Modular World Models with Competitive Independent Mechanisms

Anonymous CVPR submission

Paper ID *****

Abstract

Taking inspiration from the Independent Causal Mechanisms principle in the context of world models, we present COmpetitive Mechanisms for Efficient Transfer (COMET), a modular world model which leverages reusable, independent mechanisms across different environments. COMET is trained on multiple environments with varying dynamics via a two-step process: competition and composition. This enables the model to recognise and learn transferable mechanisms. Specifically, in the competition phase, COMET is trained with a winner-takes-all gradient allocation, encouraging the emergence of independent mechanisms. These are then re-used in the composition phase, where COMET learns to re-compose learnt mechanisms in ways that capture the dynamics of new environments. In so doing, COMET explicitly reuses prior knowledge, enabling efficient and interpretable adaptation. In contrast to competitive baselines, we demonstrate that COMET captures recognisable mechanisms without supervision and is able to adapt to new environments.

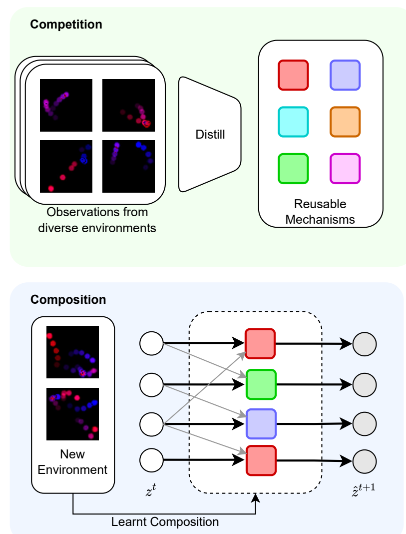

Figure 1. Illustration of training phases. In the first phase, the model learns a set of reusable mechanisms. In the second phase, the model applies these in a new environment.

1. Introduction

To reason about environments as rich and complex as our physical world requires the ability to learn efficiently and to flexibly adapt prior knowledge to unseen settings. This is of particular importance to embodied intelligence, as robots often need to adapt to changes in the environment. Whilst humans seem able to generalise knowledge across myriad tasks and situations effortlessly, building artificial agents that can do so with minimal training data remains a significant challenge. At the heart of this challenge are crucial distinctions between what and how humans and machines learn. It has been conjectured that humans represent knowledge internally in a structured and modular way, i.e., by distilling past experience into general principles about the world, which can be applied or selectively updated in novel settings [9, 23, 36, 39]. By contrast, current learning-based world models are mostly based on monolithic architectures, and the resulting entangled representations of the world limit the selective re-using of prior knowledge in new environments. Therefore, learning methods that afford modularity are key to world models that can adapt efficiently in diverse settings. In this paper, we address this challenge by developing a model capable of discovering a toolbox of generalisable concepts that can be reused across different contexts.

Recent works on object-centric world models [21, 25] have made progress towards a compositional understanding of the world. By decomposing an observed scene into discrete *object slots*, these methods model the interaction between entities in the scene. We argue that, just as the state representation of the scene can be factorised into object slots, the transition dynamics, too, can be factorised into discrete and independent *mechanisms*. Our work is motivated by the observation that, while different environments can exhibit different dynamics, we can often explain the behaviour of objects by a small set of interaction primitives, such as "A rests on top of B" and "C collides with D". The goal of this work is to acquire such a set of versatile mechanisms from observations without supervision.

Conceptually, our motivation is closely related to the *Independent Causal Mechanisms* principle (ICM) [36], which posits that the causal generative process of a system is composed of independent modules that do not inform each other. Taking inspiration from the ICM literature, we employ a competition of expert training scheme [33], where, during training, we keep a set of independently parameterised mechanisms and only update the module which best explains the observed interactions. This serves as a natural inductive bias which encourages modules to specialise in specific interaction primitives, and is consequently conducive to the emergence of reusable mechanisms. In the rest of this paper, we give details of our model and provide empirical results. We defer the discussion of related works to the Appendix.

2. COMET

In this section, we present COMET: COmpetitive Mechanisms for Efficient Transfer. The main idea of the method is to learn a set of generalisable and composable modules which encode the different modes of interaction between objects. Intuitively, while dynamics can vary across environments, the ways in which objects or entities interact with each other can be explained by a small number of independent rules. For example, in a traffic setting, whilst the behaviour of vehicles can differ across different locations, the act of stopping at a red light (the interaction between cars and traffic lights), can be used to explain behaviours in a wide range of environments. We argue that the ability to *selectively* update the modules during learning, i.e., to recognise parts of the model that are relevant to the data and perform modular updates, is instrumental to the emergence of such discrete independent mechanisms. To this end, we employ the competition of expert training procedure which only updates modules that best explains the interaction. In this scheme, modules that successfully captures reusable interactions automatically 'wins' more relevant training data and further specialises, which encourages a set of diverse modules that partitions the training data.

COMET learns from a dataset of observed sequences $\{\mathbf{x}^{1:T}\}_N$. Importantly, these sequences are sampled from environments with varying dynamics where objects can exhibit different behaviours. We assume that each observation, $\mathbf{x}^t$, can be factorised into latent object-slots, $\{\mathbf{z}_0^t, \mathbf{z}_1^t, ..., \mathbf{z}_K^t\}$, where the subscript denotes the object-id. These object representations can be based on ground-truth masks or obtained from object-centric encoders [5, 8, 29]. COMET consists of two components trained in two phases: the *mechanisms*, which models the interactions between objects and is trained in phase 1, and the *composition module*, which composes the mechanisms to explain a given environment and is trained is phase 2.

Mechanisms contains M independently parameterised feedforward networks, $f^m_{mech} : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$, with parameters θ_m, where d is the dimension of the object representations. Each mechanism predicts updates to all objects at every timestep, given the state of the object itself and another context object:

$$\Delta\mathbf{z}_i^t(m,j) = f^m_{mech}([\mathbf{z}_i^t \oplus \mathbf{z}_j^t]), \tag{1}$$

where $\oplus$ denotes concatenation, i is the index of the object to be predicted and j is the index of the *context* object, i.e. the object with which object i interacts. The mechanisms are trained during the *competition* phase where each mechanism learns to specialise to cover a particular mode of interaction between objects. Concretely, for each object, the model makes predictions using all possible mechanism-context pairs in parallel. Comparing the predictions, we update the mechanism-context pair with the most accurate prediction for a given object. Given a state transition pair, $(\mathbf{z}_{1:K}^t, \mathbf{z}_{1:K}^{t+1})$, the loss function can be written as:

$$\mathcal{L}(\theta_{1:M}) = \sum_{i=0}^{K} \min_{m,j} \left[d\left(\mathbf{z}_i^t + \Delta\mathbf{z}_i^t(m,j), \mathbf{z}_i^{t+1}\right)\right], \tag{2}$$

where d is a function that measures the prediction error. In this work, we use the Euclidean distance. Importantly, when performing back-propagation on this loss function, only the parameters of the competition winner are updated.

In order to predict transitions for objects using the trained mechanisms, the *composition module* picks the relevant context object and the active mechanism. The composition module acts as a classifier which picks the optimal mechanism-context pair for each object in the scene. Given the state of object i, for each mechanism-context pair, we compute the confidence score:

$$c_i^t(m,j) = f^m_{conf}([\mathbf{z}_i^t \oplus \mathbf{z}_j^t]), \tag{3}$$

where f^m_{conf} is an independently parameterised MLP for mechanism m, i.e. each mechanism has a corresponding f^m_{conf}. We predict a categorical distribution over all mechanism-context pairs by taking the softmax over the confidence scores for object i at time step t. Given a small number of observation sequences in a new environment, we obtain the best performing mechanism-context pair $(m^*, j^*)_i^t$ for each object at each time step by investigating which pair minimises the loss function of the competition scheme in Eq. 2. The composition module is then trained using the negative log-likelihood loss.

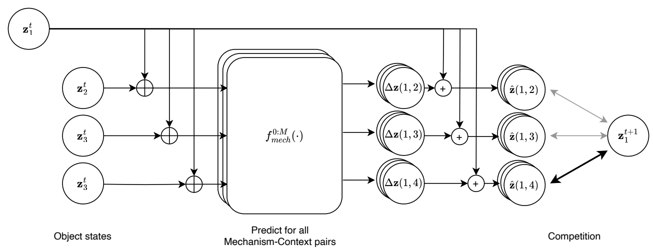

Figure 2. In the *competition* phase, predictions are made using all possible mechanism-context pairs for each object. Gradients are only allocated to the mechanism-context pair with the most accurate prediction. The figure describes the prediction step for a single object.

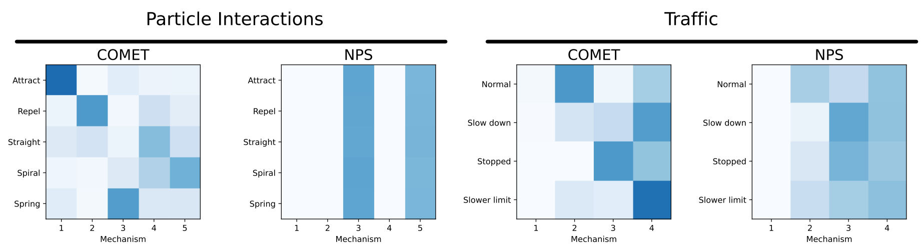

Figure 3. Disentanglement plots showing the correlation between mechanisms chosen by the models and ground-truth interaction modes. In the ideal case, the matrices should look like permutation matrices. Here, COMET is able to learn disentangled mechanisms that correspond to ground-truth behaviours in all three domains, as indicated by the fact that each interaction mode has one main corresponding learnt mechanism. In contrast, NPS does not exhibit the same structure.

3. Experiments

In this section, we demonstrate that COMET is able to disentangle different modes of interaction between objects and can efficiently reuse learnt mechanisms during adaptation.

Baselines. We evaluate COMET against two competitive baselines, C-SWM [21] and Neural Production Systems (NPS) [11]. C-SWM learns a world model from observation via contrastive learning with a GNN-based dynamics model. Similar to COMET, C-SWM operates on an object-factorised representation and achieves state-of-the-art prediction accuracy. COMET further disentangles the interactions between objects as independent mechanisms rather than learning a monolithic model that captures all interactions. We also compare against NPS, which learns to capture object interactions as independent mechanisms. Architecturally, NPS is similar to COMET, except that the mechanism-context pair is selected using dot-product attention [42] which is trained jointly with the mechanisms. In contrast, COMET deploys a competitive training scheme which allows the model to recognise shared mechanisms across environments. In our experiments, we show that competition serves as a strong inductive bias that enables the emergence of reusable mechanisms.

Datasets. We evaluate COMET on two problem domains: *Particle Interactions* and *Traffic*. For each of these domains, we define a set of environments where objects can exhibit different behaviours. These environments are designed to test whether COMET can extract meaningful mechanisms and adapt to unseen environments via composition. The *Particle Interactions* dataset consists of coloured particles that can interact with each other in different ways such as attraction and repulsion. Environments are defined by a combination of rules such as ”red particles repel each other”. The *Traffic* dataset contains observation sequences of traffic scenarios generated with the CARLA simulator [6]. Here, the environments are defined by traffic rules that apply to different vehicles such as ”blue cars do not need to stop at red lights”. In both cases, the observatiions are given as RGB images with ground-truth object masks such that each object can be encoded into an object-slot.

Disentanglement of Mechanisms. We investigate the emergence of recognisable mechanisms from competition. Here, both COMET and NPS are trained on a mixture of environments in each domain. We obtain the ground-truth labels for the object interactions and use these to directly investigate whether the learnt mechanisms correspond to actual interactions without any supervision. These labels are not accessible to the models during training. Fig. 3

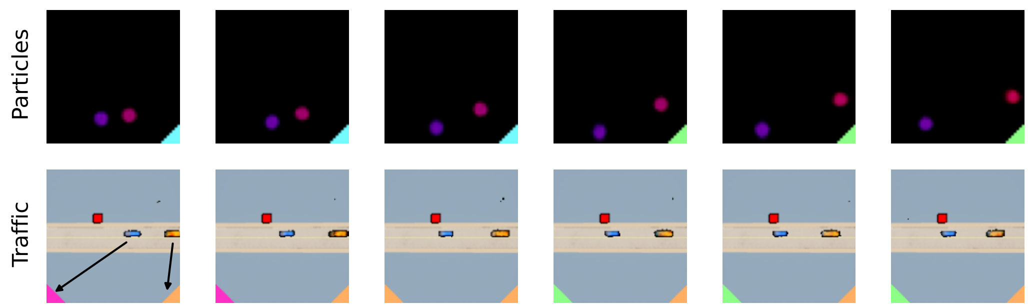

Figure 4. Qualitative rollouts. The colour of the tabs on the bottom of each frame indicates the 'winning' mechanism at each time step. Across all environments, the competition winner changes as the underlying interaction mode changes. **Top**: The particles repel each other when they are close (blue) and moves independently when they are apart (green). **Bottom**: In this traffic environment, the orange car obeys a slower speed limit and always pick the slow mechanism (orange). The blue car approaches the red light with normal driving (pink) → slow down (orange) → stop (green). Note that the orange mechanism is used as slow driving for both cars.

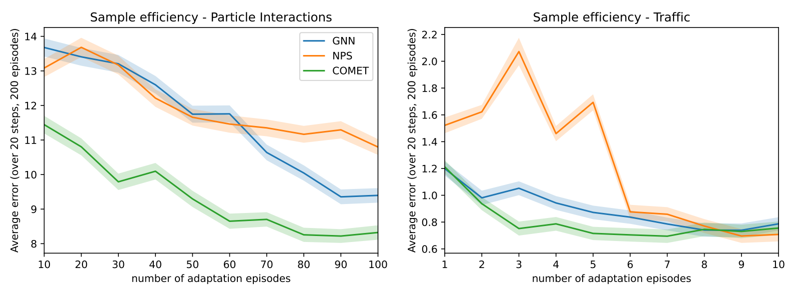

Figure 5. The average rollout error in an unseen environment with different amount of observed data in the new environment (lower is better). Shaded areas represent the standard errors of the mean. All models eventually converge to similar errors given enough data. However, COMET is able to achieve lower errors with few adaptation episodes. This means that COMET can learn to use the correct mechanisms with a small amount of data, thus corroborates our hypothesis that composing learnt mechanisms enables efficient transfer.

shows the correlation between the ground-truth interactions and the winning mechanisms in the competition process. COMET achieves successful disentanglement and learns mechanisms that corresponds to the ground-truth interactions, COMET recovering the ground-truth mode of interactions such as stopping before a red light. In contrast, the mechanisms learnt by NPS show no correspondence with the ground-truth interactions. This is likely because NPS cannot learn from a mixture of environments with varying dynamics as it employs a simple dot-product attention for picking mechanisms during training. To this end, COMET's ability to learn from diverse environments is uniquely afforded by the competition scheme which assigns relevant data to update each mechanism. Fig. 4 qualitatively illustrate that the 'winning' mechanisms switches as the underlying interaction type changes.

Adaptation Efficiency One of our main hypotheses is that learning to compose learnt mechanisms leads to data-efficient adaptation. For each domain, we train all of the models on a mixture of environments and adapt the models to unseen environments. COMET adapts via the composition module, whereas the baselines adapt by finetuning the entire model on new data. Fig. 5 shows the performance of the models when trained on different amounts of data in the new environment. COMET outperforms the baselines in the low-data regime, illustrating that reusing mechanisms improves sample efficiency compared to finetuning.

4. Conclusion

In this paper, we introduce COMET, a structured world model which encodes discrete abstract mechanisms explicitly from observations. Our model performs *selective* updates during the training phase, a central capability which facilitates the emergence of recognisable and reusable mechanisms. We show experimentally that the proposed method is indeed able to disentangle shared mechanisms across different environments from image observations, and thus enables sample-efficient and interpretable adaptation to novel situations. Looking forward, we believe that the method developed here opens up several promising avenues of research, such as designing agents that learn a growing repertoire of re-usable interaction behaviours and agents that explore the world through the lens of mechanism discovery.

References

[1] Rahaf Aljundi, Punarjay Chakravarty, and Tinne Tuytelaars. Expert gate: Lifelong learning with a network of experts. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3366–3375, 2017. 1

[2] Rim Assouel, Lluis Castrejon, Aaron Courville, Nicolas Ballas, and Yoshua Bengio. VIM: Variational independent modules for video prediction. In *Proceedings of the First Conference on Causal Learning and Reasoning*, pages 70–89. PMLR, 2022. 1

[3] David Baraff. Physically based modeling: Rigid body simulation. *SIGGRAPH Course Notes, ACM SIGGRAPH*, 2(1): 2–1, 2001. 1

[4] Philip Becker-Ehmck, Jan Peters, and Patrick Van Der Smagt. Switching linear dynamics for variational bayes filtering. In *International conference on machine learning*, pages 553–562. PMLR, 2019. 1

[5] Christopher P. Burgess, Loic Matthey, Nicholas Watters, Rishabh Kabra, Irina Higgins, Matt Botvinick, and Alexander Lerchner. MONet: Unsupervised Scene Decomposition and Representation. *arXiv preprint arXiv:1901.11390*, 2019. 2, 1

[6] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. CARLA: An open urban driving simulator. In *Proceedings of the 1st Annual Conference on Robot Learning*, pages 1–16, 2017. 3

[7] Patrick Emami, Pan He, Anand Rangarajan, and Sanjay Ranka. A Symmetric and Object-Centric World Model for Stochastic Environments. In *NeurIPS 2020 Workshop on Object Representations for Learning and Reasoning*, 2020. 1

[8] Martin Engelcke, Oiwi Parker Jones, and Ingmar Posner. GENESIS-V2: Inferring Unordered Object Representations without Iterative Refinement. In *Advances in Neural Information Processing Systems*, pages 8085–8094. Curran Associates, Inc., 2021. 2, 1

[9] Anirudh Goyal and Yoshua Bengio. Inductive biases for deep learning of higher-level cognition. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 478(2266):20210068, 2022. 1

[10] Anirudh Goyal, Alex Lamb, Phanideep Gampa, Philippe Beaudoin, Sergey Levine, Charles Blundell, Yoshua Bengio, and Michael Mozer. Object Files and Schemata: Factorizing Declarative and Procedural Knowledge in Dynamical Systems. *arXiv preprint arXiv:2006.16225*, 2020. 1

[11] Anirudh Goyal, Aniket Rajiv Didolkar, Nan Rosemary Ke, Charles Blundell, Philippe Beaudoin, Nicolas Heess, Michael Curtis Mozer, and Yoshua Bengio. Neural production systems. In *Advances in Neural Information Processing Systems*, 2021. 3, 1

[12] Anirudh Goyal, Alex Lamb, Jordan Hoffmann, Shagun Sodhani, Sergey Levine, Yoshua Bengio, and Bernhard Schölkopf. Recurrent Independent Mechanisms. In *International Conference on Learning Representations*, 2021. 1

[13] Klaus Greff, Raphaël Lopez Kaufman, Rishabh Kabra, Nick Watters, Christopher Burgess, Daniel Zoran, Loic Matthey, Matthew Botvinick, and Alexander Lerchner. Multi-Object Representation Learning with Iterative Variational Inference. In *Proceedings of the 36th International Conference on Machine Learning*, pages 2424–2433. PMLR, 2019. 1

[14] David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. In *Advances in Neural Information Processing Systems 31*, pages 2451–2463. Curran Associates, Inc., 2018. 1

[15] Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning Latent Dynamics for Planning from Pixels. In *Proceedings of the 36th International Conference on Machine Learning*, pages 2555–2565. PMLR, 2019. 1

[16] Patrick J Hayes. The naive physics manifesto. *Expert systems in the microelectronic age*, 1979. 1

[17] Biwei Huang, Fan Feng, Chaochao Lu, Sara Magliacane, and Kun Zhang. AdaRL: What, where, and how to adapt in transfer reinforcement learning. In *International Conference on Learning Representations*, 2022. 1

[18] Robert Jacobs and Michael Jordan. A competitive modular connectionist architecture. *Advances in neural information processing systems*, 3, 1990. 1

[19] Jindong Jiang, Sepehr Janghorbani, Gerard De Melo, and Sungjin Ahn. SCALOR: Generative World Models with Scalable Object Representations. In *International Conference on Learning Representations*, 2020. 1

[20] Michael I Jordan and Robert A Jacobs. Hierarchical mixtures of experts and the em algorithm. *Neural computation*, 6(2): 181–214, 1994. 1

[21] Thomas Kipf, Elise van der Pol, and Max Welling. Contrastive learning of structured world models. In *International Conference on Learning Representations*, 2020. 1, 3

[22] Adam Kosiorek, Hyunjik Kim, Yee Whye Teh, and Ingmar Posner. Sequential Attend, Infer, Repeat: Generative Modelling of Moving Objects. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2018. 1

[23] Brenden M. Lake, Tomer D. Ullman, Joshua B. Tenenbaum, and Samuel J. Gershman. Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40:e253, 2017. 1

[24] Anson Lei, Bernhard Schölkopf, and Ingmar Posner. Variational causal dynamics: Discovering modular world models from interventions. *CoRR*, abs/2206.11131, 2022. 1

[25] Zhixuan Lin, Yi Fu Wu, Skand Peri, Bofeng Fu, Jindong Jiang, and Sungjin Ahn. Improving generative imagination in object-centricworld models. In *37th International Conference on Machine Learning, ICML 2020*, pages 6096–6105, 2020. 1

[26] Zhixuan Lin, Yi-Fu Wu, Skand Vishwanath Peri, Weihao Sun, Gautam Singh, Fei Deng, Jindong Jiang, and Sungjin Ahn. SPACE: Unsupervised Object-Oriented Scene Representation via Spatial Attention and Decomposition. In *International Conference on Learning Representations*, 2020. 1

[27] Phillip Lippe, Sara Magliacane, Sindy Löwe, Yuki M Asano, Taco Cohen, and Stratis Gavves. Citris: Causal identifiability from temporal intervened sequences. In *International Conference on Machine Learning*, pages 13557–13603. PMLR, 2022. 1

[28] Francesco Locatello, Damien Vincent, Ilya Tolstikhin, Gunnar Rätsch, Sylvain Gelly, and Bernhard Schölkopf. Competitive training of mixtures of independent deep generative models. *arXiv preprint arXiv:1804.11130*, 2018. 1

[29] Francesco Locatello, Dirk Weissenborn, Thomas Unterthiner, Aravindh Mahendran, Georg Heigold, Jakob Uszkoreit, Alexey Dosovitskiy, and Thomas Kipf. Object-Centric Learning with Slot Attention. In *Advances in Neural Information Processing Systems*, pages 11525–11538. Curran Associates, Inc., 2020. 2, 1

[30] Cheol-Hui Min, Jinseok Bae, Junho Lee, and Young Min Kim. GATSBI: Generative Agent-Centric Spatio-Temporal Object Interaction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3074–3083, 2021. 1

[31] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018. 1

[32] Oleksiy Ostapenko, Pau Rodriguez, Massimo Caccia, and Laurent Charlin. Continual learning via local module composition. In *Advances in Neural Information Processing Systems*, 2021. 1

[33] Giambattista Parascandolo, Niki Kilbertus, Mateo Rojas-Carulla, and Bernhard Schölkopf. Learning independent causal mechanisms. In *International Conference on Machine Learning*, pages 4036–4044. PMLR, 2018. 2, 1

[34] Cansu Sancaktar, Sebastian Blaes, and Georg Martius. Curious Exploration via Structured World Models Yields Zero-Shot Object Manipulation. In *Advances in Neural Information Processing Systems*, 2022. 1

[35] Bernhard Schölkopf, Dominik Janzing, Jonas Peters, Eleni Sgouritsa, Kun Zhang, and Joris Mooij. On causal and anticausal learning. In *Proceedings of the 29th International Conference on Machine Learning*, pages 459–466, 2012. 1

[36] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Towards causal representation learning. *Proceedings of the IEEE*, 109(5):612–634, 2021. 1, 2

[37] Noam Shazeer, *Azalia Mirhoseini, *Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *International Conference on Learning Representations*, 2017. 1

[38] Gautam Singh, Skand Peri, Junghyun Kim, Hyunseok Kim, and Sungjin Ahn. Structured World Belief for Reinforcement Learning in POMDP. In *Proceedings of the 38th International Conference on Machine Learning*, pages 9744–9755. PMLR, 2021. 1

[39] Elizabeth S Spelke and Katherine D Kinzler. Core knowledge. *Developmental science*, 10(1):89–96, 2007. 1

[40] Andrea Tacchetti, H. Francis Song, Pedro A. M. Mediano, Vinicius Zambaldi, János Kramár, Neil C. Rabinowitz, Thore Graepel, Matthew Botvinick, and Peter W. Battaglia. Relational Forward Models for Multi-Agent Learning. In *International Conference on Learning Representations*, 2018. 1

[41] Sjoerd van Steenkiste, Michael Chang, Klaus Greff, and Jürgen Schmidhuber. Relational Neural Expectation Maximization: Unsupervised Discovery of Objects and their Interactions. In *International Conference on Learning Representations*, 2018. 1

[42] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2017. 3

[43] Rishi Veerapaneni, John D. Co-Reyes, Michael Chang, Michael Janner, Chelsea Finn, Jiajun Wu, Joshua Tenenbaum, and Sergey Levine. Entity Abstraction in Visual Model-Based Reinforcement Learning. In *Proceedings of the Conference on Robot Learning*, pages 1439–1456. PMLR, 2020. 1

[44] Ruben Villegas, Arkanath Pathak, Harini Kannan, Dumitru Erhan, Quoc V Le, and Honglak Lee. High fidelity video prediction with large stochastic recurrent neural networks. *Advances in Neural Information Processing Systems*, 32, 2019. 1

[45] Julius von Kügelgen, Ivan Ustyuzhaninov, Peter Gehler, Matthias Bethge, and Bernhard Schölkopf. Towards causal generative scene models via competition of experts. *arXiv preprint arXiv:2004.12906*, 2020. 1

[46] Nicholas Watters, Loic Matthey, Matko Bosnjak, Christopher P. Burgess, and Alexander Lerchner. COBRA: Data-Efficient Model-Based RL through Unsupervised Object Discovery and Curiosity-Driven Exploration. *arXiv preprint arXiv:1905.09275*, 2019. 1

[47] Yizhe Wu, Oiwi Parker Jones, Martin Engelcke, and Ingmar Posner. APEX: Unsupervised, Object-Centric Scene Segmentation and Tracking for Robot Manipulation. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3375–3382, 2021. 1

Modular World Models with Competitive Independent Mechanisms

Supplementary Material

5. Related Works

Learning internal models of the world enables decision-making agents to plan, predict and reason about the world [14, 15, 21]. As such, latent world models have attracted significant interest in recent years. These methods [e.g. 21, 31, 44] in general involve learning latent representations of the state and forward prediction models. Our work situates in this broad context of world model learning and we focus our contribution on learning dynamics models which are factorised into composable mechanisms. We take recent works form object-centric state representations as a starting point.

Object-Centric Representations There has been a growing interest in models that reflect the compositional nature of real-world scenarios and aim to use object-centric representations to leverage recurring features in scenes. Prior works have investigated unsupervised object-centric representation learning from static images [5, 8, 13, 26, 29]. Motivated by the assumption that dynamics tend to manifest themselves at the object-level [3, 16], subsequent works extend this capability to video data via factorised dynamics models which operate on object-centric latent spaces. While most of these object-centric world models (OCWMs) are geared towards using temporal inputs to generate future video rollouts [7, 19, 21, 22, 25, 30], some more explicitly consider their use in model-based reinforcement learning and planning [38, 43, 46, 47]. In particular, graph neural networks (GNNs) are often used as a natural way to predict future states of objects and enable the modelling of interactions between objects via message passing [21, 25, 34, 38, 40, 41, 43]. We build on these approaches by further factorising the dynamics into reusable interaction primitives.

Mechanism-based Models Our work is motivated by the conjecture that the organisation of knowledge into high-level abstract concepts is crucial to systematic generalisation [9]. This idea is similar in spirit to the *Independent Causal Mechanisms* principle [35] and the *Sparse Mechanism Shift* hypothesis [36] in the causality literature, which respectively posit that data-generating causal mechanisms operate independently from one another, and that changes in the environment can be attributed to sparse changes to such mechanisms. Several works [17, 24, 27] have leveraged causal discovery techniques, e.g., sparsity regularisation, to learn dynamics models that are factorised into structural causal models.

Similar to our approach is a class of models which represents the learned dynamics in OCWMs not as a monolithic module, but rather as a collection of independently acting mechanisms – each focusing on a different aspect of the environment's dynamics. Becker-Ehmck et al. [4] use a variational approach to learn to pick different transition models conditioned on the state, but is limited to linear transitions. RIMs [12] constitute an approach where parts of the state space are represented by independent and sparsely interacting recurrent units. Building on this, [10] use a GNN to model environment dynamics but reflect the concept of independent mechanisms by using different sets of GNN parameters depending on an object's current state. Another approach that follows this line of work is VIM [2] which considers the disentanglement of mechanisms and objects in the setting where object move independently to each other. Closer to our method are Neural Production Systems (NPS) [11], another descendant of RIMs, learning a set of independent mechanisms capturing the interaction between objects. Our method differs from NPS in the application of competition training which, as we demonstrate empirically in Sec. 3, is instrumental to the emergence of composable mechanisms. Furthermore, we propose a novel method for adapting to changes in the environment.

Competition of Experts The backbone of our learning algorithm draws from mixture of experts methods [18, 20, 37] and in particular from the algorithm of Parascandolo et al. [33]. In the context of learning independent causal mechanisms, Parascandolo et al. [33] demonstrate that the competition of experts algorithm induces the emergence of mechanisms that explain transformations in the data. The idea of utilising a competitive training scheme on modular model architectures has been applied on diverse settings such as lifelong learning [1, 32], generative models [28] and object-centric scene composition [45]. Taking inspiration from this line of work, COMET uses a similar competitive training scheme as an inductive bias for disentangling modes of interaction in the setting of world model learning.