Novelty Assessment Report

Paper: Regularized Latent Dynamics Prediction is a Strong Baseline For Behavioral Foundation Models

PDF URL: https://openreview.net/pdf?id=jdL6WB5jHZ

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-01

Abstract

Behavioral Foundation Models (BFMs) have been recently successful in producing agents with the capabilities to adapt to any unknown reward or task. In reality, these methods are only able to produce near-optimal policies for the reward functions that are in the span of some pre-existing state features. Naturally, their efficiency relies heavily on the choice of state features that they use. As a result, these BFMs have used a wide variety of complex objectives, often sensitive to environment coverage, to train task spanning features with different inductive properties. With this work, our aim is to examine the question: are these complex representation learning objectives necessary for zero-shot RL? Specifically, we revisit the objective of self-supervised next-state prediction in latent space for state feature learning, but observe that such an objective alone is prone to increasing state-feature similarity, and subsequently reducing span of reward functions that we can represent optimal policies for. We propose an approach, RLDP, that adds a simple regularization to maintain feature diversity and can match or surpass state-of-the-art complex representation learning methods for zero-shot RL. Furthermore, we demonstrate the prior approaches diverge in low-coverage scenarios where RLDP still succeeds.

Core Task Landscape

This paper addresses: **Representation Learning for Zero-Shot Reinforcement Learning**

A total of **50 papers** were analyzed and organized into a taxonomy with **27 categories**.

Taxonomy Overview

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

- **Latent Dynamics and Predictive Representation Learning**
- **Contrastive and Self-Supervised Representation Learning**
- **Decoupled and Disentangled Representation Learning**
- **Invariant and Robust Representation Learning**
- **Cross-Modal and Multi-Modal Representation Learning**
- **Meta-Learning and Task Representation for Zero-Shot Transfer**
- **Cross-Trajectory and Cross-Domain Generalization**
- **Sim-to-Real Transfer and Domain Adaptation**
- **Language-Conditioned and Semantic Representation Learning**
- **Object-Centric and Structured Representation Learning**
- ... and 7 more categories

Complete Taxonomy Tree

- Representation Learning for Zero-Shot Reinforcement Learning Survey Taxonomy
- Latent Dynamics and Predictive Representation Learning
  - Temporal Difference and Forward-Backward Representations (4 papers)
  - [6] Tackling the Zero-Shot Reinforcement Learning Loss Directly (Ollivier, 2025) View paper
  - [10] TD-JEPA: Latent-predictive Representations for Zero-Shot Reinforcement Learning (Bagatella, 2025) View paper
  - [21] Unsupervised Zero-Shot Reinforcement Learning via Dual-Value Forward-Backward Representation (J Sun, 2025) View paper
  - [45] Does zero-shot reinforcement learning exist? (Touati, 2022) View paper
  - Next-State Prediction and Latent Dynamics Modeling ★ (2 papers)
  - [0] Regularized Latent Dynamics Prediction is a Strong Baseline For Behavioral Foundation Models (Anon et al., 2026) View paper
  - [12] Towards Robust Zero-Shot Reinforcement Learning (Zheng KeXin, 2025) View paper
- Contrastive and Self-Supervised Representation Learning
  - Contrastive Predictive and Temporal Contrast Methods (3 papers)
  - [4] Decoupling Representation Learning from Reinforcement Learning (Stooke, 2020) View paper
  - [11] Representation Learning with Contrastive Predictive Coding (Oord, 2018) View paper
  - [46] On the Importance of Feature Decorrelation for Unsupervised Representation Learning in Reinforcement Learning (Lee Hojoon, 2023) View paper
  - General Self-Supervised and Unsupervised Learning (3 papers)
  - [34] Unsupervised state representation learning in atari (Kendrick Tan, 2019) View paper
  - [36] Building more efficient AI models through unsupervised representation learning (S Mishra, 2024) View paper
  - [39] PointUR-RL: Unified Self-Supervised Learning Method Based on Variable Masked Autoencoder for Point Cloud Reconstruction and Representation Learning (Kang Li, 2024) View paper
- Decoupled and Disentangled Representation Learning
  - Decoupling Policy from Representation Learning (1 papers)
  - [18] Disentangling policy from offline task representation learning via adversarial data augmentation (Jia, 2024) View paper

- Disentangled Task and Attribute Representations (2 papers)
  - [8] DARLA: Improving Zero-Shot Transfer in Reinforcement Learning (Higgins, 2017) View paper
  - [26] Zero-Shot Policy Transfer with Disentangled Task Representation of Meta-Reinforcement Learning (Zheng, 2022) View paper
- Invariant and Robust Representation Learning
  - Behavioral Similarity and Bisimulation Metrics (1 papers)
  - [3] Learning invariant representations for reinforcement learning without reconstruction (Zhang, 2020) View paper
  - Visual and Domain Invariance Methods (3 papers)
  - [23] The role of pretrained representations for the OOD generalization of RL agents (TrÃ¤uble, 2022) View paper
  - [30] SECANT: Self-Expert Cloning for Zero-Shot Generalization of Visual Policies (Fan, 2021) View paper
  - [31] Prompt-based Visual Alignment for Zero-shot Policy Transfer (Zhang Rui, 2024) View paper
- Cross-Modal and Multi-Modal Representation Learning (3 papers)
  - [1] Cross-modal representation learning for zero-shot action recognition (Lin, 2022) View paper
  - [19] Zero-shot cross-modal transfer of Reinforcement Learning policies through a Global Workspace (MaytiÃ©, 2024) View paper
  - [35] M2CURL: Sample-Efficient Multimodal Reinforcement Learning via Self-Supervised Representation Learning for Robotic Manipulation (Lygerakis, 2024) View paper
- Meta-Learning and Task Representation for Zero-Shot Transfer
  - Offline Meta-RL with Task Representations (1 papers)
  - [7] Generalizable task representation learning for offline meta-reinforcement learning with data limitations (Renzhe Zhou, 2024) View paper
  - Online and Compositional Meta-RL (2 papers)
  - [13] MR-Selection: A Meta-Reinforcement Learning Approach for Zero-Shot Hyperspectral Band Selection (Jie Feng, 2023) View paper
  - [33] Using task features for zero-shot knowledge transfer in lifelong learning. (David Isele, 2016) View paper
- Cross-Trajectory and Cross-Domain Generalization (2 papers)
  - [9] Cross-Trajectory Representation Learning for Zero-Shot Generalization in RL (Mazoure, 2021) View paper
  - [50] How the level sampling process impacts zero-shot generalisation in deep reinforcement learning (Garcin, 2023) View paper
- Sim-to-Real Transfer and Domain Adaptation (2 papers)
  - [2] Zero-shot sim-to-real transfer using Siamese-Q-Based reinforcement learning (Zhenyu Zhang, 2025) View paper
  - [14] Comparing Quadrotor Control Policies for Zero-Shot Reinforcement Learning under Uncertainty and Partial Observability (Sven Gronauer, 2023) View paper
- Language-Conditioned and Semantic Representation Learning (3 papers)
  - [16] LLM-Empowered State Representation for Reinforcement Learning (Wang BoYuan, 2024) View paper
  - [28] Semantic-guided reinforced region embedding for generalized zero-shot learning (Jiannan Ge, 2021) View paper
  - [32] Learning Invariable Semantical Representation from Language for Extensible Policy Generalization (Li Yihan, 2022) View paper
- Object-Centric and Structured Representation Learning (1 papers)
  - [20] Zero-shot object-centric representation learning (Didolkar, 2024) View paper
- Hierarchical and Subtask-Based Representation Learning (1 papers)
  - [25] Hierarchical reinforcement learning for zero-shot generalization with subtask dependencies (Sohn, 2018) View paper
- Unified Frameworks and Theoretical Foundations (1 papers)
  - [5] A Unified Framework for Zero-Shot Reinforcement Learning (Jacopo Di Ventura, 2025) View paper
- Behavioral Foundation Models and Unsupervised Pre-Training (1 papers)
  - [29] Zero-shot whole-body humanoid control via behavioral foundation models (Tirinzoni, 2025) View paper
- Imitation and Behavior Cloning for Representation Learning (4 papers)
  - [15] Learning Task-Agnostic Skill Bases to Uncover Motor Primitives in Animal Behaviors (J Wang, 2025) View paper
  - [42] Learning compound tasks without task-specific knowledge via imitation and self-supervised learning (Sang Hyun Lee, 2020) View paper
  - [43] Guided reinforcement learning with learned skills (K Pertsch, 2021) View paper
  - [44] Behavior Prior Representation learning for Offline Reinforcement Learning (Zang Hong-yu, 2022) View paper
- Representation Learning for Specific RL Domains
  - Robotic Manipulation and Dexterous Control (4 papers)
  - [27] RL-100: Performant Robotic Manipulation with Real-World Reinforcement Learning (Lei Kun, 2025) View paper
  - [37] Reset-free reinforcement learning via multi-task learning: Learning dexterous manipulation behaviors without human intervention (Gupta, 2021) View paper
  - [40] Learning Generalizable Pivoting Skills (Zhang Xiang, 2023) View paper
  - [49] Visual-policy learning through multi-camera view to single-camera view knowledge distillation for robot manipulation tasks (Cihan Acar, 2023) View paper
  - Scheduling and Combinatorial Optimization (1 papers)
  - [17] Learning to Dispatch for Job Shop Scheduling via Deep Reinforcement Learning (Zhang Cong, 2020) View paper
  - Plasma and Physical System Control (1 papers)
  - [24] Plasma Shape Control via Zero-shot Generative Reinforcement Learning (Li Rongpeng, 2025) View paper
- Auxiliary Methods and Supporting Techniques
  - Safety and Task-Agnostic Constraints (1 papers)
  - [47] Task-Agnostic Safety for Reinforcement Learning (Md Asifur Rahman, 2023) View paper
  - Continual and Lifelong Learning (1 papers)
  - [22] Continual deep reinforcement learning with task-agnostic policy distillation (Hafez, 2024) View paper
  - Multi-Agent Communication and Coordination (1 papers)
  - [38] TACTIC: Task-Agnostic Contrastive pre-Training for Inter-Agent Communication (Peihong Yu, 2025) View paper
  - Latent Space Mapping and Semantic Alignment (1 papers)
  - [41] Mapping representations in Reinforcement Learning via Semantic Alignment for Zero-Shot Stitching (Maiorca, 2025) View paper

- Block-Structured and Model-Free Representation Learning (1 papers)
  - [48] Efficient Reinforcement Learning in Block MDPs: A Model-free Representation Learning Approach (Zhang Xue-zhou, 2022) View paper

Narrative

Core task: representation learning for zero-shot reinforcement learning. The field is organized around diverse strategies for building representations that enable agents to generalize to unseen tasks or environments without additional training. At the highest level, the taxonomy distinguishes between approaches that emphasize predictive modeling of environment dynamics (Latent Dynamics and Predictive Representation Learning), methods that leverage contrastive or self-supervised signals (Contrastive and Self-Supervised Representation Learning[11]), and techniques that explicitly decouple or disentangle task-relevant factors (Decoupled and Disentangled Representation Learning[4,18]). Other major branches focus on invariance and robustness to distribution shifts (Invariant and Robust Representation Learning[3]), cross-modal integration (Cross-Modal and Multi-Modal Representation Learning[1]), meta-learning for task adaptation (Meta-Learning and Task Representation for Zero-Shot Transfer[7,13]), and domain transfer challenges such as sim-to-real (Sim-to-Real Transfer and Domain Adaptation[2,8]). Additional directions include language-conditioned policies (Language-Conditioned and Semantic Representation Learning[16,28]), object-centric factorizations (Object-Centric and Structured Representation Learning[20]), hierarchical decompositions (Hierarchical and Subtask-Based Representation Learning[25]), and large-scale pre-training (Behavioral Foundation Models and Unsupervised Pre-Training[29]).

A central tension across these branches concerns the trade-off between model-based predictive accuracy and task-agnostic feature learning: some works prioritize forward dynamics or next-state prediction to capture environment structure, while others argue that contrastive or invariance-based objectives yield more transferable representations. The original paper, Latent Dynamics Baseline[0], sits squarely within the predictive modeling branch, emphasizing next-state prediction and latent dynamics as a foundation for zero-shot transfer. This places it in close proximity to methods like TD-JEPA[10] and Robust Zero-shot[12], which similarly exploit temporal structure, yet contrasts with approaches such as Invariant Representations[3] or Unified Zero-shot Framework[5] that prioritize invariance or task-agnostic distillation over explicit dynamics modeling. The landscape reveals ongoing debate about whether predictive world models or task-invariant embeddings offer a more robust path to generalization, with Latent Dynamics Baseline[0] contributing a baseline perspective on the former.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Towards Robust Zero-Shot Reinforcement Learning

Authors: Zheng KeXin, Kexin Zheng, Zheng Yi-nan, Lauriane Teyssier, Luo Yu, et al. (10 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

The recent development of zero-shot reinforcement learning (RL) has opened a new avenue for learning pre-trained generalist policies that can adapt to arbitrary new tasks in a zero-shot manner. While the popular Forward-Backward representations (FB) and related methods have shown promise in zero-shot RL, we empirically found that their modeling lacks expressivity and that extrapolation errors caused by out-of-distribution (OOD) actions during offline learning sometimes lead to biased representat...

Relationship Analysis

Both papers belong to the Next-State Prediction and Latent Dynamics Modeling category, focusing on learning state representations through predictive objectives in latent space for zero-shot RL. They overlap in addressing representation learning for behavioral foundation models using successor features, with both identifying issues in existing methods like Forward-Backward representations. The key difference is that the original paper (RLDP) proposes regularized latent dynamics prediction with orthogonal regularization to prevent feature collapse, while the candidate paper (BREEZE) focuses on behavioral regularization during policy learning, diffusion-based policy extraction, and attention-based architectures to address expressivity and OOD action issues in FB-based methods.

Contributions Analysis

Overall novelty summary. The paper proposes RLDP, a method that revisits next-state prediction in latent space for learning state features in zero-shot reinforcement learning, augmented with a regularization to prevent feature collapse. It resides in the 'Next-State Prediction and Latent Dynamics Modeling' leaf, which contains only two papers (including this one), indicating a relatively sparse research direction within the broader 'Latent Dynamics and Predictive Representation Learning' branch. This positioning suggests the paper addresses a focused niche: simple predictive objectives for zero-shot RL, rather than the more crowded contrastive or meta-learning approaches elsewhere in the taxonomy.

The taxonomy reveals that neighboring branches emphasize alternative strategies: 'Temporal Difference and Forward-Backward Representations' (four papers) explores TD-based or bidirectional dynamics, while 'Contrastive and Self-Supervised Representation Learning' (six papers across two leaves) prioritizes contrastive losses over temporal prediction. The paper's focus on next-state prediction with regularization diverges from these directions by questioning whether complex objectives are necessary, positioning it as a simplification or baseline challenge to methods in adjacent branches that employ more elaborate contrastive or invariance-based objectives.

Among the three contributions analyzed, the core RLDP method and its robustness in low-coverage settings each examined ten candidates with no clear refutations, suggesting these aspects may be relatively novel within the limited search scope of thirty papers. However, the identification and mitigation of feature collapse examined ten candidates and found three that appear to provide overlapping prior work, indicating this specific problem and its solution have been addressed to some extent in the examined literature. The analysis does not claim exhaustive coverage, so additional relevant work may exist beyond the top-30 semantic matches.

Given the limited search scope and the paper's position in a sparse taxonomy leaf, the work appears to offer a focused contribution by revisiting a simple objective with a targeted fix for feature collapse. The analysis suggests moderate novelty for the method and robustness claims, while the feature collapse insight has more substantial prior overlap among the candidates examined. A broader literature search would be needed to assess whether the simplicity argument and regularization approach represent a significant departure from state-of-the-art complex methods.

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

Contribution 1: Regularized Latent Dynamics Prediction (RLDP) method

Description: The authors introduce RLDP, a representation learning method that combines latent next-state prediction with orthogonal regularization to prevent feature collapse. This approach provides a simpler alternative to complex successor measure estimation methods while maintaining competitive performance for zero-shot reinforcement learning.

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. RARA: Zero-shot Sim2Real Visual Navigation with Following Foreground Cues

URL: View paper

Brief Assessment

RARA[57] focuses on sim-to-real visual navigation using foreground cues and background randomization, not on latent dynamics prediction or zero-shot reinforcement learning frameworks.

---

2. Zero-Shot Self-Supervised Joint Temporal Image and Sensitivity Map Reconstruction via Linear Latent Space

URL: View paper

Brief Assessment

Linear Latent Space[54] addresses temporal MRI reconstruction using self-supervised learning in a linear latent space for medical imaging, not reinforcement learning or behavioral foundation models. The domains and technical objectives are fundamentally different.

---

3. Empowering Aerial Maneuver Games Through Model-Based Constrained Reinforcement Learning

URL: View paper

Brief Assessment

Aerial Maneuver Games[53] focuses on model-based RL for air combat with safety constraints and self-play, not on representation learning for zero-shot RL across diverse reward functions. The domains and objectives are fundamentally different.

---

4. Sim-to-real via latent prediction: Transferring visual non-prehensile manipulation policies

URL: View paper

Brief Assessment

Latent Prediction Transfer[56] focuses on sim-to-real transfer for robotic manipulation using a VAE-based latent dynamics model with a frozen dynamics predictor during transfer. RLDP addresses zero-shot RL with orthogonal regularization to prevent feature collapse across diverse reward functions. The technical objectives and problem settings differ fundamentally.

---

5. Prototypical context-aware dynamics for generalization in visual control with model-based reinforcement learning

URL: View paper

Brief Assessment

Prototypical Context-aware Dynamics[51] focuses on context-aware latent dynamics for visual control generalization across different environmental contexts, not on zero-shot RL with orthogonal regularization for preventing feature collapse as in RLDP.

---

6. Cross-Trajectory Representation Learning for Zero-Shot Generalization in RL

URL: View paper

Brief Assessment

Cross-Trajectory Learning[9] focuses on SSL objectives for behavioral similarity across trajectories in generalization tasks, not on latent dynamics prediction with orthogonal regularization for zero-shot RL as in RLDP.

---

7. Zero-shot whole-body humanoid control via behavioral foundation models

URL: View paper

Brief Assessment

Behavioral Foundation Models[29] focuses on forward-backward representations with conditional-policy regularization for humanoid control using unlabeled behavior datasets, not on latent dynamics prediction with orthogonal regularization for preventing feature collapse in zero-shot RL.

---

8. DRED: Zero-shot transfer in reinforcement learning via data-regularised environment design

URL: View paper

Brief Assessment

DRED[52] focuses on environment design and level sampling for zero-shot generalization across different environment instances, not on representation learning through latent dynamics prediction with regularization for zero-shot RL as in the original paper.

---

9. Kitchenshift: Evaluating zero-shot generalization of imitation-based policy learning under domain shifts

URL: View paper

Brief Assessment

KitchenShift[55] focuses on evaluating imitation-based policy learning under domain shifts in a kitchen environment, not on latent dynamics prediction methods for zero-shot RL. The candidate paper does not address representation learning through latent next-state prediction with regularization.

---

10. Transfer RL across observation feature spaces via model-based regularization

URL: View paper

Brief Assessment

Model-based Regularization Transfer[58] focuses on transfer learning across different observation spaces (e.g., vector to image), while RLDP addresses zero-shot RL with orthogonal regularization to prevent feature collapse within a single observation space for behavioral foundation models.

Contribution 2: Identification and mitigation of feature collapse in latent dynamics prediction

Description: The authors identify that naive latent dynamics prediction leads to increasing state-feature similarity (a mild form of feature collapse) that reduces the span of representable reward functions. They propose orthogonal regularization as a solution to maintain feature diversity during representation learning.

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. Quantized Representations Prevent Dimensional Collapse in Self-predictive RL

URL: View paper

Prior Art Analysis

Quantized Representations[75] demonstrates that dimensional collapse in self-predictive RL was identified and addressed prior to the original paper's submission. The candidate explicitly defines dimensional collapse (Definition 1.2), shows empirically that it occurs during latent dynamics learning without mitigation, and proposes quantization as a solution to preserve representation rank. This directly refutes the novelty claim that the original authors were first to identify increasing state-feature similarity (mild feature collapse) in latent dynamics prediction and propose orthogonal regularization as the solution.

Evidence

Evidence 1 - **Rationale**: Quantized Representations[75] formally defines dimensional collapse in self-predictive RL, demonstrating prior identification of the feature collapse phenomenon in latent dynamics prediction. - **Original**: we identify that naive latent dynamics prediction leads to increasing state-feature similarity, and subsequently reducing span. we propose an approach, rldp, that adds a simple regularization to maintain feature diversity - **Candidate**: definition 1.2 (dimensional collapse) . let eθ : o → zbe an encoder which maps observations o ∈o to latent states zt ∈z = rd, of dimension d. given a data set of n latent states dz = {z1,..., zn}, the representation is said to be dimensionally collapsed when the latent states span a lower dimensional...

Evidence 2 - **Rationale**: Both papers identify that feature collapse in latent dynamics prediction leads to poor performance, with Quantized Representations[75] providing empirical evidence that this was a known issue requiring mitigation. - **Original**: we find that in its naive form, this objective leads to a mild form of feature collapse where the representation of different states increase in similarity over training. this collapse results in poor zero-shot rl performance - **Candidate**: our experiments show that whilst dimensional collapse is not always an issue, in some more complex environments, it can prevent agents from learning to solve a task (see fig. 3). it is worth noting that previous approaches utilize auxiliary loss terms to help prevent representation collapse

Evidence 3 - **Rationale**: Quantized Representations[75] proposes quantization as a solution to prevent dimensional collapse in latent dynamics prediction, demonstrating that mitigation strategies for this problem existed before the original paper's orthogonal regularization approach. - **Original**: we propose an approach, rldp, that adds a simple regularization to maintain feature diversity and can match or surpass state-of-the-art complex representation learning methods - **Candidate**: we accomplish this by quantizing our latent representation with finite scalar quantization [17], without using any reconstruction loss. as a result, our latent space is bounded and associated with an implicit codebook, whose size we can control

Evidence 4 - **Rationale**: Quantized Representations[75] provides empirical evidence of identifying and mitigating dimensional collapse in latent dynamics prediction through quantization, showing this was addressed prior to the original paper's work. - **Original**: we investigate latent-dynamics prediction as a simple alternative to learning state features for zero-shot rl. we identify as well as mitigate feature collapse with learned representation plaguing latent dynamics prediction - **Candidate**: fig. 3 shows the orthogonality-preserving effect of our quantization scheme as the matrix rank stays close to the maximum. without quantization, a dimensional collapse occurs, which can have significant harmful effects as the representational power of the latent state diminishes

---

2. Sim2real transfer for deep reinforcement learning with stochastic state transition delays

URL: View paper

Brief Assessment

Stochastic Transition Delays[74] focuses on handling variable timing delays in sim2real transfer for robotics, not on feature collapse in latent dynamics prediction for representation learning in zero-shot RL.

---

3. TD-JEPA: Latent-predictive Representations for Zero-Shot Reinforcement Learning

URL: View paper

Brief Assessment

TD-JEPA[10] focuses on TD-based latent prediction for multi-policy learning from offline data, not on identifying or mitigating feature collapse through orthogonal regularization as the original paper does.

---

4. A reliable representation with bidirectional transition model for visual reinforcement learning generalization

URL: View paper

Brief Assessment

Bidirectional Transition Model[71] focuses on visual reinforcement learning with bidirectional transition models for reliability. The candidate's context fragments do not provide sufficient detail about feature collapse mechanisms or orthogonal regularization approaches that would refute the original paper's novelty claim regarding identification and mitigation of feature collapse in latent dynamics prediction.

---

5. Ilpo-mp: Mode priors prevent mode collapse when imitating latent policies from observations

URL: View paper

Brief Assessment

ILPO-MP[73] addresses mode collapse in generative forward dynamics models for imitation learning from observations, not feature collapse in latent dynamics prediction for reinforcement learning. The contexts differ fundamentally: ILPO-MP deals with discrete latent action spaces and forward model collapse, while the original work focuses on continuous state-feature similarity in zero-shot RL.

---

6. Derl: Coupling decomposition in action space for reinforcement learning task

URL: View paper

Brief Assessment

Decomposition Action Space[72] focuses on decomposing action space for policy learning, not on feature collapse in latent dynamics prediction or representation learning for state features.

---

7. On the Importance of Feature Decorrelation for Unsupervised Representation Learning in Reinforcement Learning

URL: View paper

Prior Art Analysis

Feature Decorrelation[46] directly addresses the same problem of representational collapse in latent dynamics prediction for RL. Both papers identify that predicting future states in latent space leads to feature collapse where representations collapse into a low-dimensional manifold. Both propose regularization techniques to maintain feature diversity - the original paper uses orthogonal regularization while Feature Decorrelation[46] uses feature decorrelation. The candidate paper explicitly states this is 'an important challenge' and proposes a solution, demonstrating prior work exists on this exact problem.

Evidence

Evidence 1 - **Rationale**: Both papers identify the same core problem: latent dynamics prediction leads to feature collapse. The original describes it as 'increasing state-feature similarity' while the candidate describes it as 'representational collapse' into a 'low-dimensional manifold' - these are equivalent formulations of the same phenomenon. - **Original**: we observe that such an objective alone is prone to increasing state-feature similarity, and subsequently reducing span - **Candidate**: an important challenge of this approach is the representational collapse, where the subspace of the latent representations collapses into a low-dimensional manifold

Evidence 2 - **Rationale**: Both papers work on the same fundamental approach (predicting future states in latent space) and identify the same fundamental problem (collapse). This shows Feature Decorrelation[46] was already addressing this issue. - **Original**: we revisit the objective of self-supervised next-state prediction in latent space for state feature learning, but observe that such an objective alone is prone to increasing state-feature similarity - **Candidate**: the underlying principle of these methods is to learn temporally predictive representations by predicting future states in the latent space. however, an important challenge of this approach is the representational collapse

Evidence 3 - **Rationale**: Both papers propose regularization-based solutions to maintain feature diversity during latent dynamics prediction. The original uses 'regularization to maintain feature diversity' while the candidate uses 'decorrelating the features' - both are methods to prevent the same collapse problem, showing prior work on this mitigation strategy. - **Original**: we propose an approach, rldp, that adds a simple regularization to maintain feature diversity - **Candidate**: we propose a novel url framework that causally predicts future states while increasing the dimension of the latent manifold by decorrelating the features in the latent space

8. Constrained latent action policies for model-based offline reinforcement learning

URL: View paper

Brief Assessment

Constrained Latent Actions[69] focuses on learning joint distributions of observations and actions in offline RL with latent action spaces, not on feature collapse in latent dynamics prediction for representation learning in zero-shot RL.

9. Understanding self-predictive learning for reinforcement learning

URL: View paper

Prior Art Analysis

Self-predictive Learning[68] demonstrates that naive latent dynamics prediction leads to representation collapse and proposes semi-gradient updates and optimal predictors as solutions. The paper explicitly identifies that 'trivial representations (such as constants) minimize the prediction error' and shows that 'careful designs of the optimization dynamics are critical to learning meaningful representations.' The candidate provides theoretical analysis (Theorem 1) proving non-collapse properties and empirically validates that removing semi-gradient updates or optimal predictors causes collapse, directly addressing the same problem space as the original paper's contribution.

Evidence

Evidence 1 - **Rationale**: Both papers propose solutions to prevent collapse in latent dynamics prediction, though with different specific mechanisms (orthogonal regularization vs. semi-gradient updates). - **Original**: we identify that naive latent dynamics prediction leads to increasing state-feature similarity, and subsequently reducing span. we propose an approach, rldp, that adds a simple regularization to maintain feature diversity - **Candidate**: we identify that a faster paced optimization of the predictor and semi-gradient updates on the representation, are crucial to preventing the representation collapse.

Evidence 2 - **Rationale**: Both papers empirically demonstrate and measure feature collapse in latent dynamics prediction using cosine similarity metrics, showing increasing similarity over training iterations. - **Original**: figure 1 shows that while the solutions do not collapse, there is an increase in feature similarity over the course of learning, which we refer to as a mild form of collapse. - **Candidate**: figure 2 shows the absolute value of the inner product between the two (normalized) columns of φ, also known as the cosine similarity, versus the number of iterations... when either semi-gradient or optimal predictor are removed from the learning algorithm, the representation columns start to collap...

10. Analyzing and overcoming degradation in warm-start reinforcement learning

URL: View paper

Brief Assessment

Warm-start Degradation[70] focuses on value function extrapolation error and policy degradation in warm-start RL settings, not on feature collapse in latent dynamics prediction for representation learning. The technical contexts are fundamentally different.

Contribution 3: Demonstration of robustness in low-coverage settings

Description: The authors demonstrate that RLDP, being a policy-independent representation learning objective, succeeds in low-coverage scenarios where prior approaches that rely on explicit Bellman backups struggle due to out-of-distribution action selection issues.

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. Disentangling policy from offline task representation learning via adversarial data augmentation

URL: View paper

Brief Assessment

Disentangling Policy[18] addresses offline meta-RL with limited behavior policies for task identification, not general zero-shot RL with low-coverage datasets for successor feature learning.

2. The state of sparse training in deep reinforcement learning

URL: View paper

Brief Assessment

Sparse Training[59] focuses on sparse neural network architectures in deep RL across various environments (Atari, MuJoCo), not on representation learning methods for low-coverage settings or behavioral foundation models as in the original paper.

3. Locality Sensitive Sparse Encoding for Learning World Models Online

URL: View paper

Brief Assessment

Locality Sensitive Encoding[65] addresses online learning under data nonstationarity and covariate shift, not the low-coverage offline RL setting where RLDP demonstrates advantages over Bellman backup methods.

4. Chip Floorplanning Optimization Using Deep Reinforcement Learning

URL: View paper

Brief Assessment

Chip Floorplanning[62] focuses on chip design optimization using DRL and GNNs for spatial layout problems, not on representation learning for general RL in low-coverage settings or policy-independent objectives.

5. Learning future representation with synthetic observations for sample-efficient reinforcement learning

URL: View paper

Brief Assessment

Synthetic Observations[64] focuses on image-based RL with synthetic visual observations for representation learning, not on low-coverage dataset scenarios or policy-independent objectives that avoid Bellman backup issues.

6. Topological identification and interpretation for single-cell epigenetic regulation elucidation in multi-tasks using scAGDE

URL: View paper

Brief Assessment

scAGDE[66] addresses single-cell epigenetic data sparsity in chromatin accessibility analysis, not reinforcement learning policy optimization in low-coverage scenarios. The technical domains are entirely distinct.

7. Solving Offline Reinforcement Learning with Decision Tree Regression

URL: View paper

Brief Assessment

Decision Tree Regression[61] focuses on reformulating offline RL as regression tasks using decision trees, evaluated on standard benchmarks. It does not specifically address low-coverage scenarios or compare representation learning approaches in such settings.

8. Pretraining representations for data-efficient reinforcement learning

URL: View paper

Brief Assessment

Pretraining Representations[60] focuses on offline pretraining for data-efficient RL using self-supervised objectives, not specifically on low-coverage settings or policy-independent representation learning that avoids Bellman backup issues.

9. Offline multitask representation learning for reinforcement learning

URL: View paper

Brief Assessment

Offline Multitask Representation[63] focuses on multitask representation learning across multiple offline tasks, not on low-coverage scenarios within single-task settings. The candidate does not address policy-independent representation learning objectives or compare against methods that rely on explicit Bellman backups in low-coverage settings.

10. Representation learning for online and offline rl in low-rank mdps

URL: View paper

Brief Assessment

Low-rank MDPs[67] addresses offline RL under partial coverage conditions in low-rank MDPs, while the original paper demonstrates RLDP's robustness in low-coverage scenarios for behavioral foundation models with policy-independent representation learning. These are distinct technical approaches to different problem settings.

Appendix: Text Similarity Detection

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

References

- [0] Regularized Latent Dynamics Prediction is a Strong Baseline For Behavioral Foundation Models View paper
- [1] Cross-modal representation learning for zero-shot action recognition View paper
- [2] Zero-shot sim-to-real transfer using Siamese-Q-Based reinforcement learning View paper
- [3] Learning invariant representations for reinforcement learning without reconstruction View paper
- [4] Decoupling Representation Learning from Reinforcement Learning View paper
- [5] A Unified Framework for Zero-Shot Reinforcement Learning View paper
- [6] Tackling the Zero-Shot Reinforcement Learning Loss Directly View paper
- [7] Generalizable task representation learning for offline meta-reinforcement learning with data limitations View paper
- [8] DARLA: Improving Zero-Shot Transfer in Reinforcement Learning View paper
- [9] Cross-Trajectory Representation Learning for Zero-Shot Generalization in RL View paper
- [10] TD-JEPA: Latent-predictive Representations for Zero-Shot Reinforcement Learning View paper
- [11] Representation Learning with Contrastive Predictive Coding View paper
- [12] Towards Robust Zero-Shot Reinforcement Learning View paper
- [13] MR-Selection: A Meta-Reinforcement Learning Approach for Zero-Shot Hyperspectral Band Selection View paper
- [14] Comparing Quadrotor Control Policies for Zero-Shot Reinforcement Learning under Uncertainty and Partial Observability View paper
- [15] Learning Task-Agnostic Skill Bases to Uncover Motor Primitives in Animal Behaviors View paper
- [16] LLM-Empowered State Representation for Reinforcement Learning View paper
- [17] Learning to Dispatch for Job Shop Scheduling via Deep Reinforcement Learning View paper
- [18] Disentangling policy from offline task representation learning via adversarial data augmentation View paper
- [19] Zero-shot cross-modal transfer of Reinforcement Learning policies through a Global Workspace View paper
- [20] Zero-shot object-centric representation learning View paper
- [21] Unsupervised Zero-Shot Reinforcement Learning via Dual-Value Forward-Backward Representation View paper
- [22] Continual deep reinforcement learning with task-agnostic policy distillation View paper
- [23] The role of pretrained representations for the OOD generalization of RL agents View paper
- [24] Plasma Shape Control via Zero-shot Generative Reinforcement Learning View paper
- [25] Hierarchical reinforcement learning for zero-shot generalization with subtask dependencies View paper
- [26] Zero-Shot Policy Transfer with Disentangled Task Representation of Meta-Reinforcement Learning View paper
- [27] RL-100: Performant Robotic Manipulation with Real-World Reinforcement Learning View paper
- [28] Semantic-guided reinforced region embedding for generalized zero-shot learning View paper
- [29] Zero-shot whole-body humanoid control via behavioral foundation models View paper

- [30] SECANT: Self-Expert Cloning for Zero-Shot Generalization of Visual Policies View paper
- [31] Prompt-based Visual Alignment for Zero-shot Policy Transfer View paper
- [32] Learning Invariable Semantical Representation from Language for Extensible Policy Generalization View paper
- [33] Using task features for zero-shot knowledge transfer in lifelong learning. View paper
- [34] Unsupervised state representation learning in atari View paper
- [35] M2CURL: Sample-Efficient Multimodal Reinforcement Learning via Self-Supervised Representation Learning for Robotic Manipulation View paper
- [36] Building more efficient AI models through unsupervised representation learning View paper
- [37] Reset-free reinforcement learning via multi-task learning: Learning dexterous manipulation behaviors without human intervention View paper
- [38] TACTIC: Task-Agnostic Contrastive pre-Training for Inter-Agent Communication View paper
- [39] PointUR-RL: Unified Self-Supervised Learning Method Based on Variable Masked Autoencoder for Point Cloud Reconstruction and Representation Learning View paper
- [40] Learning Generalizable Pivoting Skills View paper
- [41] Mapping representations in Reinforcement Learning via Semantic Alignment for Zero-Shot Stitching View paper
- [42] Learning compound tasks without task-specific knowledge via imitation and self-supervised learning View paper
- [43] Guided reinforcement learning with learned skills View paper
- [44] Behavior Prior Representation learning for Offline Reinforcement Learning View paper
- [45] Does zero-shot reinforcement learning exist? View paper
- [46] On the Importance of Feature Decorrelation for Unsupervised Representation Learning in Reinforcement Learning View paper
- [47] Task-Agnostic Safety for Reinforcement Learning View paper
- [48] Efficient Reinforcement Learning in Block MDPs: A Model-free Representation Learning Approach View paper
- [49] Visual-policy learning through multi-camera view to single-camera view knowledge distillation for robot manipulation tasks View paper
- [50] How the level sampling process impacts zero-shot generalisation in deep reinforcement learning View paper
- [51] Prototypical context-aware dynamics for generalization in visual control with model-based reinforcement learning View paper
- [52] DRED: Zero-shot transfer in reinforcement learning via data-regularised environment design View paper
- [53] Empowering Aerial Maneuver Games Through Model-Based Constrained Reinforcement Learning View paper
- [54] Zero-Shot Self-Supervised Joint Temporal Image and Sensitivity Map Reconstruction via Linear Latent Space View paper
- [55] Kitchenshift: Evaluating zero-shot generalization of imitation-based policy learning under domain shifts View paper
- [56] Sim-to-real via latent prediction: Transferring visual non-prehensile manipulation policies View paper
- [57] RARA: Zero-shot Sim2Real Visual Navigation with Following Foreground Cues View paper
- [58] Transfer RL across observation feature spaces via model-based regularization View paper
- [59] The state of sparse training in deep reinforcement learning View paper
- [60] Pretraining representations for data-efficient reinforcement learning View paper
- [61] Solving Offline Reinforcement Learning with Decision Tree Regression View paper
- [62] Chip Floorplanning Optimization Using Deep Reinforcement Learning View paper
- [63] Offline multitask representation learning for reinforcement learning View paper
- [64] Learning future representation with synthetic observations for sample-efficient reinforcement learning View paper
- [65] Locality Sensitive Sparse Encoding for Learning World Models Online View paper
- [66] Topological identification and interpretation for single-cell epigenetic regulation elucidation in multi-tasks using scAGDE View paper
- [67] Representation learning for online and offline rl in low-rank mdps View paper
- [68] Understanding self-predictive learning for reinforcement learning View paper
- [69] Constrained latent action policies for model-based offline reinforcement learning View paper
- [70] Analyzing and overcoming degradation in warm-start reinforcement learning View paper
- [71] A reliable representation with bidirectional transition model for visual reinforcement learning generalization View paper
- [72] Derl: Coupling decomposition in action space for reinforcement learning task View paper
- [73] Ilpo-mp: Mode priors prevent mode collapse when imitating latent policies from observations View paper
- [74] Sim2real transfer for deep reinforcement learning with stochastic state transition delays View paper
- [75] Quantized Representations Prevent Dimensional Collapse in Self-predictive RL View paper