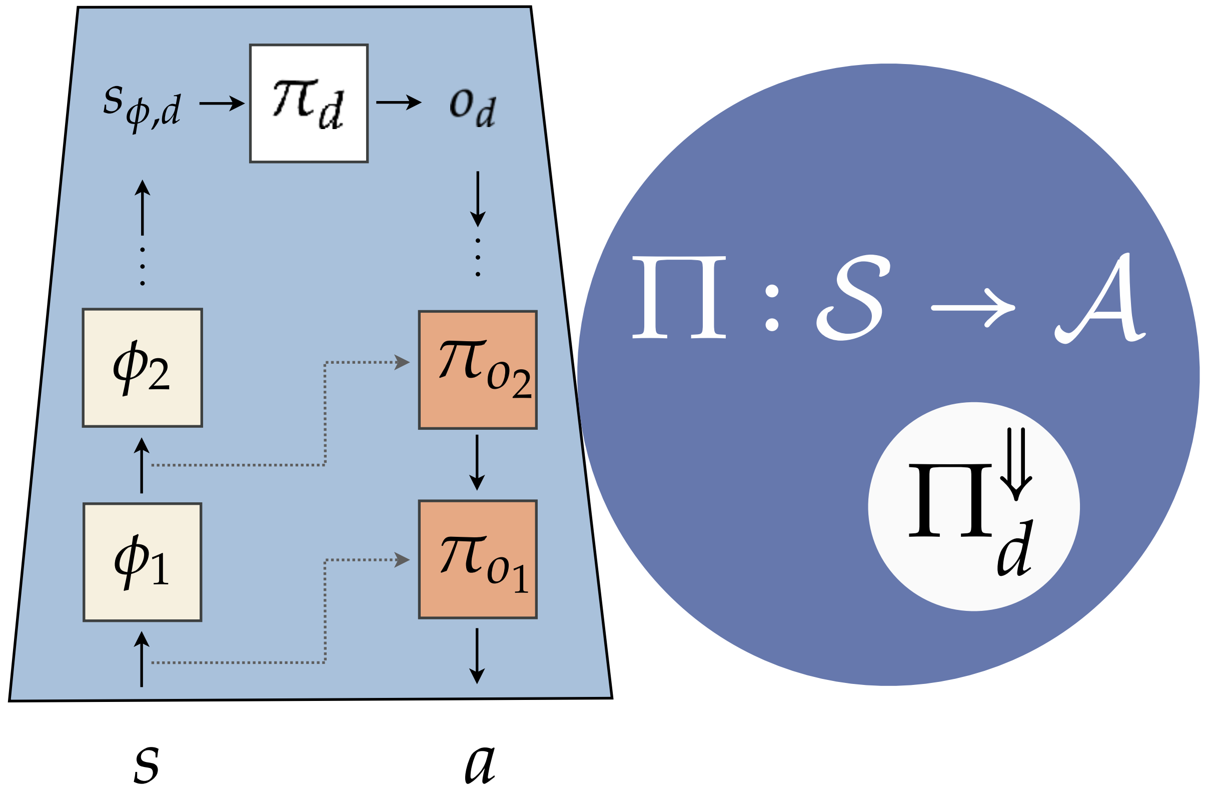


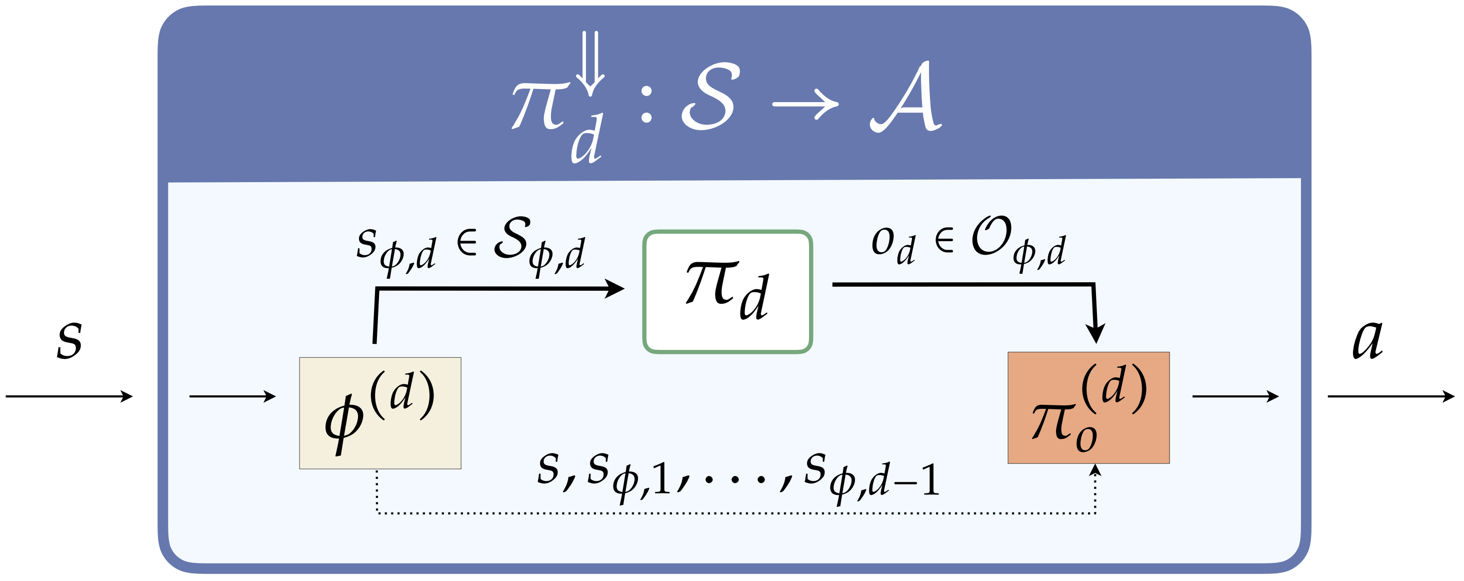
Definition. A depth d hierarchy H_d is defined by the pair

$$\phi^{(d)} = (\phi_1, \phi_2, \ldots, \phi_d),$$

$$\mathcal{O}^{(d)} = (\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_d).$$

Value-Preserving Hierarchies




$\pi_d^{\Downarrow} : \mathcal{S} \to \mathcal{A}$
$s_{\phi,d} \in \mathcal{S}_{\phi,d}$
π_d
$o_d \in \mathcal{O}_{\phi,d}$
s
$\phi^{(d)}$
$\pi_o^{(d)}$
a
$s, s_{\phi,1}, \ldots, s_{\phi,d-1}$

Value-Preserving Hierarchies

Assumption 1. *The value function is consistent throughout the hierarchy.*

▯ **value expressivity**

Assumption 2. *Subsequent levels of the hierarchy can represent policies similar in value to the best policy at the previous level.*

▯ **policy expressivity**

Value-Preserving Hierarchies

value expressivity

Assumption 1. *The value function is consistent throughout the hierarchy.*

policy expressivity

Assumption 2. *Subsequent levels of the hierarchy can represent policies similar in value to the best policy at the previous level.*

Theorem. Any hierarchy H_d that satisfies Assumptions 1 and 2 has bounded value loss:

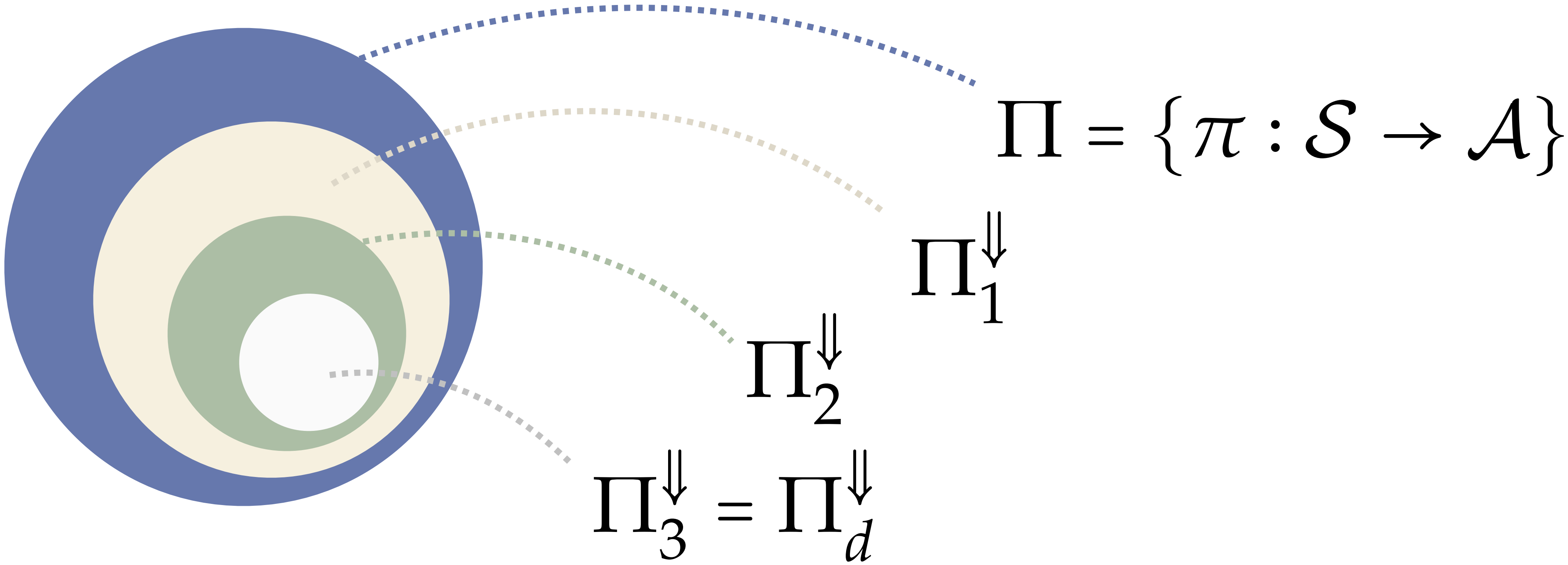
value expressivity

$$\min_{\pi_d^{\Downarrow}\in\Pi_d^{\Downarrow}} \max_{s\in\mathcal{S}} \left(V^*(s) - V^{\pi_d^{\Downarrow}}(s)\right) \leq \boxed{d}(\boxed{\kappa}+\boxed{\ell})$$

depth

policy expressivity

Value-Preserving Hierarchies



$$\Pi = \{\pi : \mathcal{S} \rightarrow \mathcal{A}\}$$

$$\Pi_1^{\Downarrow}$$

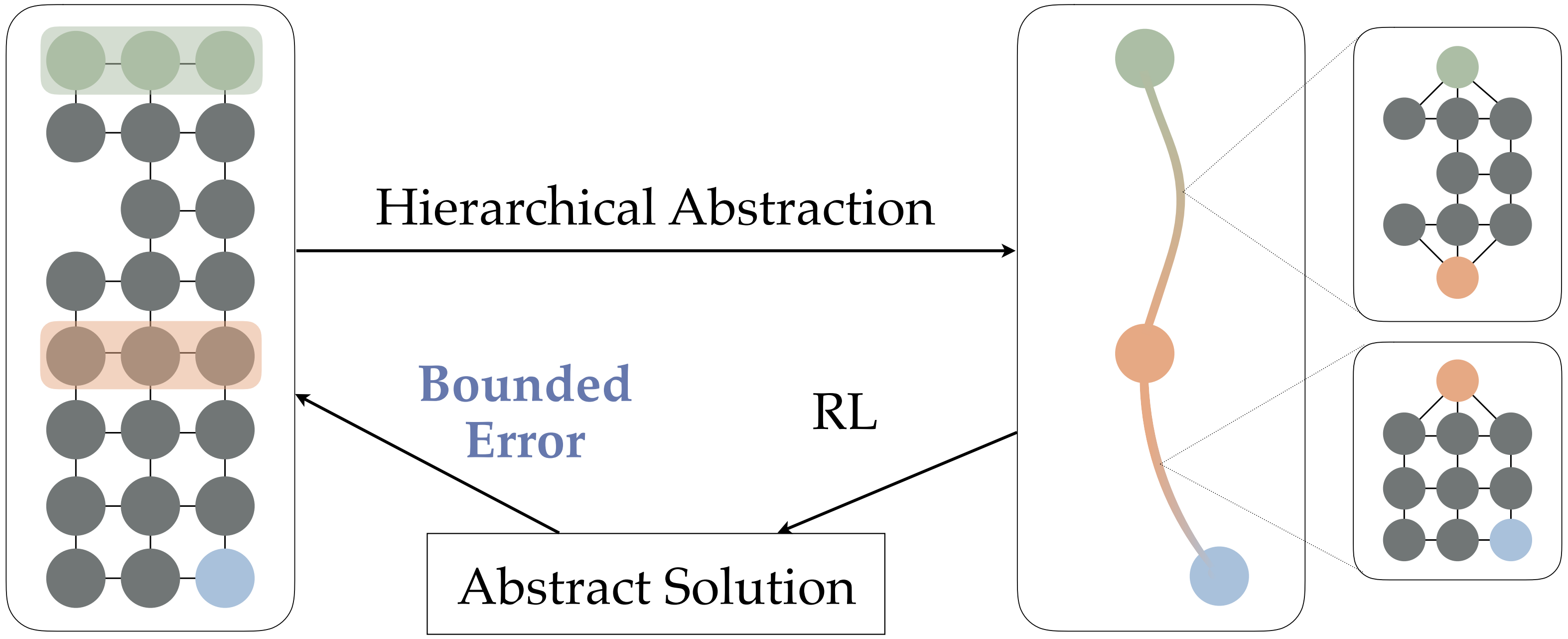
$$\Pi_2^{\Downarrow}$$

$$\Pi_3^{\Downarrow} = \Pi_d^{\Downarrow}$$

Theorem. Any hierarchy H_d that satisfies Assumptions 1 and 2 has bounded value loss:

$$\min_{\pi_d^{\Downarrow} \in \Pi_d^{\Downarrow}} \max_{s \in \mathcal{S}} \left(V^*(s) - V^{\pi_d^{\Downarrow}}(s) \right) \leq \underset{\text{depth}}{\boxed{d}}\left(\overset{\text{value expressivity}}{\boxed{\kappa}} + \underset{\text{policy expressivity}}{\boxed{\ell}}\right)$$

Value-Preserving Hierarchies



Hierarchical Abstraction

Bounded Error

RL

Abstract Solution

Theorem. Any hierarchy H_d that satisfies Assumptions 1 and 2 has bounded value loss:

$$\min_{\pi_d^{\Downarrow}\in\Pi_d^{\Downarrow}} \max_{s\in\mathcal{S}} \left(V^*(s) - V^{\pi_d^{\Downarrow}}(s)\right) \le \underset{\text{depth}}{\boxed{d}}\left(\overset{\text{value expressivity}}{\boxed{\kappa}}+\underset{\text{policy expressivity}}{\boxed{\ell}}\right)$$

Thanks to Mentors!

Masters

Ph.D

Undergrad

Joshua Schechter

Stefanie Tellex

Michael L. Littman

David Liben-Nowell

Ana Moltchanova

George Konidaris

Peter Stone

Will Dabney

Fernando Diaz

Owain Evans

Committee

Internships

Thanks to Collaborators!

Alekh
Agarwal

Cam
Allen

Dilip
Arumugam

Kavosh
Asadi

Gabriel
Barth-Maron

Stephen
Brawner

Jonathon
Cohen

Marie
desJardins

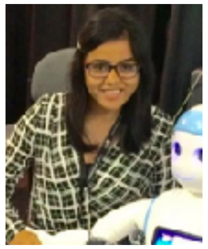
Tom
Griffiths

Yue
Guo

D. Ellis
Hershkowitz

Mark
Ho

Yuu
Jinnai


Khimya
Khetarpal

Akshay
Krishnamurthy

Lucas
Lehnert

James
MacGlashan

Jee Won
Park

Doina
Precup

Emily
Reif

Mark
Rowland

John
Salvatier

Robert
Schapire

Andreas
Stuhlmüeller

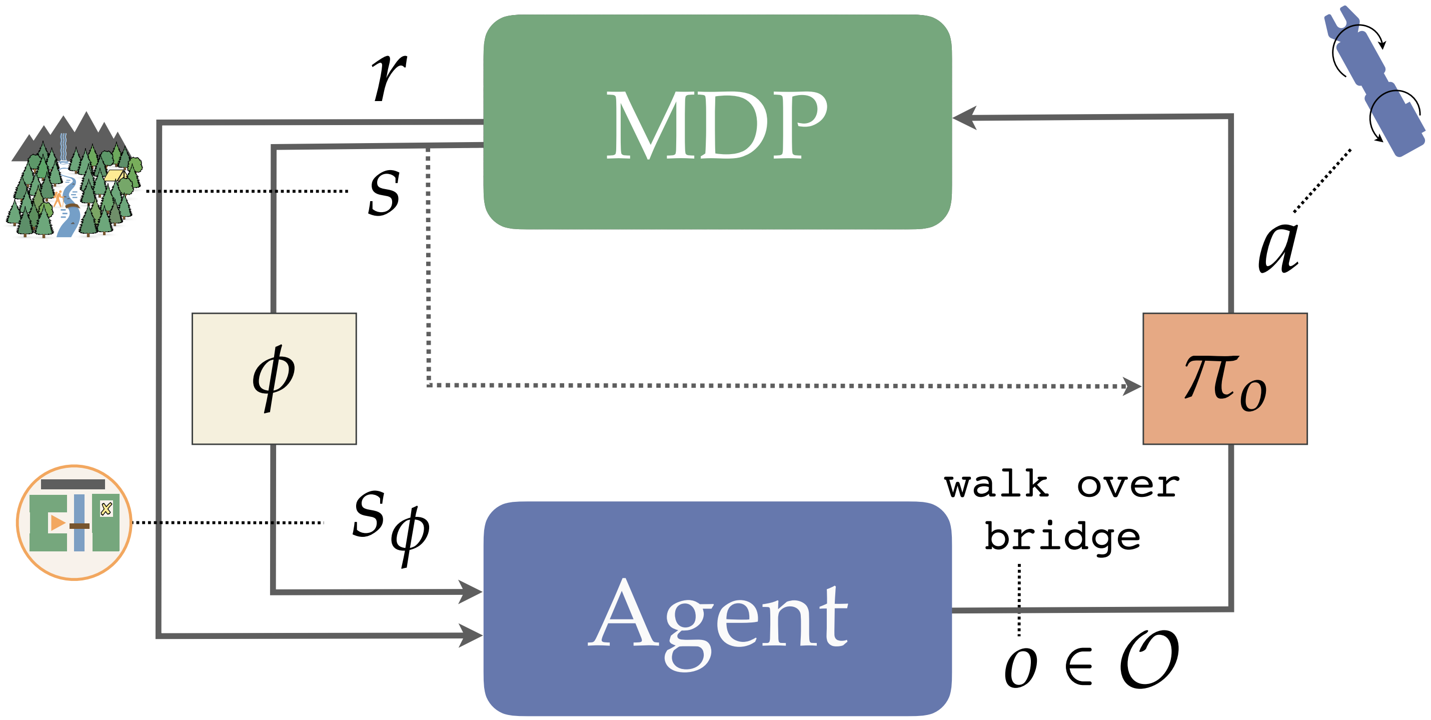
Nathan
Umbanhower

Edward
Williams

John
Winder

Lawson
Wong

Summary



Question: *How do effective RL* ***agents*** *come up with the right state and action abstractions of the* ***MDPs*** *they inhabit?*

Dissertation: david-abel.github.io/thesis.pdf

Contact: dmabel@deepmind.com