Novelty Assessment Report

Paper: TD-JEPA: Latent-predictive Representations for Zero-Shot Reinforcement Learning
PDF URL: https://openreview.net/pdf?id=SzXDuBN8M1
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-27

Abstract

Latent prediction–where agents learn by predicting their own latents–has emerged as a powerful paradigm for training general representations in machine learning. In reinforcement learning (RL), this approach has been explored to define auxiliary losses for a variety of settings, including reward-based and unsupervised RL, behavior cloning, and world modeling. While existing methods are typically limited to single-task learning, one-step prediction, or on-policy trajectory data, we show that temporal difference (TD) learning enables learning representations predictive of long-term latent dynamics across multiple policies from offline, reward-free transitions. Building on this, we introduce TD-JEPA, which leverages TD-based latent-predictive representations into unsupervised RL. TD-JEPA trains explicit state and task encoders, a policy-conditioned multi-step predictor, and a set of parameterized policies directly in latent space. This enables zero-shot optimization of any reward function at test time. Theoretically, we show that an idealized variant of TD-JEPA avoids collapse with proper initialization, and learns encoders that capture a low-rank factorization of long-term policy dynamics, while the predictor recovers their successor features in latent space. Empirically, TD-JEPA matches or outperforms state-of-the-art baselines on locomotion, navigation, and manipulation tasks across 13 datasets in ExoRL and OGBench, especially in the challenging setting of zero-shot RL from pixels.

Core Task Landscape

This paper addresses: **Learning Latent-Predictive Representations for Zero-Shot Reinforcement Learning**

A total of **50 papers** were analyzed and organized into a taxonomy with **17 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Latent Dynamics Prediction and World Modeling**
- **Generalization and Transfer Across Tasks and Domains**
- **Reward-Predictive and Task-Conditioned Representations**
- **Unsupervised and Self-Supervised Representation Learning**
- **Cross-Modal and Multi-Modal Representation Learning**
- **Latent Space Alignment and Communication**
- **Domain-Specific Applications and Extensions**
- **Uncertainty Quantification and Robustness**

Complete Taxonomy Tree

- Learning Latent-Predictive Representations for Zero-Shot Reinforcement Learning Survey Taxonomy
- Latent Dynamics Prediction and World Modeling
  - Self-Predictive Latent Representations ★ (5 papers)
  - [0] TD-JEPA: Latent-predictive Representations for Zero-Shot Reinforcement Learning (Anon et al., 2026) View paper
  - [1] Self-Predictive Representations for Combinatorial Generalization in Behavioral Cloning (Lawson Daniel, 2025) View paper
  - [2] Data-Efficient Reinforcement Learning with Self-Predictive Representations (Schwarzer, 2020) View paper
  - [43] Bootstrap Latent-Predictive Representations for Multitask Reinforcement Learning (Zhaohan Daniel Guo, 2022) View paper
  - [45] Disentangled Predictive Representation for Meta-Reinforcement Learning (Sephora Madjiheurem, 2021) View paper
  - World Model-Based Planning and Control (5 papers)
  - [7] Dynamics-Aligned Latent Imagination in Contextual World Models for Zero-Shot Generalization (RÃ¶der, 2025) View paper
  - [12] Learning from Reward-Free Offline Data: A Case for Planning with Latent Dynamics Models (Sobal, 2025) View paper
  - [17] Latent imagination facilitates zero-shot transfer in autonomous racing (Brunnbauer, 2022) View paper
  - [25] LUMOS: Language-Conditioned Imitation Learning with World Models (Iman Nematollahi, 2025) View paper
  - [31] Procedural generalization by planning with self-supervised world models (Anand, 2021) View paper
  - Reward-Free and Passive Data Learning (2 papers)
  - [3] Reinforcement learning from passive data via latent intentions (Ghosh, 2023) View paper
  - [5] Regularized latent dynamics prediction is a strong baseline for behavioral foundation models (P Jajoo, 2025) View paper
- Generalization and Transfer Across Tasks and Domains
  - Cross-Task and Multi-Task Generalization (5 papers)
  - [4] Towards generalizable reinforcement learning via causality-guided self-adaptive representations (Yang Yupei, 2024) View paper
  - [15] The Role of Diverse Replay for Generalisation in Reinforcement Learning (Weltevrede, 2023) View paper
  - [24] Foundation Policies with Hilbert Representations (Park, 2024) View paper
  - [27] Cross-Trajectory Representation Learning for Zero-Shot Generalization in RL (Mazoure, 2021) View paper
  - [34] PI-QT-Opt: Predictive Information Improves Multi-Task Robotic Reinforcement Learning at Scale (Lee, 2022) View paper
  - Domain and Context Adaptation (4 papers)

- [8] Context-Based Meta-Reinforcement Learning With Bayesian Nonparametric Models (Zhenshan Bing, 2024) View paper
    - [11] In-Context Reinforcement Learning via Communicative World Models (Li Tao, 2025) View paper
    - [38] Observations Meet Actions: Learning Control-Sufficient Representations for Robust Policy Generalization (Gu Yuliang, 2025) View paper
    - [48] Zero-shot Domain Adaptation without Domain Semantic Descriptors (Kumagai, 2018) View paper
    - Cross-Embodiment and Morphology Transfer (2 papers)
    - [9] Cross-embodiment robot manipulation skill transfer using latent space alignment (Wang Tianyu, 2024) View paper
    - [16] Heterogeneous multi-agent reinforcement learning for zero-shot scalable collaboration (Xudong Guo, 2025) View paper
    - Zero-Shot Visual and Perceptual Generalization (3 papers)
    - [10] Diverse Policy Learning via Random Obstacle Deployment for Zero-Shot Adaptation (Seokjin Choi, 2025) View paper
    - [20] Zero Shot Generalization of Vision-Based RL Without Data Augmentation (Batra, 2024) View paper
    - [28] Improving Zero-Shot Generalization in Reinforcement Learning Through Abstract Representations (Ichise, 2025) View paper
- Reward-Predictive and Task-Conditioned Representations
    - Reward-Predictive Successor Representations (2 papers)
    - [39] Successor Clusters: A Behavior Basis for Unsupervised Zero-Shot Reinforcement Learning (L Bagot, 2025) View paper
    - [42] Reward-predictive representations generalize across tasks in reinforcement learning. (Lucas Lehnert, 2021) View paper
    - Task and Goal-Conditioned Representations (3 papers)
    - [30] DiactTOD: Learning Generalizable Latent Dialogue Acts for Controllable Task-Oriented Dialogue Systems (Qingyang Wu, 2023) View paper
    - [44] Zero-Shot Skill Composition and Simulation-to-Real Transfer by Learning Task Representations (Zhanpeng He, 2018) View paper
    - [50] ZAS-F: A Zero-Shot Abstract Sub-Goal Framework Empowers Robots for Long Horizontal Inventory Tasks (Yongshuai Wu, 2025) View paper
- Unsupervised and Self-Supervised Representation Learning
    - Contrastive and Predictive Coding (2 papers)
    - [19] Representation Learning with Contrastive Predictive Coding (Oord, 2018) View paper
    - [26] Zero-shot Musical Stem Retrieval with Joint-Embedding Predictive Architectures (Alain Riou, 2024) View paper
    - Behavioral Foundation Models and Skill Discovery (2 papers)
    - [32] Zero-Shot Whole-Body Humanoid Control via Behavioral Foundation Models (Tirinzoni, 2025) View paper
    - [33] Zero-Shot Imitation Policy Via Search In Demonstration Dataset (Leopold, 2024) View paper
- Cross-Modal and Multi-Modal Representation Learning
    - Vision-Language Alignment for Policy Learning (2 papers)
    - [13] From Foresight to Forethought: VLM-In-the-Loop Policy Steering via Latent Alignment (Yilin Wu, 2025) View paper
    - [22] Zero-shot model-based reinforcement learning using large language models (Thomas Albert, 2024) View paper
    - Semantic Attribute-Based Zero-Shot Learning (4 papers)
    - [6] Self-supervised embedding for generalized zero-shot learning in remote sensing scene classification (Rambabu Damalla, 2023) View paper
    - [18] Learning Modality-Invariant Latent Representations for Generalized Zero-shot Learning (Jingjing Li, 2020) View paper
    - [36] Zero-Shot Visual Recognition via Bidirectional Latent Embedding (Wang Qian, 2016) View paper
    - [47] Learning MLatent Representations for Generalized Zero-Shot Learning (Yalan Ye, 2022) View paper
- Latent Space Alignment and Communication (2 papers)
    - [46] Relative representations enable zero-shot latent space communication (Moschella, 2022) View paper
    - [49] Learning Nearest Neighbour Informed Latent Word Embeddings to Improve Zero-Shot Machine Translation (Nishant Kambhatla, 2023) View paper
- Domain-Specific Applications and Extensions
    - Robotics and Manipulation (2 papers)
    - [35] Zero-shot Safety Prediction for Autonomous Robots with Foundation World Models (Ivan, 2024) View paper
    - [40] Grammarization-Based Grasping with Deep Multi-Autoencoder Latent Space Exploration by Reinforcement Learning Agent (Askianakis, 2024) View paper
    - Non-RL Domains (5 papers)
    - [21] StereoCrafter-Zero: Zero-Shot Stereo Video Generation with Noisy Restart (Shi Jian, 2024) View paper
    - [23] Neuro-symbolic Zero-Shot Code Cloning with Cross-Language Intermediate Representation (Hasija, 2023) View paper
    - [29] CSLP-AE: A Contrastive Split-Latent Permutation Autoencoder Framework for Zero-Shot Electroencephalography Signal Conversion (MÃ¸rup, 2023) View paper
    - [37] Accelerating Bayesian inference of structural damage detection based on physics-guided CNN and zero-shot transfer learning (Rongrong Hou, 2025) View paper
    - [41] Zero-Shot Learning for S&P 500 Forecasting via Constituent-Level Dynamics: Latent Structure Modeling Without Index Supervision (Yoonjae Noh, 2025) View paper
- Uncertainty Quantification and Robustness (1 papers)
    - [14] Uncertainties of Latent Representations in Computer Vision (Kirchhof, 2024) View paper

Narrative

Core task: learning latent-predictive representations for zero-shot reinforcement learning. This field centers on building compact latent encodings that capture predictive structure in sequential decision problems, enabling agents to generalize to novel tasks or environments without task-specific fine-tuning. The taxonomy reveals several complementary research directions. Latent Dynamics Prediction and World Modeling focuses on learning forward models that simulate future states in latent space, often through self-predictive objectives as seen in Self-Predictive Representations[2] and Bootstrap Latent-Predictive[43]. Generalization and Transfer Across Tasks and Domains addresses how learned representations can be reused across different problem instances, while Reward-Predictive and Task-Conditioned Representations emphasize encoding goal-relevant information. Unsupervised and Self-Supervised Representation Learning explores methods like Contrastive Predictive Coding[19] that extract structure without explicit reward signals. Cross-Modal and Multi-Modal Representation Learning tackles scenarios where agents must integrate diverse sensory inputs, and Uncertainty Quantification and Robustness examines how to handle distributional shift and model confidence.

Within the world modeling branch, a handful of works explore different strategies for self-prediction in latent space. Self-Predictive Combinatorial[1] investigates compositional structure, while Disentangled Predictive[45] aims to separate independent factors of variation. TD-JEPA[0] sits naturally in this cluster, emphasizing temporal-difference style objectives for learning predictive embeddings that support zero-shot transfer. Compared to Bootstrap Latent-Predictive[43], which relies on bootstrapping target networks, TD-JEPA[0] integrates temporal-difference learning more directly into the representation objective. Meanwhile, Regularized Latent Dynamics[5] highlights the importance of regularization to prevent overfitting in learned world models. These contrasting approaches reflect ongoing questions about how best to balance predictive accuracy, computational efficiency, and generalization: whether to prioritize disentanglement, compositional reasoning, or robust temporal consistency when building latent representations for zero-shot RL.

Related Works in Same Category

The following **4 sibling papers** share the same taxonomy leaf node with the original paper:

1. Self-Predictive Representations for Combinatorial Generalization in Behavioral Cloning

Authors: Lawson Daniel, Berseth, Glen, Khetarpal, Khimya | **Year/Venue**: 2025 | **URL**: View paper

Abstract

While goal-conditioned behavior cloning (GCBC) methods can perform well on in-distribution training tasks, they do not necessarily generalize zero-shot to tasks that require conditioning on novel state-goal pairs, i.e. combinatorial generalization. In part, this limitation can be attributed to a lack of temporal consistency in the state representation learned by BC; if temporally correlated states are properly encoded to similar latent representations, then the out-of-distribution gap for novel ...

Relationship Analysis

Both papers belong to the Self-Predictive Latent Representations category, training agents to predict future latent states for improved representation learning in reinforcement learning contexts. They overlap in using self-predictive objectives (JEPA-style architectures) to learn temporally consistent representations that capture dynamics, with both connecting their methods to successor representations. The key differences are that the original paper (TD-JEPA) focuses on zero-shot RL across multiple tasks using temporal-difference learning with policy-conditioned predictors for offline reward-free data, while the candidate paper (BYOL-γ) targets combinatorial generalization in behavioral cloning by predicting geometrically-sampled future states without TD learning or negative samples.

---

2. Data-Efficient Reinforcement Learning with Self-Predictive Representations

Authors: Schwarzer, Max | **Year/Venue**: 2020 • International Conference on Learning Representations | **URL**: View paper

Abstract

While deep reinforcement learning excels at solving tasks where large amounts of data can be collected through virtually unlimited interaction with the environment, learning from limited interaction remains a key challenge. We posit that an agent can learn more efficiently if we augment reward maximization with self-supervised objectives based on structure in its visual input and sequential interaction with the environment. Our method, Self-Predictive Representations(SPR), trains an agent to pre...

Relationship Analysis

Both papers belong to the Self-Predictive Latent Representations category, training agents to predict future latent states for improved reinforcement learning. SPR focuses on sample-efficient single-task RL from pixels by predicting future latent states with data augmentation as an auxiliary objective to reward maximization, achieving strong results on Atari with limited data. TD-JEPA differs by using temporal difference learning for multi-step, policy-conditioned latent prediction across multiple tasks from offline data, enabling zero-shot policy optimization for any downstream reward function without requiring reward signals during pre-training.

---

3. Bootstrap Latent-Predictive Representations for Multitask Reinforcement Learning

Authors: Zhaohan Daniel Guo, Bernardo Avila Pires, Bernardo Ãvila Pires, Mohammad Gheshlaghi Azar, Bilal Piot, et al. (8 authors total) | **Year/Venue**: 2022 • arXiv (Cornell University) | **URL**: View paper

Abstract

Learning a good representation is an essential component for deep reinforcement learning (RL). Representation learning is especially important in multitask and partially observable settings where building a representation of the unknown environment is crucial to solve the tasks. Here we introduce Prediction of Bootstrap Latents (PBL), a simple and flexible self-supervised representation learning algorithm for multitask deep RL. PBL builds on multistep predictive representations of future observa...

Relationship Analysis

Both papers belong to the Self-Predictive Latent Representations category, training agents to predict future latent states through bootstrapped prediction mechanisms. They overlap in using latent-predictive objectives to learn representations that capture environment dynamics without reconstruction, and both employ multi-step prediction horizons. The key difference is that TD-JEPA focuses on zero-shot RL through policy-conditioned TD learning with separate state and task encoders for offline training, while PBL uses bidirectional bootstrapping (forward and reverse predictions) as an auxiliary task for online multitask RL with a single representation.

---

4. Disentangled Predictive Representation for Meta-Reinforcement Learning

Authors: Sephora Madjiheurem, Laura Toni | **Year/Venue**: 2021 • International Conference on Machine Learning | **URL**: View paper

Abstract

N/A

Relationship Analysis

Both papers belong to the Self-Predictive Latent Representations category, focusing on learning representations by predicting future latent states in reinforcement learning. While TD-JEPA trains policy-conditioned multi-step predictors using temporal difference learning for zero-shot RL across multiple tasks from offline data, the candidate paper focuses on disentangled predictive representations specifically for meta-reinforcement learning contexts. The key difference lies in TD-JEPA's emphasis on zero-shot generalization through successor features and its off-policy TD formulation, versus the candidate's focus on disentanglement for meta-learning scenarios.

Contributions Analysis

Overall novelty summary. The paper introduces TD-JEPA, which applies temporal-difference learning to train latent-predictive representations for zero-shot reinforcement learning. It resides in the 'Self-Predictive Latent Representations' leaf, which contains five papers including the original work. This leaf sits within the broader 'Latent Dynamics Prediction and World Modeling' branch, indicating a moderately populated research direction focused on learning forward models in latent space. The sibling papers explore related themes such as compositional structure, disentanglement, and bootstrapping-based prediction, suggesting an active but not overcrowded subfield where different architectural and objective choices are still being explored.

The taxonomy reveals that TD-JEPA's leaf is adjacent to 'World Model-Based Planning and Control', which emphasizes model-predictive control rather than representation learning, and 'Reward-Free and Passive Data Learning', which focuses on learning from observational data without reward signals. The paper's emphasis on policy-conditioned multi-step prediction and zero-shot task adaptation also connects it to the 'Cross-Task and Multi-Task Generalization' branch, though it remains distinct by prioritizing latent dynamics over explicit task encoders. The taxonomy's scope and exclude notes clarify that TD-JEPA's focus on TD-based objectives differentiates it from planning-centric world models and from methods requiring reward signals during training.

Among 23 candidates examined across three contributions, none were flagged as clearly refuting the paper's claims. The first contribution (TD-based latent-predictive representations) examined three candidates with no refutations, suggesting limited prior work directly combining TD learning with policy-conditioned multi-step latent prediction. The second contribution (TD-JEPA algorithm) and third contribution (theoretical analysis) each examined ten candidates, again with no refutations. This indicates that within the limited search scope, the specific combination of TD objectives, explicit state and task encoders, and zero-shot optimization in latent space appears relatively unexplored, though the search scale is modest and may not capture all relevant prior work.

Based on the limited literature search of 23 candidates, TD-JEPA appears to occupy a distinct position within the self-predictive latent representations subfield. The absence of refutable prior work among examined candidates suggests novelty in its specific technical approach, though the search scope does not guarantee exhaustive coverage of related methods in successor features, world modeling, or unsupervised RL. The taxonomy context indicates the paper contributes to an active but not saturated research direction, where different strategies for learning predictive latent dynamics are still being actively developed and compared.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: TD-based latent-predictive representations for multi-step, policy-conditioned dynamics

Description: The authors introduce a novel temporal-difference loss for latent-predictive representation learning that models multi-step, policy-conditioned dynamics from offline data. Unlike prior methods limited to single-step prediction or on-policy data, this approach learns representations that capture long-term features relevant for value estimation across multiple policies.

This contribution was assessed against **3 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. TOPS: Transition-Based Volatility-Reduced Policy Search

URL: View paper

Brief Assessment

TOPS[69] focuses on risk-averse RL using reward volatility and transition-based learning, not on temporal difference learning for latent-predictive representations or multi-step policy-conditioned dynamics prediction from offline data.

---

2. Multi-agent LLMs with Offline Reinforcement Learning for Hierarchical Multi-turn Decision-making

URL: View paper

Brief Assessment

Multi-Agent LLMs[67] focuses on hierarchical multi-agent LLM architectures for multi-turn decision-making using offline RL (IQL-based), not on temporal-difference learning for latent-predictive representation learning or multi-step policy-conditioned dynamics prediction.

---

3. Uncertainty-driven exploration in sparse model-based reinforcement learning

URL: View paper

Brief Assessment

Uncertainty-Driven Exploration[68] focuses on uncertainty-driven exploration in sparse reward environments, not on temporal-difference learning for multi-step policy-conditioned latent dynamics prediction from offline data.

Contribution 2: TD-JEPA algorithm for zero-shot unsupervised RL

Description: The authors propose TD-JEPA, a zero-shot unsupervised RL algorithm that jointly trains state encoders, task encoders, policy-conditioned predictors, and parameterized policies end-to-end from offline reward-free transitions. The method enables zero-shot optimization of any reward function at test time entirely in latent space.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. ZSL-RPPO: Zero-Shot Learning for Quadrupedal Locomotion in Challenging Terrains using Recurrent Proximal Policy Optimization

URL: View paper

Brief Assessment

ZSL-RPPO[56] focuses on zero-shot learning for quadrupedal robot locomotion using recurrent policy optimization in partially observable environments, not on unsupervised RL with latent-predictive representations for general task optimization.

---

2. Dynamics-Aligned Latent Imagination in Contextual World Models for Zero-Shot Generalization

URL: View paper

Brief Assessment

Dynamics-Aligned Imagination[7] focuses on contextual MDPs with latent context inference for adaptation to environmental variations, not on zero-shot unsupervised RL with latent-predictive representations and successor features as in the original paper.

---

3. Data-driven latent space representation for robust bipedal locomotion learning

URL: View paper

Brief Assessment

Data-Driven Latent Space[55] focuses on bipedal locomotion control using autoencoders for dimensionality reduction, not zero-shot unsupervised RL with latent-predictive representations for multi-task policy optimization.

---

4. Learning to generalize with latent embedding optimization for few- and zero-shot cross domain fault diagnosis

URL: View paper

Brief Assessment

Latent Embedding Optimization[52] focuses on cross-domain fault diagnosis in industrial systems, not reinforcement learning or zero-shot policy optimization. The domains are entirely different.

---

5. Distributional Successor Features Enable Zero-Shot Policy Optimization

URL: View paper

Brief Assessment

Distributional Successor Features[57] focuses on learning distributions of successor features using diffusion models for zero-shot policy optimization, while TD-JEPA uses temporal difference learning for latent-predictive representations. The candidate does not demonstrate that similar TD-based latent-predictive zero-shot RL existed prior to the original work.

6. Right Question is Already Half the Answer: Fully Unsupervised LLM Reasoning Incentivization

URL: View paper

Brief Assessment

Unsupervised LLM Reasoning[51] focuses on unsupervised reasoning enhancement for large language models through entropy minimization, not on zero-shot reinforcement learning with latent-predictive representations for control tasks.

7. Diverse Policy Learning via Random Obstacle Deployment for Zero-Shot Adaptation

URL: View paper

Brief Assessment

Random Obstacle Deployment[10] focuses on zero-shot adaptation to dynamically changing obstacles through diverse policy learning, not on unsupervised RL with latent-predictive representations and offline reward-free training as in the original paper.

8. From Parameters to Behavior: Unsupervised Compression of the Policy Space

URL: View paper

Brief Assessment

Policy Space Compression[54] focuses on compressing policy parameter spaces into low-dimensional latent spaces for efficient fine-tuning, not on zero-shot unsupervised RL with latent-predictive representations and temporal difference learning for successor features.

9. Constrained Skill Discovery: Quadruped Locomotion with Unsupervised Reinforcement Learning

URL: View paper

Brief Assessment

Constrained Skill Discovery[53] focuses on quadruped locomotion using constrained mutual information maximization for skill discovery, not on latent-predictive TD learning or zero-shot reward optimization in latent space as proposed in the original paper.

10. CSLP-AE: A Contrastive Split-Latent Permutation Autoencoder Framework for Zero-Shot Electroencephalography Signal Conversion

URL: View paper

Brief Assessment

CSLP-AE[29] focuses on EEG signal conversion between subjects and tasks using contrastive learning for biological signals, not reinforcement learning or policy optimization in latent space.

Contribution 3: Theoretical analysis connecting TD-JEPA to successor features and policy evaluation

Description: The authors provide theoretical guarantees showing that TD-JEPA with linear predictors avoids representation collapse, recovers a low-rank factorization of successor measures, and minimizes an upper bound on policy evaluation error. These results build on a novel gradient matching argument that generalizes existing analyses of latent-predictive representations.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. A neurally plausible model learns successor representations in partially observable environments

URL: View paper

Brief Assessment

Neurally Plausible Successor[59] focuses on learning successor representations in partially observable environments using distributional codes for neural implementation, not on latent-predictive TD learning or zero-shot RL frameworks as in the original paper.

2. Policy-Oriented Cognitive Risk Map Modeling for Lane Change via Deep Successor Representation

URL: View paper

Brief Assessment

Cognitive Risk Map[63] focuses on risk assessment for autonomous driving using successor representations for modeling driver cognition, not on theoretical guarantees for latent-predictive representation learning or policy evaluation error bounds in general RL settings.

3. Learning structures: predictive representations, replay, and generalization

URL: View paper

Brief Assessment

Predictive Representations Replay[58] is a review paper focusing on cognitive maps and successor representations in neuroscience contexts. It does not present novel theoretical analyses connecting latent-predictive TD learning to successor features or policy evaluation error bounds.

4. Accelerating learning in constructive predictive frameworks with the successor representation

URL: View paper

Brief Assessment

Constructive Predictive Frameworks[66] focuses on using successor representations to accelerate learning of general value functions in a continual learning setting, not on latent-predictive representations or TD-based learning frameworks for zero-shot RL.

5. Learning Successor Feature Representations to Train Robust Policies for Multi-task Learning

URL: View paper

Brief Assessment

Robust Multi-Task Policies[61] focuses on learning successor feature representations for multi-task continuous control using actor-critic methods, without the temporal difference latent-predictive framework or the specific theoretical guarantees about representation collapse and gradient matching that TD-JEPA provides.

6. A New Representation of Universal Successor Features for Enhancing the Generalization of Target-Driven Visual Navigation

URL: View paper

Brief Assessment

Universal Successor Features[60] focuses on target-driven visual navigation with parsimonious dynamics models, not on theoretical guarantees for latent-predictive representations or policy evaluation error bounds in general RL settings.

7. Combining Behaviors with the Successor Features Keyboard

URL: View paper

Brief Assessment

Successor Features Keyboard[65] focuses on transfer learning by combining known behaviors using successor features in a different context (task transfer with discovered representations), not on theoretical guarantees for latent-predictive TD learning or policy evaluation error bounds.

8. Eigenoption discovery through the deep successor representation

URL: View paper

Brief Assessment

Eigenoption Discovery[64] focuses on discovering options through successor representations in a hierarchical RL framework, not on latent-predictive representations for zero-shot policy evaluation across multiple tasks as in the original paper.

9. Successor Clusters: A Behavior Basis for Unsupervised Zero-Shot Reinforcement Learning

URL: View paper

Brief Assessment

Successor Clusters[39] focuses on clustering state spaces for reward feature construction in zero-shot RL, not on latent-predictive TD learning or gradient matching arguments for representation collapse analysis.

10. Accounting for sensitivity of latent learning to behavioral statistics with successor representations

URL: View paper

Brief Assessment

Behavioral Statistics Sensitivity[62] focuses on how successor representations are affected by exploration statistics in spatial navigation tasks, not on theoretical guarantees for TD-based latent-predictive learning or policy evaluation error bounds.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] TD-JEPA: Latent-predictive Representations for Zero-Shot Reinforcement Learning View paper
- [1] Self-Predictive Representations for Combinatorial Generalization in Behavioral Cloning View paper
- [2] Data-Efficient Reinforcement Learning with Self-Predictive Representations View paper
- [3] Reinforcement learning from passive data via latent intentions View paper
- [4] Towards generalizable reinforcement learning via causality-guided self-adaptive representations View paper
- [5] Regularized latent dynamics prediction is a strong baseline for behavioral foundation models View paper
- [6] Self-supervised embedding for generalized zero-shot learning in remote sensing scene classification View paper
- [7] Dynamics-Aligned Latent Imagination in Contextual World Models for Zero-Shot Generalization View paper
- [8] Context-Based Meta-Reinforcement Learning With Bayesian Nonparametric Models View paper
- [9] Cross-embodiment robot manipulation skill transfer using latent space alignment View paper
- [10] Diverse Policy Learning via Random Obstacle Deployment for Zero-Shot Adaptation View paper
- [11] In-Context Reinforcement Learning via Communicative World Models View paper
- [12] Learning from Reward-Free Offline Data: A Case for Planning with Latent Dynamics Models View paper
- [13] From Foresight to Forethought: VLM-In-the-Loop Policy Steering via Latent Alignment View paper
- [14] Uncertainties of Latent Representations in Computer Vision View paper
- [15] The Role of Diverse Replay for Generalisation in Reinforcement Learning View paper
- [16] Heterogeneous multi-agent reinforcement learning for zero-shot scalable collaboration View paper
- [17] Latent imagination facilitates zero-shot transfer in autonomous racing View paper
- [18] Learning Modality-Invariant Latent Representations for Generalized Zero-shot Learning View paper
- [19] Representation Learning with Contrastive Predictive Coding View paper
- [20] Zero Shot Generalization of Vision-Based RL Without Data Augmentation View paper
- [21] StereoCrafter-Zero: Zero-Shot Stereo Video Generation with Noisy Restart View paper
- [22] Zero-shot model-based reinforcement learning using large language models View paper
- [23] Neuro-symbolic Zero-Shot Code Cloning with Cross-Language Intermediate Representation View paper
- [24] Foundation Policies with Hilbert Representations View paper
- [25] LUMOS: Language-Conditioned Imitation Learning with World Models View paper
- [26] Zero-shot Musical Stem Retrieval with Joint-Embedding Predictive Architectures View paper
- [27] Cross-Trajectory Representation Learning for Zero-Shot Generalization in RL View paper
- [28] Improving Zero-Shot Generalization in Reinforcement Learning Through Abstract Representations View paper
- [29] CSLP-AE: A Contrastive Split-Latent Permutation Autoencoder Framework for Zero-Shot Electroencephalography Signal Conversion View paper
- [30] DiactTOD: Learning Generalizable Latent Dialogue Acts for Controllable Task-Oriented Dialogue Systems View paper
- [31] Procedural generalization by planning with self-supervised world models View paper
- [32] Zero-Shot Whole-Body Humanoid Control via Behavioral Foundation Models View paper
- [33] Zero-Shot Imitation Policy Via Search In Demonstration Dataset View paper
- [34] PI-QT-Opt: Predictive Information Improves Multi-Task Robotic Reinforcement Learning at Scale View paper
- [35] Zero-shot Safety Prediction for Autonomous Robots with Foundation World Models View paper

- [36] Zero-Shot Visual Recognition via Bidirectional Latent Embedding View paper
- [37] Accelerating Bayesian inference of structural damage detection based on physics-guided CNN and zero-shot transfer learning View paper
- [38] Observations Meet Actions: Learning Control-Sufficient Representations for Robust Policy Generalization View paper
- [39] Successor Clusters: A Behavior Basis for Unsupervised Zero-Shot Reinforcement Learning View paper
- [40] Grammarization-Based Grasping with Deep Multi-Autoencoder Latent Space Exploration by Reinforcement Learning Agent View paper
- [41] Zero-Shot Learning for S&P 500 Forecasting via Constituent-Level Dynamics: Latent Structure Modeling Without Index Supervision View paper
- [42] Reward-predictive representations generalize across tasks in reinforcement learning. View paper
- [43] Bootstrap Latent-Predictive Representations for Multitask Reinforcement Learning View paper
- [44] Zero-Shot Skill Composition and Simulation-to-Real Transfer by Learning Task Representations View paper
- [45] Disentangled Predictive Representation for Meta-Reinforcement Learning View paper
- [46] Relative representations enable zero-shot latent space communication View paper
- [47] Learning MLatent Representations for Generalized Zero-Shot Learning View paper
- [48] Zero-shot Domain Adaptation without Domain Semantic Descriptors View paper
- [49] Learning Nearest Neighbour Informed Latent Word Embeddings to Improve Zero-Shot Machine Translation View paper
- [50] ZAS-F: A Zero-Shot Abstract Sub-Goal Framework Empowers Robots for Long Horizontal Inventory Tasks View paper
- [51] Right Question is Already Half the Answer: Fully Unsupervised LLM Reasoning Incentivization View paper
- [52] Learning to generalize with latent embedding optimization for few- and zero-shot cross domain fault diagnosis View paper
- [53] Constrained Skill Discovery: Quadruped Locomotion with Unsupervised Reinforcement Learning View paper
- [54] From Parameters to Behavior: Unsupervised Compression of the Policy Space View paper
- [55] Data-driven latent space representation for robust bipedal locomotion learning View paper
- [56] ZSL-RPPO: Zero-Shot Learning for Quadrupedal Locomotion in Challenging Terrains using Recurrent Proximal Policy Optimization View paper
- [57] Distributional Successor Features Enable Zero-Shot Policy Optimization View paper
- [58] Learning structures: predictive representations, replay, and generalization View paper
- [59] A neurally plausible model learns successor representations in partially observable environments View paper
- [60] A New Representation of Universal Successor Features for Enhancing the Generalization of Target-Driven Visual Navigation View paper
- [61] Learning Successor Feature Representations to Train Robust Policies for Multi-task Learning View paper
- [62] Accounting for sensitivity of latent learning to behavioral statistics with successor representations View paper
- [63] Policy-Oriented Cognitive Risk Map Modeling for Lane Change via Deep Successor Representation View paper
- [64] Eigenoption discovery through the deep successor representation View paper
- [65] Combining Behaviors with the Successor Features Keyboard View paper
- [66] Accelerating learning in constructive predictive frameworks with the successor representation View paper
- [67] Multi-agent LLMs with Offline Reinforcement Learning for Hierarchical Multi-turn Decision-making View paper
- [68] Uncertainty-driven exploration in sparse model-based reinforcement learning View paper
- [69] TOPS: Transition-Based Volatility-Reduced Policy Search View paper