State Representations for Metrics in RL

Work done with Han Wang, Adam White and Martha White

Raksha Kumaraswamy

Outline

- Definition of metrics
- Motivate utility of metrics in State Representations
- Quick review of some RL terms
- Discuss desirable property of Representation Space
- Incorporate the property in Representation Space
- Empirical evidence for the above
- Conclusion

What are Metrics?

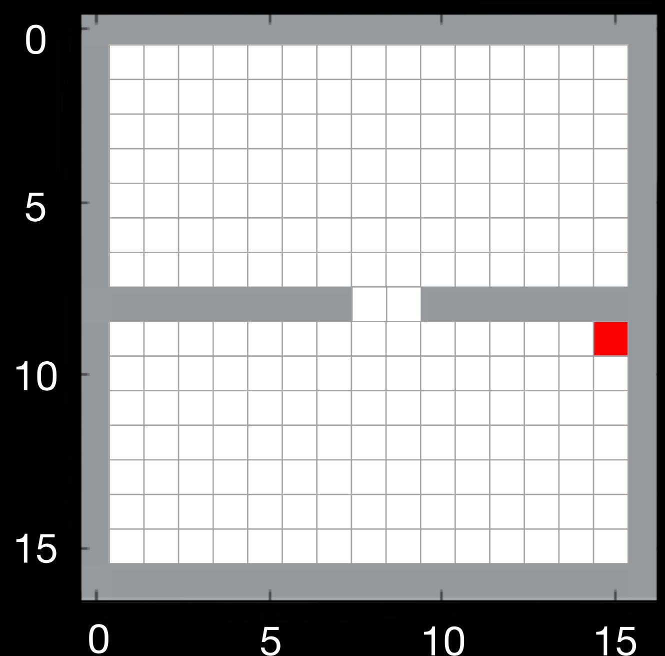
A metric/distance function, d, is a pairwise function over a set $\mathcal{X}$ which satisfies certain properties:

- Non-negativity: $d(x, x') \geq 0$
- Identity of indiscernibles: $d(x, x') = 0 \iff x = x'$
- Symmetry: $d(x, x') = d(x', x)$
- Triangle Inequality: $d(x, x'') \leq d(x, x') + d(x', x'')$

Popular metrics in ML: Mahalanobis distance, Minkowski distance, Euclidean distance etc.

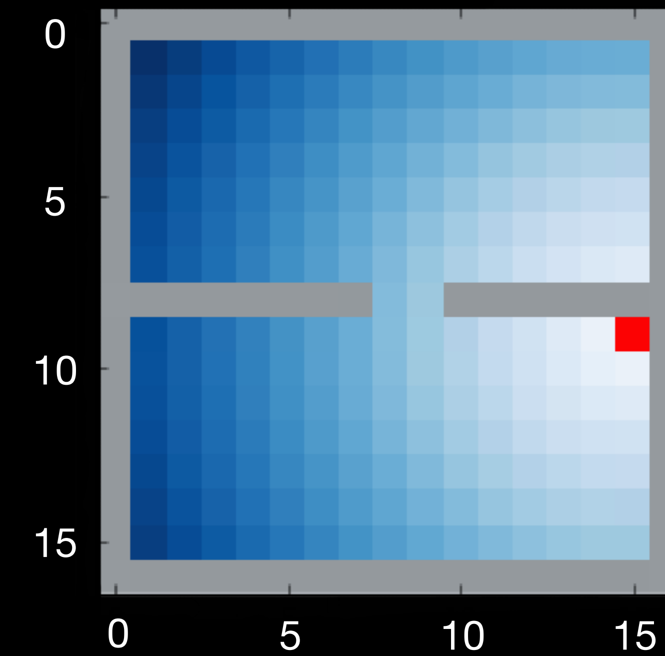
Sure. But, how are Metrics related to State Representations?

Motivating Example



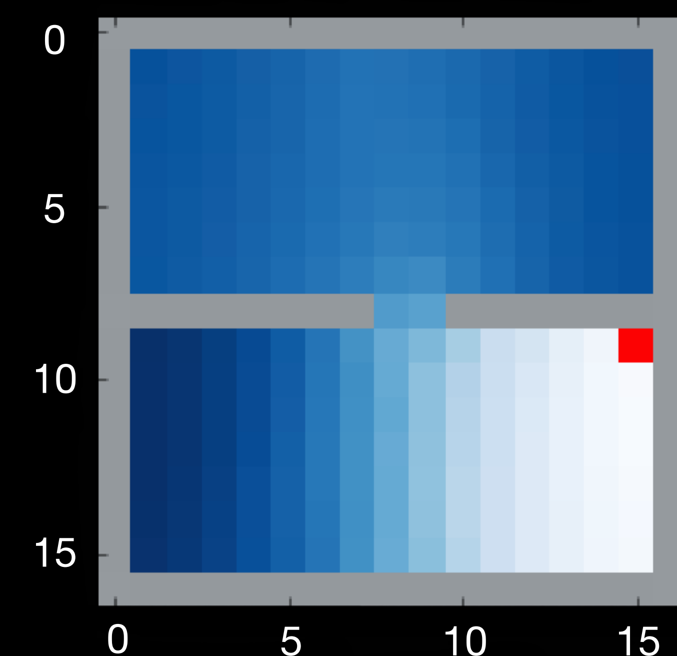
Distances in Observation space

$d(x_i, x_g)$



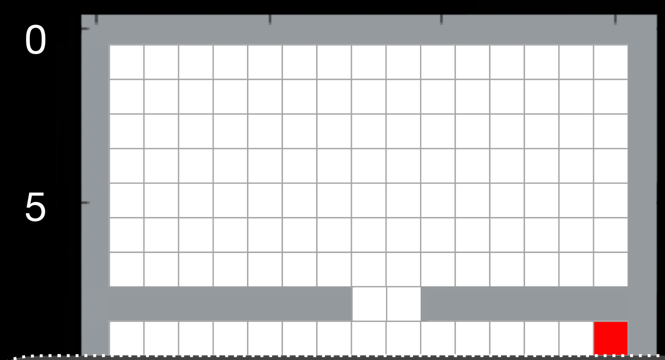
Distances in Representation space

$d(\Phi(x_i), \Phi(x_g))$



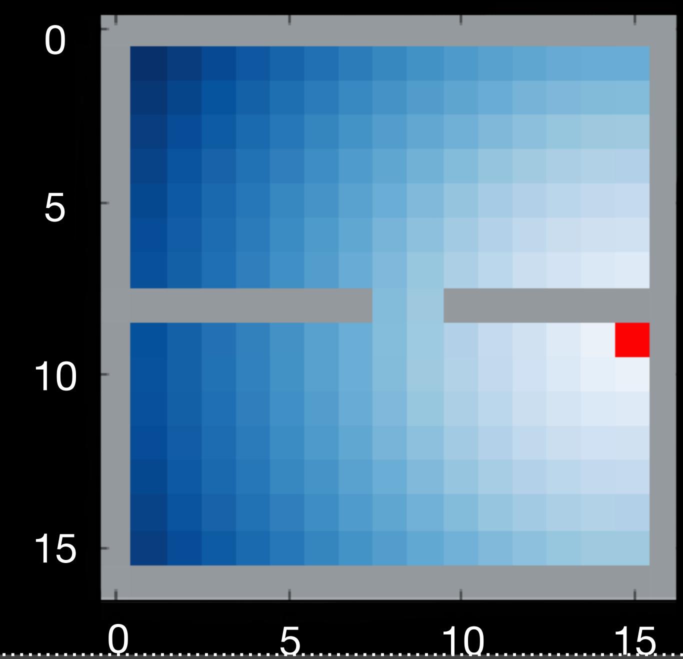
Plots from [Wu et. al, 2018]

Motivating Example



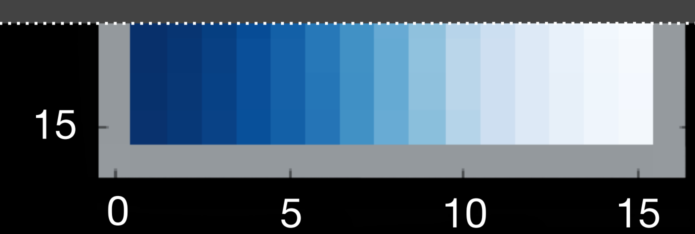
Distances in Observation space

$d(x_i, x_g)$



Here, Φ, captures the geometry of the problem making regular metrics — like Euclidean distance — useful.

$d(\Phi(x_i), \Phi(x_g))$



Plots from [Wu et. al, 2018]

What was this magical Φ encoding?

The Laplacian

A space where states close in time are embedded to be close, and states far in time and embedded to be far.

What was this magical Φ encoding?

Recent work that makes learning them incrementally feasible:
***"The Laplacian in RL: Learning Representations with Efficient Approximations"*, Wu et. al, ICLR, 2018.**

The Laplacian

A space where states close in time are embedded to be close, and states far in time and embedded to be far.

So should Φ be based only on transition dynamics for RL agents?

Arguably not.. So what would other good options be?

And more importantly, do we need a different *learning algorithm* for each?

So should Φ be based only on transition dynamics for RL agents?

Key questions:

- How do we learn a good Φ, such that the usual *d*'s are meaningful?

- Can we learn all Φ's the same way despite the metric they are trying to reflect?

her
be?

ed a
ch?

Some RL details

Given a d-dimensional representation space $\phi \in \Re^{d \times 1}$:

Successor Feature

$$\Psi_{\pi,i}(s) = \mathbb{E}_{\pi,P}\left[\sum_{t=0}^{\infty} \gamma^t \phi_i(S_t) \middle| S_0 = s\right]$$

Sample of Successor Feature

$$\Psi^{\sim}_{\pi,i}(s) = \left[\sum_{t=0}^{\infty} \gamma^t \phi_i(S_t) \middle| S_0 = s, \pi, P\right]$$

Successor Features

$$\Psi_{\pi}(s) = [\Psi_{\pi,0}(s), \Psi_{\pi,1}(s), \ldots, \Psi_{\pi,d}(s)]^T$$

Sample of Successor Features

$$\Psi^{\sim}_{\pi}(s) = [\Psi^{\sim}_{\pi,0}(s), \Psi^{\sim}_{\pi,1}(s), \ldots, \Psi^{\sim}_{\pi,d}(s)]^T$$

State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

State Representations for Metrics via Supervision

Let’s assume w … pace (Ψ).

Supervision space?

A space that has good metric properties.

State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

State Representations for Metrics via Supervision

Let's assume we … pace (Ψ).

With this, we w… e (Φ), that captures the m…

Why don't we just use the supervision space as the representation space?

- generalization
- dimensionality
- availability

State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

- Represents observations that are close in the supervision space (Ψ), to be close in the representation space (Φ).

State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

- Represe ce (Ψ), to be clo

Turns out, learning top eigenvectors (or left singular vectors) of Ψ (or $\Psi\Psi^T$) as Φ would be great for this.

State Representations for Metrics via Supervision

Let’s assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

- Represe … ce (Ψ), to be clo …

But, how do we do that?

State Representations for Metrics via Supervision

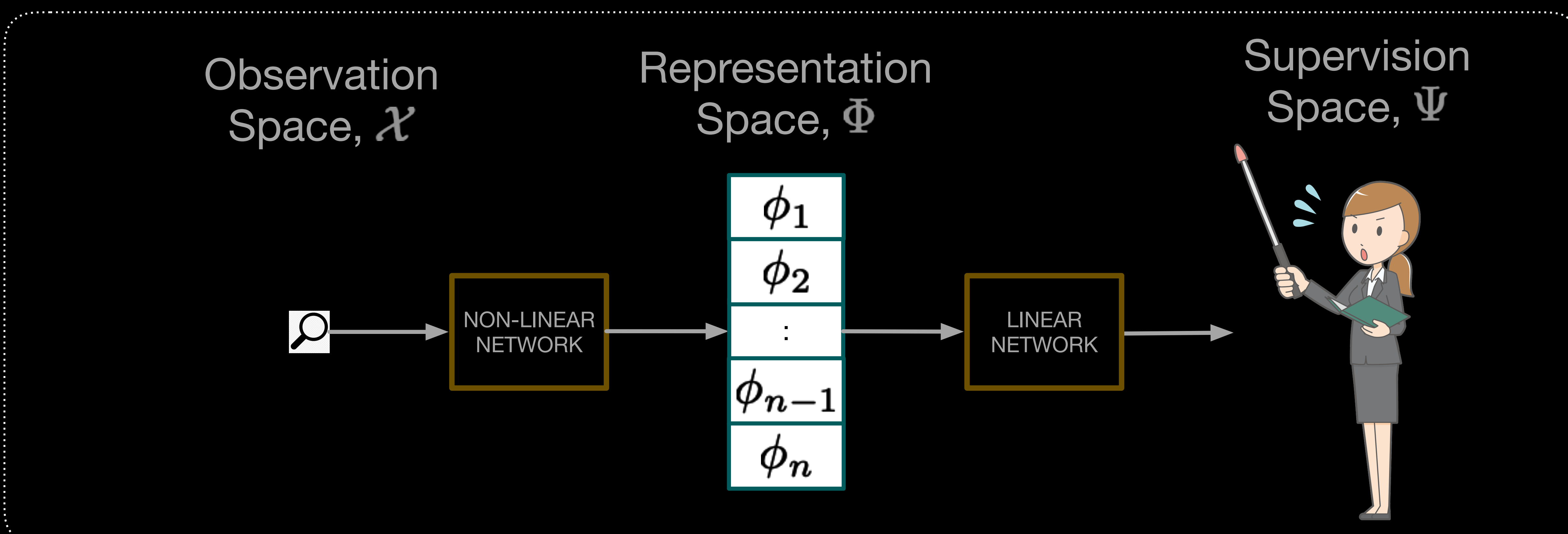
Let's assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.



State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

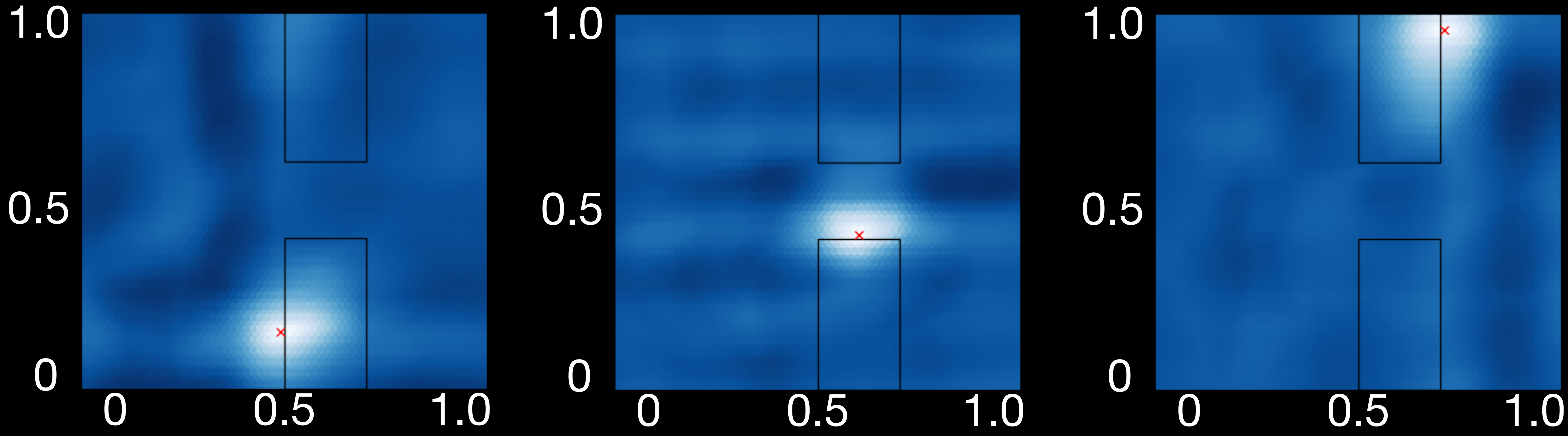
With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

The loss optimized with backprop:

$$\min_W \min_\Phi E_{u\sim\rho_\pi, y\sim P_\pi^\psi(.|u)}\left[||y - W(\Phi(u))||_2^2\right] \\ +\beta E_{u\sim\rho_\pi, v\sim\rho_\pi}\left[(\phi(u)^T\phi(v))^2 - ||\phi(u)||_2^2 - ||\phi(v)||_2^2\right]$$

Great! What do the results look like for this Φ?



But, the wall.. why does that happen?
.. Neural networks *generalize.*

Okay, how do we prevent it?

State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

- Represents observations that are close in the supervision space (Ψ), to be close in the representation space (Φ).
- Takes into account observations that have not been seen.

State Representations for Metrics via Supervision

Let'

With
cap

We

- R),
  to

- Ta

Maybe by negative sampling/hallucinating observations?

That is, force the neural net optimization to account for these states.

State Representations for

Le

Wi

ca

We

•

•

How does that modify the loss?

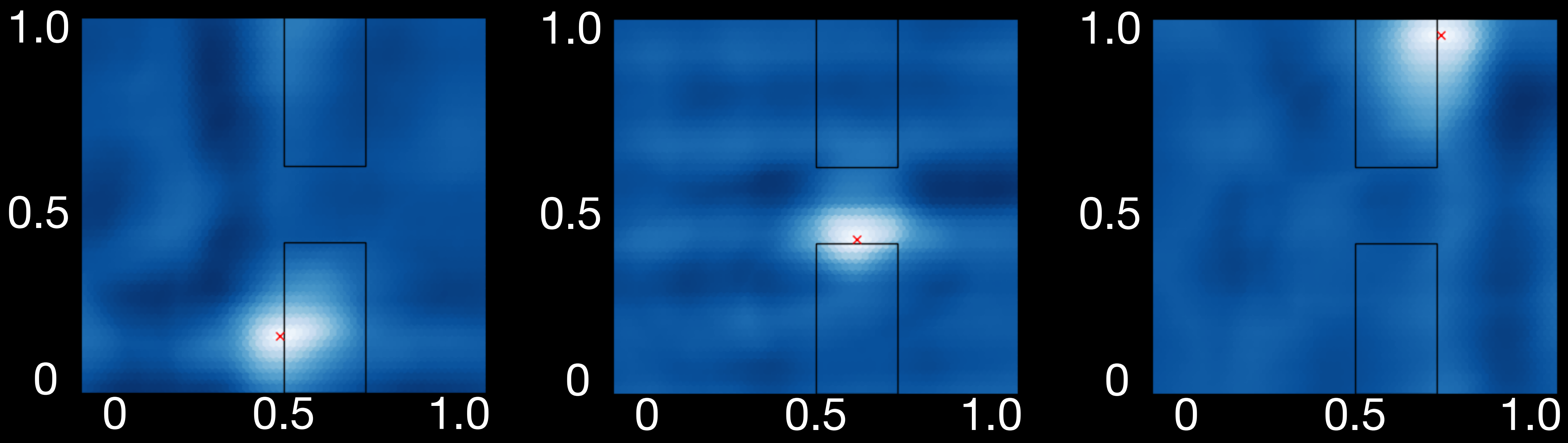
Changes

$$\min_W \min_\Phi E_{u\sim\rho_\pi, y\sim P_\pi^\psi(.|u)}\left[||y - W(\Phi(u))||_2^2\right]$$
$$+\beta E_{u\sim\rho_\pi, v\sim\rho_\pi}\left[(\phi(u)^T\phi(v))^2 - ||\phi(u)||_2^2 - ||\phi(v)||_2^2\right]$$

to

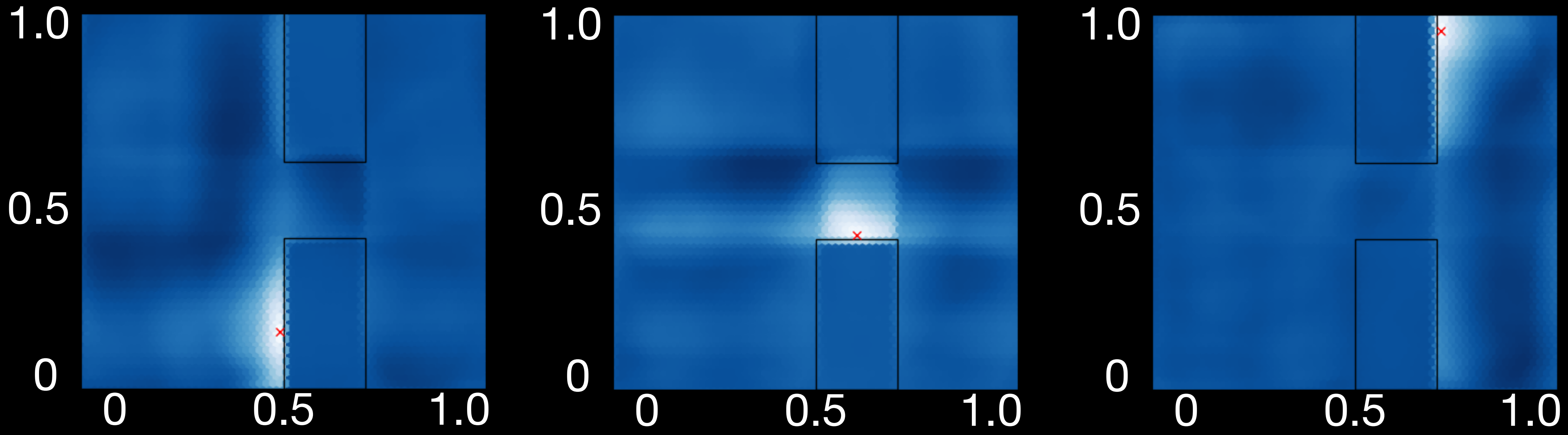
$$\min_W \min_\Phi E_{u\sim\rho_\pi, y\sim P_\pi^\psi(.|u)}\left[||y - W(\Phi(u))||_2^2\right]$$
$$+\beta E_{u\sim\rho_\pi, v\sim\rho}\left[(\phi(u)^T\phi(v))^2 - ||\phi(u)||_2^2 - ||\phi(v)||_2^2\right]$$

Does this Φ account for the walls?



Without negative sampling

Does this Φ account for the walls?



With negative sampling

(Maybe) Interesting. But how does an interactive agent have access to good supervision spaces?

One solution may be to use supervision spaces that can be learned incrementally too?

State Representations for Metrics via Supervision

Let's assume we have access to a good supervision space (Ψ).

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

- Represents observations that are close in the supervision space (Ψ), to be close in the representation space (Φ).
- Takes into account observations that have not been seen.

State Representations for Metrics via Supervision

~~Let's assume we have access to a good supervision space (Ψ).~~

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

- Represents observations that are close in the supervision space (Ψ), to be close in the representation space (Φ).
- Takes into account observations that have not been seen.

State Representations for Metrics via Supervision

Use a supervision space that is amenable to incremental learning.

With this, we would like to learn a representation space (Φ), that captures the metric in Ψ.

We would like a Φ that:

- Represents observations that are close in the supervision space (Ψ), to be close in the representation space (Φ).
- Takes into account observations that have not been seen.

State Representations for

How does that modify the loss?

Changes

$$\min_W \min_\Phi E_{u\sim\rho_\pi, y\sim P_\pi^\psi(.|u)}\left[||y - W(\Phi(u))||_2^2\right]$$
$$+\beta E_{u\sim\rho_\pi, v\sim\rho}\left[(\phi(u)^T\phi(v))^2 - ||\phi(u)||_2^2 - ||\phi(v)||_2^2\right]$$

to

$$\min_W \min_\Phi E_{u\sim\rho_\pi, v\sim P_\pi(.|u)}\left[||\Psi(u) - W(\Phi(u))||_2^2\right]$$
$$+\beta E_{u\sim\rho_\pi, v\sim\rho}\left[(\phi(u)^T\phi(v))^2 - ||\phi(u)||_2^2 - ||\phi(v)||_2^2\right]$$

where, $\Psi(u) = \Omega(u) + \gamma W(\Phi(v))$, TD target!

Conclusions

- ✓ Compact representation of information
- ✓ Possible to use multiple supervision signals
- ✓ Amenable for incremental learning (depending on the supervision)

- Access to good supervision information.
- Negative sampling in higher dimensions?
- Are representations with these properties sufficient?

Thank you!