DO WORLD MODELS LEARN GLOBAL UNDERSTANDING?

Anonymous authors
Paper under double-blind review

ABSTRACT

AI systems often feel frustratingly brittle and fragmented. A large language model (LLM) may correctly explain a concept but fail to apply it, or follow safety instructions in one context but not another. This behavior suggests a general failure to lift local information to a global understanding. To gain fundamental insight, we frame "understanding" as learning constraints and propagating their consequences. We construct learning tasks on monoid worlds, sets of states connected by action transitions $(\text{state}, \text{action}, \text{next state})$, where observed training transitions and an underlying unseen constraint jointly determine held-out transitions. Measuring generalization tests whether models can learn global constraints from local transitions and propagate their consequences. We consider inverse, commutativity, composition, and periodicity constraints relevant to spatial and semantic structure. Across attention, recurrent, and state-space architectures, next-state training on paths of observed transitions fits the data but fails to propagate non-trivial constraints. Compositional training, which uses identical paths but hides intermediate states from the input, achieves 96% accuracy on inverse, commutativity, and composition constraints across architectures, yields corresponding improvements in geometric generalization of world models trained on embodied environments ($+42\%$ on propagating commutative constraints) and relational generalization in Wikidata-finetuned LLMs ($+73\%$ on composition propagation). How far do models propagate constraints when inferring an unseen fact may depend on first inferring others? We define proof depth d of a held-out transition, measuring the minimum number of inference rounds to infer the transition, and find that model generalization decreases sharply with increasing proof depth. A compositionally trained transformer on length $T = 2$ paths generalizes at 95% for transitions derivable directly from observations ($d = 1$) but only at 49% for those requiring an additional inference round ($d = 2$). Increasing compositional path length T improves generalization, and $T = 4$ compositional models achieve 95% accuracy on $d = 2$ transitions. These results provide a formal way to investigate global understanding in language and world models and demonstrate that compositional training promotes information propagation and integration.

1 INTRODUCTION

What does it mean for a system to understand? Do language models understand semantics when training on "Tom's parent is Mary" fails to teach "Mary's child is Tom" (Berglund et al., 2024)? Do world models understand space if a left rotation followed by a right rotation produces a new scene (Kwon et al., 2026)? In both cases, a constraint (parent and child are converses; left and right rotations are inverses) lets a few seen facts determine many unseen ones. This suggests framing "understanding" as learning constraints and propagating their consequences globally.

To study constraint learning and propagation systematically, we construct monoid worlds where constraints are controllable and propagation consequences can be mathematically determined. A monoid world is a set of states with actions defining transitions between states. In the context of space, the set could be points on a line, and the actions, left/right, could correspond to opposite translations. These spatial actions satisfy an inverse constraint: moving left and then right returns you to the original state. Analogously, semantic relations obey constraints like composition: the parent of a parent is a grandparent. We capture these constraints as equalities over action sequences, enabling us to determine from data, manipulate, and directly specify a large class of constraints governing each world.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Since our world is precisely defined, we can directly evaluate whether models propagate constraints by constructing held-out sets of transitions which are uniquely determined (semantically entailed) by joint knowledge of the constraint and observed transitions. For example, observing $(1, \text{left}, 2)$ entails the unseen transition $(2, \text{right}, 1)$ when the constraint $\text{left} = \text{right}^{-1}$ is learned. Generalization determines whether models learn constraints from data and propagate their consequences.

Not all consequences of constraints are created equal, and some require multiple inference steps. Consider learning a new fact which sets in motion a chain of factual updates. We frame this formally by defining a miniature proof system and measure the inference complexity of an unseen transition by its proof depth d: the minimum number of rounds in our system to derive the transition. We construct worlds with unseen transitions with varied proof depths and explore how generalization and training dynamics change with the required inference depth.

Next-state training on paths of monoid transitions fails to propagate non-trivial constraints across transformer, recurrent, and state-space models. Compositional training, learning action compositions by hiding intermediate state inputs, improves generalization across worlds at an average accuracy of 96% on inverse, commutative, and composition constraints. We apply our monoid framing to study constraint propagation in learning visual environments with world models and relational structure with pretrained large language models (LLMs). We find traditional next-state-style objectives poorly propagate geometric and semantic constraints, while compositional training improves generalization by $+42\%$ on commutative constraints and $+73\%$ on composition constraints. Finally, we find that generalization decreases sharply with proof depth d, implying that models struggle to iteratively apply constraints. Increasing compositional path length T, however, improves deeper propagation.

2 A MONOID FRAMEWORK FOR GLOBAL CONSTRAINTS

We construct worlds governed by specified constraints and hold out transitions which are jointly implied by observed transitions and associated constraints. We train models on paths of observed transitions and evaluate them on held-out transitions, measuring a model's ability to learn global constraints from local data and propagate their consequences.

2.1 MONOID WORLD AND GLOBAL CONSTRAINTS

A monoid world is a finite set of states $\mathcal{S} := [N] := \{0, \ldots, N-1\}$ with actions $\mathcal{A} := \{a_0, \ldots, a_{n-1}\}$, each a function $a : \mathcal{S} \to \mathcal{S}$ on the state space. Consider the world shown in Figure 1a, which has a discrete $90°$ rotation interpretation with states $\mathcal{S} := [4]$ and actions $L : x \mapsto x - 1 \mod 4$ and $R : x \mapsto x + 1 \mod 4$. Actions can be composed to form action sequences, and with the identity action 1, form a monoid. Local observations are transitions $(s, a, a(s))$, which correspond to the change of state $s \in \mathcal{S}$ under the action $a \in \mathcal{A}$. In our rotation world, $(0, R, 1)$ is a transition.

The global constraints of our monoid world are equality relations between action sequences: two distinct sequences of actions induce the same map on the entire state space. In other words, at every initial state, the paths of transitions for the two action sequences converge to the same terminal state. The global nature comes from the requirement that the constraint on transitions applies uniformly over the entire state space. In our rotation world, we have two essential constraints: inverse $LR = RL = 1$ and periodicity $L^4 = 1$, from which all other constraints are derived.

When building a monoid world with a specified constraint, we choose a state-space size N and randomly sample action transitions $a \in \mathcal{A}$ such that the constraint is satisfied. For example, when constructing inverse worlds $LR = RL = 1$ as in Figure 1a, we choose $L : \mathcal{S} \to \mathcal{S}$ as a random bijection and derive R as its inverse. Monoid theory warns us that finite worlds may inadvertently introduce unwanted constraints, so we audit all our worlds by enumerating $u = v$ global constraints for $|u| + |v| \leq 8$ and reject impure worlds. We explore inverse, commutativity, composition, periodicity, and copy constraints shown as a presentation, $\langle \text{Actions} \mid \text{Constraints} \rangle$, in Figure 2.

2.2 EVALUATION OF CONSTRAINT PROPAGATION BY ENTAILED HOLD-OUTS

We hold out transitions which are jointly inferable from observed training transitions and the world's global constraint. Formally, inferability is captured by semantic entailment. A transition $t = (s, a, s')$ is semantically entailed by constraint $\mathcal{C}$ and seen transitions $\mathcal{T}$ if and only if all possible assignments

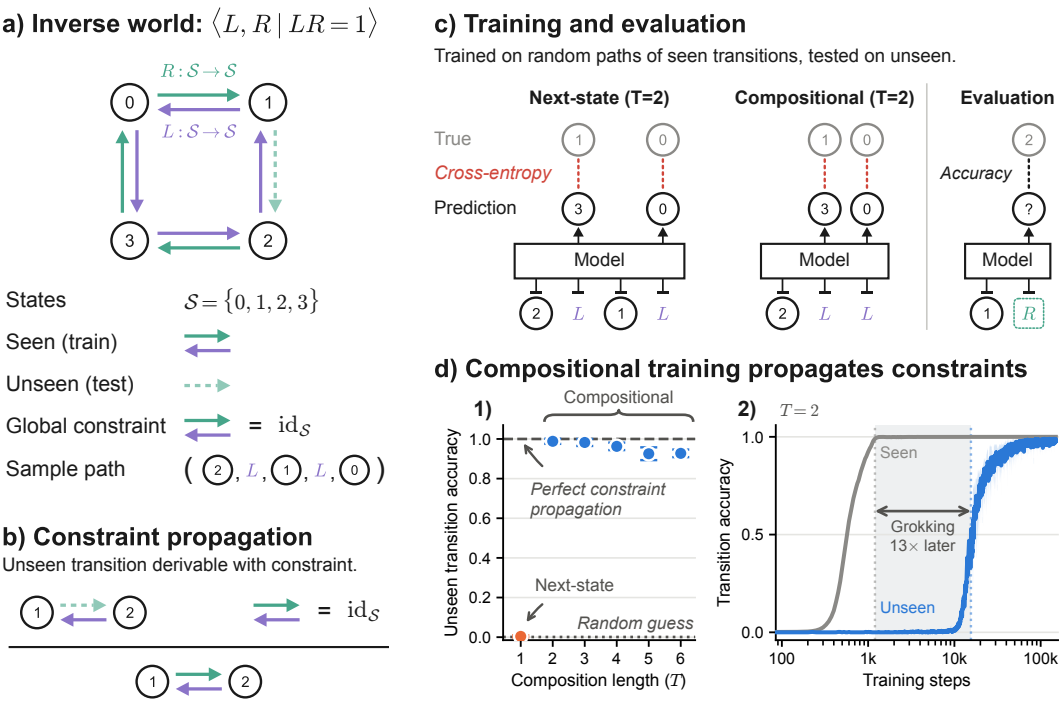


Figure 1: **Toy instantiation of an inverse monoid world with training and evaluation setup. a)** A four-state world with actions L and R satisfying inverse constraint $LR = RL = 1$, seen and unseen transitions, and sample training path. **b)** Constraint propagation uses seen transitions and the constraint to infer unseen transitions. **c)** Next-state training shows intermediate states, while compositional training does not. Evaluation is generalization over unseen transitions. **d)** 1) Next-state training fails to propagate inverse constraint, while compositional training succeeds ($\geq 93\%$). 2) Training curves for $T = 2$ show sharp grokking-like generalization. Plots averaged over 5 seeds.

of transitions that satisfy $\mathcal{C}$ and agree on $\mathcal{T}$ also agree on t. Consider the cyclic inverse world in Figure 1a. The unseen transition $t = (1, R, 2)$ is semantically entailed since any world that contains the seen transition $(2, L, 1) \in \mathcal{T}$ and the inverse constraint must agree on t. In practice, we perform greedy sampling with semantic entailment checks until we reach a specified hold-out size.

2.3 Next-state and compositional training

We train models on paths of (T) transitions, $(s_0, a_0, s_1, \ldots, a_{T-1}, s_T)$, where $s_{t+1} = a_t(s_t)$. At each step we sample uniformly from training transitions originating from the current state. Every transition in the training path belongs to the training set of transitions (no held-out transition is ever shown). Models receive only paths of observed local transitions, never the global constraint directly.

Under next-state training, we show intermediate states as inputs and for steps $1 \leq t \leq T$ train to predict s_t given $(s_0, a_0, s_1, \ldots, s_{t-1}, a_{t-1})$ as input (Figure 1c).

Under compositional training, we hide intermediate states from the inputs and for $1 \leq t \leq T$ train to predict s_t given $(s_0, a_0, a_1, \ldots, a_{t-1})$ as input, forcing the model to compose actions internally with T controlling the composition length. At $T = 1$, next-state and compositional training coincide.

2.4 Measuring inference complexity of a transition with proof depth

Certain held-out transitions are intuitively harder to infer than others. In the commutative world of Figure 4, both held-out transitions are semantically entailed, but only one can be immediately derived from observed transitions; the other requires first inferring an intermediate transition. We quantify this dependency by proof depth, defined relative to a fixed set of inference rules based on the world's algebraic constraints. Consider the rules of the commutative world depicted in Figure 4b. Knowing

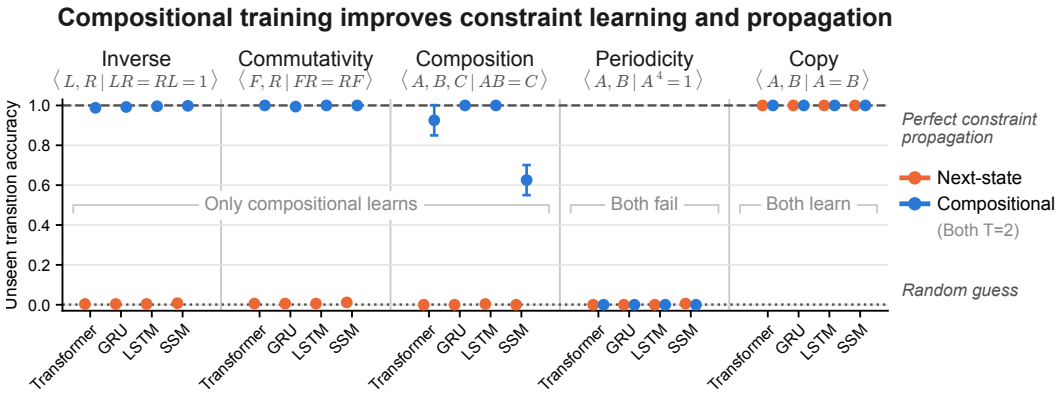


Figure 2: **Across monoid worlds and architectures, compositional training generalizes better than next-state training with equal optimizer steps.** Compositional objective propagates inverse, commutative, and composition constraints with average accuracy of 96%, and copy at 100%. Next-state training only succeeds on copy and fails on other tested families. Neither propagates periodicity. Points averaged over 5 seeds and bars show standard deviations.

transitions $(1, F, 0)$, $(0, R, 2)$, and $(1, R, 3)$ derives $(3, F, 2)$ since actions F and R commute, so the paths $1 \xrightarrow{F} 0 \xrightarrow{R} 2$ and $1 \xrightarrow{R} 3 \xrightarrow{F} 2$ must end at the same state. Constraint cascading then happens in rounds, each round using the derived transitions of previous rounds, then inference rules infer transitions (Figure 4b). The minimal number of rounds to infer a transition is its proof depth.

3 Constraint Propagation in Controlled Monoid Worlds

We construct monoid worlds with $N = 1024$ states satisfying specified constraints and train models on 95% of state-action transitions $(s, a, a(s))$ such that the remaining 5% of held-out transitions, which we evaluate model generalization on, are entailed by seen transitions and the world's constraint. Unless stated otherwise, models have four layers of width 64 and train for 150k steps at $T = 2$ with reported results averaged over five seeds.

3.1 Next-state training fails to propagate constraints

Under a next-state objective, where intermediate states are shown, transformers on inverse worlds fail to generalize with accuracy at chance (Figure 1d.1). Scaling model, world, and compute does not save next-state training, and models continue to fail at propagating inverse constraints (Appendix A).

Across attention, recurrent, and state-space architectures, next-state training $T = 2$ fails, indistinguishable from chance, at propagating inverse, commutative, composition, and periodicity constraints in their associated worlds (Figure 2). Only in the trivial copy world where actions are identical ($A = B$) does next-state training generalize.

We find a systematic "copy" failure mode of general constraint propagation by next-state trained models, where models behave as if every world were the copy world. Averaged across architectures and worlds, 64% of incorrectly inferred unseen transitions are explained by copy propagation, assuming the unseen $(s, A, A(s))$ transition matches seen $(s, B, B(s))$, although seen transitions contradict the $A = B$ constraint at almost every state (Appendix B). These results suggest the next-state objective poorly propagates constraints from data.

3.2 Compositional training improves constraint propagation

Under a compositional objective, where intermediate states are hidden, transformers successfully propagate inverse constraints, generalizing unseen transitions with $\geq 93\%$ accuracy (Figure 1d.1). We find a sharp jump in generalization of 98% when varying the composition length from $T = 1$, a reduction to next-state training, to $T = 2$, when training first includes two-action compositions.

Across architectures, compositional training at $T = 2$ enables models to propagate inverse, commutativity, and composition constraints with an average accuracy of 96% and copy at 100% (Figure 2). Extending composition length to $T = 4$ yields nearly identical generalization results of 96% averaged over the same worlds (Appendix D). Periodicity constraints, however, are not propagated by next-state or compositional training at $T = 2$ or 4. Across periodicity worlds $A^k = 1$ for $k = 3, 4, 5$, compositionally trained transformers at $1 \leq T \leq 8$ all fail to generalize, suggesting that the failure is not due to training paths too short to contain the constraint. Compositionally trained models also have reduced "copy" failure modes, hallucinating a copy constraint for 36% of failures (Appendix B).

Compositional training fits observed transitions an order of magnitude earlier than it recovers held-out ones. On the inverse world, transformers learn 100% of seen transitions at 1.2k steps but achieve 50% generalization at 16k, approximately 13 times as many training steps, evidence of a grokking-like delay (Power et al., 2022). Finally, generalization accuracy, fixing world size, increases monotonically with "constraint evidence", the number of seen transitions satisfying the constraint (Appendix C).

4 Constraint propagation in vision and language

We explore learning and propagating constraints towards geometric generalization in world models trained on visual environments and relational generalization in LLMs trained on knowledge-based tasks. We identify domain-relevant constraints, which often hold only approximately, and build entailed hold-outs which test whether models propagate the constraints.

4.1 Geometric constraints in visual world models

An agent moves (L)eft/(R)ight/(F)orward/(B)ackward through an AI2-THOR virtual environment, generating a dataset of observation-action paths of length $T = 2$ (Figure 3a.1) (Kolve et al., 2022). Euclidean geometry imposes constraints on motion requiring that the final destination be independent of the order of translations: stepping forward and then right or right and then forward takes you to the same endpoint. Actions are also invertible since you can undo a previous action by stepping in the opposite direction. In environments with object collisions, these constraints do not hold for all states. We can compute these approximate constraints by identifying paths which start from the same initial observation and recording whether certain action sequences converge to the same endpoint observation. If two action sequences frequently ($\geq 50\%$) map between the same initial and terminal observations, we declare them an approximate constraint. We find the Euclidean constraints: inverse $LR = 1$ and commutativity $FR = RF$, which hold roughly 80% of the time.

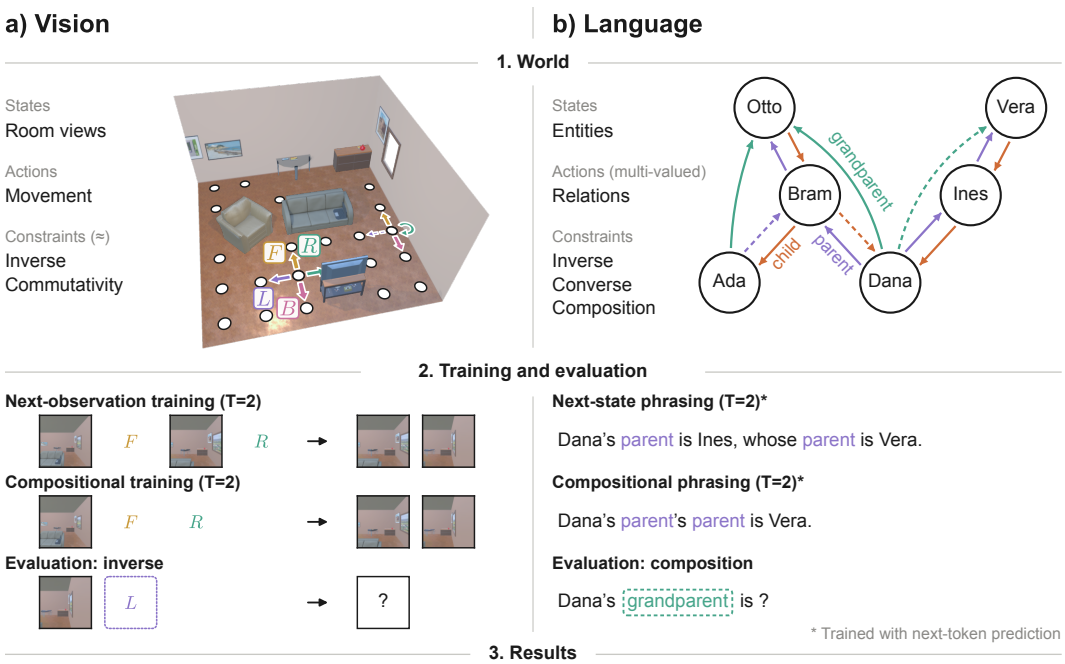
We hold out transitions $(\text{obs}, \text{action}, \text{next obs})$ that test propagation of inverse and commutative constraints. These hold-outs are taken at states away from object collisions where constraints are valid. Transformer world models are trained to predict the next observation given previous observations and actions. Accuracy is measured by choosing the observation in the dataset with the highest predicted likelihood. Since visual similarity may enable held-out recovery without learning geometric constraints, we compare to a baseline of shifting observation pixels.

Under next-observation training, where intermediate observations are shown, world models poorly generalize inverse and commutative constraints, with an accuracy of only 50% and 51%, respectively, compared to the pixel-shift baseline of 41% and 44% (Figure 3a.2, 3a.3). Trained compositionally, hiding intermediate observations, world models achieve 83% accuracy on inverse and 93% on commutative constraints, averaging a $+38\%$ improvement over next-observation trained models.

To examine recovery without visual shortcuts, we train transformers to predict the next state directly, instead of its observation. Generalization for next-state trained models on inverse and commutative constraints collapses to random guessing, while compositionally trained models achieve 94% and 98%, respectively (Figure 3a.3). These results suggest next-observation/state training poorly propagates inverse and commutative constraints in visual domains while compositional training successfully propagates geometric constraints even without visual cues.

4.2 Relational constraints in large language models

We examine when LLMs, trained on facts, propagate relational constraints without premises supplied in context. We study inverse constraints between successor and predecessor, the converse relation



Training	Inverse State	Inverse Observation	Commutative State	Commutative Observation
Baseline (pixel-shift)	–	41%	–	44%
Next-observation	0±0%	50±2%	0±0%	51±2%
Compositional	**94±1%**	**83±3%**	**98±1%**	**93±2%**

Model	Training	Inverse	Converse	Composition
Gemma 270M	Next-state	0%	0%	0%
	Compositional	**55%**	1%	**68%**
Qwen 600M	Next-state	0%	0%	0%
	Compositional	**11%**	0%	**79%**
Gemma 1B	Next-state	3%	1%	0%
	Compositional	**61%**	0%	**71%**
Gemma 4B	Next-state	3%	0%	0%
	Compositional	**53%**	3%	**72%**

Figure 3: **Compositional objectives improve constraint propagation over next-state-style training in vision and language a)** 1) AI2-THOR virtual 3D environment, where an embodied agent translates between observed states. Inverse $LR = RL = 1$ and commutative $FR = RF$ constraints hold approximately due to collisions. 2) Training and evaluation follow Figure 1c with image observations replacing abstract states. 3) Next-observation training marginally outperforms pixel-shifting baseline ($+8\%$) while compositional training improves by ($+46\%$) (observation). Training to predict state, minimizing visual cues, collapses next-state generalization to chance, while compositional training generalizes with 96% (state). Results averaged over three seeds. **b)** 1) Wikidata knowledge graph world, where entities are states and relations are multi-valued actions. 2) LLMs are finetuned with next-token prediction, the objective changing the sentence phrasing. 3) Compositional, not next-state, phrasing propagates inverse constraint at average accuracy of 45% and composition at 73%.

between parent and child, and the composition of parent and grandparent. Unlike inverse relations, converse relations branch, and the child of a parent may be a sibling, not necessarily a strict identity. For each constraint, we extract a Wikidata knowledge graph of 512 renamed individuals and associated relations (Figure 3b.1). We hold out transitions $(\text{person A}, \text{relation}, \text{person B})$ jointly entailed by training transitions and constraint. For example, the observed relationship $(\text{Dana}, \text{parent}, \text{Ines})$ entails the held-out relationship $(\text{Ines}, \text{child}, \text{Dana})$ under the converse constraint.

We finetune pretrained base LLMs (270M to 4B parameters) with next-token prediction on natural-language sentences generated from (T=2) training paths. We compare two phrasings the same underlying path. Next-state phrasing allows models to condition on intermediate entities, e.g., "Dana's parent is Ines, whose parent is Vera", while compositional phrasing omits intermediate entities, "Dana's parent's parent is Vera" (Figure 3b.2). Unlike previous compositional objectives, next-token prediction over compositional phrasing never uses intermediate entities as targets, so we include (T=1) paths in both conditions. We evaluate on a held-out relation such as $(\text{Ines}, \text{child}, \text{Dana})$ by prompting "Ines's child is" and choosing the highest-likelihood name from the graph. For relations

with multiple correct answers (converse and grandparent), we report filtered accuracy where all other correct answers (e.g., Dana's siblings) are excluded from the likelihood comparison.

Training with next-state phrasing fails to propagate inverse and composition constraints with an average generalization of 1%, marginally better than random guessing (Figure 3b.3). Compositional phrasing improves inverse propagation with an average accuracy of 45% and composition generalization at 73%. Converse is not propagated by next-state or compositional methods.

Composition learning over paths presents a shortcut: while "Dana's grandparent is Vera" is never shown during training, models see the $T = 2$ path "Dana's parent is Ines, whose parent is Vera" under next-state phrasing and "Dana's parent's parent is Vera" under compositional phrasing. Can propagation still occur when such shortcut paths are removed? We train a transformer on $T = 2$ paths in the monoid composition world $N = 1024$ with and without shortcuts (Appendix F). Next-state trained models fail in both settings, while compositionally trained models still recover without shortcuts (50% versus 92% with shortcuts). These results suggest that while shortcuts help compositionally trained models recover, they are not required for composition propagation.

To understand the effect of branching on converse constraint propagation, we construct an $N = 1024$ branching world that interpolates between line $\beta = 0$, where parent and child are true inverses, and binary tree $\beta = 1$, where each parent has two children and only a converse relationship remains (Appendix E). For comparison, parents in our Wikidata world have an average of 2.5 children. A transformer trained compositionally with $T = 2$ achieves 99% on $\beta = 0$, but adding only a 10% chance a parent might have another child ($\beta = 0.1$) drops accuracy to 63% and $\leq 3\%$ for $\beta \geq 0.75$.

Compositional phrasing improves inverse and composition constraint propagation but is limited in converse generalization on branching worlds.

5 How far do constraints propagate?

We measure an unseen transition's derivation complexity by proof depth and construct commutative monoid worlds of size $N = 1024$ with 10% of unseen transitions spread nearly uniformly across depths $d = 1$ to 8 (Figure 4a). Previous results concerned only $d = 1$ generalization. We now explore the effect of proof depth on constraint propagation by compositionally trained transformers. We train with composition lengths $T = 1, 2, 3, 4$ ($T = 1$ is next-state training) for 1.5M steps over five seeds.

5.1 Constraint propagation decreases sharply with proof depth

We find generalization of compositionally trained models decreases sharply with proof depth, while next-state trained models do not generalize at any proof depth. A transformer trained with composition length $T = 2$ recovers $d = 1$ transitions, whose derivation depends directly on seen transitions, at 95%, while $d = 2$ recovery, transitions whose derivation also depends on unseen $d = 1$ transitions, is only at 49%. The steep decreasing trend in generalization continues to deeper transitions with $d = 3$ at 15%, $d = 4$ at 5%, and $d = 5$ onward indistinguishable from random chance. Similar step-to-step accuracy ratios over depths $1 \leq d \leq 4$ suggest approximate geometric decay with rate 0.4 ± 0.1. Even with compositional training, models struggle to propagate constraints deeply.

5.2 Propagation dynamics suggest internal constraint cascades

Models trained compositionally propagate constraints to infer transitions iteratively by proof depth: first $d = 1$ transitions are learned, then $d = 2$, $d = 3$, and so on (Figure 4d). Transformers trained with composition length $T = 4$ reach half their final accuracy for $d = 1$ at 26k training steps, $d = 2$ at 50k steps, $d = 3$ at 102k, and $d = 4$ at 236k, so each depth takes about twice as many steps as the one before ($2.1\times$, geometric mean). Monotonic ordering holds for all $1 \leq d \leq 8$, but spacing beyond $d = 4$ shrinks, and $d = 5$ to $d = 8$ all reach half their final accuracy between 364k and 376k steps.

Timing suggests that models may be propagating constraints in a qualitatively similar way to constraint cascades, where local rules are applied repeatedly to derive held-out transitions in order of proof depth d. Unlike general constraint propagation, a cascade predicts that transitions are inferred only after all their derivation premises, the other unseen transitions required for the derivation, are inferred (Figure 4b). We examine 150 checkpoints for each of 5 seeds trained over composition lengths $T = 2, 3, 4$ on

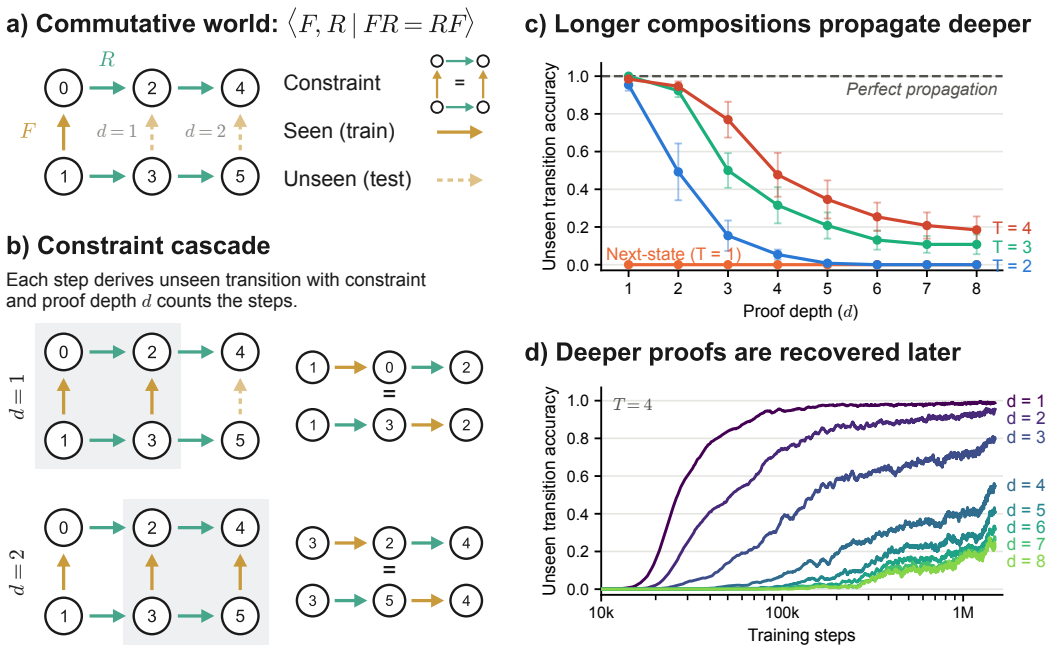


Figure 4: **Section of a commutative monoid world with unseen transitions of proof depths $d = 1$ and $d = 2$. Proof depth measures propagation complexity and predicts generalization and training dynamics. a)** Section of a larger commutative world containing two unseen transitions. **b)** A constraint cascade iteratively applies a simple derivation rule (here, commutative equations) to infer unseen transitions. Proof depth is the minimal number of iterations. The $d = 1$ transition is a premise for deriving $d = 2$ **c)** Generalization decreases geometrically with proof depth, but increasing composition length improves propagation to higher depths. **d)** Compositional training propagates constraints sequentially in proof-depth order, suggesting an internal constraint cascade.

the commutative world to find the percentage of inferred transitions missing their derivation premises. To remove timing confounds (on average $d = n + 1$ is learned after $d = n$), we compare to a null which shuffles the identity of $d = n$ transitions, preserving the timing but removing any dependency structure. We find that only 3% of correct inferences made by the transformer were made without having correctly inferred their derivation premises, compared to 11% for the null model, a difference of $3.7\times$. Throughout training, the model's correct held-out predictions are almost always derivable from seen transitions and other predictions, satisfying the constraint cascade condition.

5.3 Longer composition length enables deeper propagation

Training on longer composition lengths T with the same number of training steps improves generalization at higher proof depths (Figure 4c). Increasing T from 2 to 3 maintains $d = 1$ recovery at $\geq 95\%$ but significantly improves $d = 2$ by $+43\%$ and similarly improves recovery of $3 \leq d \leq 8$ with an average increase of $+19\%$. Raising composition length T from 3 to 4 maintains $d = 1$ and $d = 2$ recovery while again accumulating improvement for $3 \leq d \leq 8$ over $T = 3$ by $+14\%$. Decay curves also change qualitatively, from the geometric-like decay of $T = 2$, which is indistinguishable from random guessing by $d = 5$, to plateau-like behavior at $T = 3$ and $T = 4$, where generalization appears to flatten out at roughly 10% and 20%, respectively. These results suggest training on longer composition lengths may be key to deep propagation of constraints.

6 Discussion

Identical data, learned under a different objective, can turn a model that fails to generalize into one that learns and propagates global constraints. Across attention, recurrent, and state-space architectures, compositionally trained models propagate non-trivial constraints to recover unseen data while next-

state trained models do not. Compositional training also shows superior propagation in vision and language domains. Both objectives use identical path data, the only difference is that compositional training hides intermediate states. Global understanding therefore depends on how training requires local observations to be combined. Providing intermediate states encourages models to solve the task locally, while compositional training encourages integration across actions.

Training objective, perhaps more than architecture or task domain, determines whether a model propagates certain constraints. Across monoid worlds, constraints that are (not) propagated are consistently (not) propagated across the disparate architectures tested. Similarly, constraints propagated compositionally in monoid worlds are also propagated by compositional approaches in visual and language domains. This pattern suggests that controlled monoid worlds capture general learning biases that extend to richer tasks and may provide simplified environments for understanding and designing better-propagating algorithms.

Even when models propagate constraints, propagation is shallow, with generalization decreasing sharply with proof depth. Higher proof depth transitions are also recovered later and more slowly. Together with training dynamics, this suggests an internal process resembling a slow, inaccurate constraint cascade. Increasing composition length improves generalization to deeper consequences, suggesting that propagation reach depends on the organization of local information it learns from. Propagating constraint consequences deeply is essential for consistent information integration.

7 Related work

LLMs often learn facts without propagating their consequences. The reversal curse shows that training on one direction of a relation can fail to recover its reverse (Berglund et al., 2024). Similar failures exist in implicit reasoning, where models poorly compose relations (Balesni et al., 2025). Changes to training can address individual failures such as reversal (Ma et al., 2026) and composition (Feng et al., 2025) have been proposed, but remain specialized. We frame these failures and improvements in the context of learning and propagating constraints. Across controlled monoid worlds, visual environments, and knowledge-based language tasks, we study how objective determines which constraints models propagate. We further measure how far consequences are propagated using proof depth and explore associated training dynamics.

Accurate local predictions do not imply coherent global understanding (Vafa et al., 2024). Previous work improves consistency through reverse-prediction training (Kwon et al., 2026) or architectural modifications that recover group structures in synthetic environments (Keurti et al., 2023). Our constraint propagation framework extends beyond group structure to include semantic constraints such as the converse relation between parent and child.

Training without intermediate states has improved generalization in knowledge-graph embedding models (Guu et al., 2015) and discouraged local prediction shortcuts in planning (Bachmann & Nagarajan, 2025). We study how training objective affects constraint propagation across architectures, visual environments, and language tasks. Compositional training improves propagation for many geometric and semantic constraints but fails on others. Even when immediate consequences are recovered, generalization decreases sharply with proof depth. Composition length increases propagation reach to higher proof depths, suggesting that how much local information models combine during training shapes how far they propagate what they learn.

8 Conclusion

We frame global understanding as learning constraints from local observations and propagating their consequences deeply. Our monoid worlds provide controlled environments where generalization is entailed by observed data and unobserved global constraints. Across architectures, next-state training fails to propagate non-trivial constraints, while compositional training substantially improves propagation. We apply our constraint framing to visual and language tasks and find corresponding improvements under compositional-style objectives as well as regularities in which constraints propagate and under what conditions across domains. Deep propagation remains limited, though longer composition lengths help. These results suggest that global understanding requires training methods that encourage models to integrate local observations and deeply propagate their consequences.

AI USE STATEMENT

In this work, we used generative AI tools to implement methods, including code that generates the vision and language worlds, extracts the Wikidata graphs, trains and evaluates the models, and organize our code base. Additionally, we used generative AI tools for improving readability, suggest improvements, and proofread drafts. We have reviewed all AI-assisted work. We reviewed and tested all AI-written code and reviewed every AI-suggested edit to the text. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The supplementary material contains source code for replication, the AI2-THOR rooms with their rendered observations, and the extracted Wikidata graphs. Scripts in it regenerate every run configuration, including all architectures, hyperparameters, and seeds, and reproduce every figure and table.

REFERENCES

Gregor Bachmann and Vaishnavh Nagarajan. The pitfalls of next-token prediction, 2025. URL https://arxiv.org/abs/2403.06963.

Mikita Balesni, Tomek Korbak, and Owain Evans. Lessons from studying two-hop latent reasoning, 2025. URL https://arxiv.org/abs/2411.16353.

Lukas Berglund, Meg Tong, Maximilian Kaufmann, Mikita Balesni, Asa Stickland, Tomek Korbak, and Owain Evans. The reversal curse: LLMs trained on "A is B" fail to learn "B is A". In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/5178b2f2d7c44aa390c0777dc77b3f0c-Abstract-Conference.html.

Jiahai Feng, Stuart Russell, and Jacob Steinhardt. Extractive structures learned in pretraining enable generalization on finetuned facts. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 16853–16883. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/feng25m.html.

Kelvin Guu, John Miller, and Percy Liang. Traversing knowledge graphs in vector space. In Lluís Màrquez, Chris Callison-Burch, and Jian Su (eds.), *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 318–327, Lisbon, Portugal, September 2015. Association for Computational Linguistics. doi: 10.18653/v1/D15-1038. URL https://aclanthology.org/D15-1038/.

Hamza Keurti, Hsiao-Ru Pan, Michel Besserve, Benjamin F Grewe, and Bernhard Schölkopf. Homomorphism AutoEncoder – learning group structured representations from observed transitions. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 16190–16215. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/keurti23a.html.

Eric Kolve, Roozbeh Mottaghi, Winson Han, Eli VanderBilt, Luca Weihs, Alvaro Herrasti, Matt Deitke, Kiana Ehsani, Daniel Gordon, Yuke Zhu, Aniruddha Kembhavi, Abhinav Gupta, and Ali Farhadi. Ai2-thor: An interactive 3d environment for visual ai, 2022. URL https://arxiv.org/abs/1712.05474.

Soonwoo Kwon, Jin-Young Kim, Hyojun Go, and Kyungjune Baek. Toward stable world models: Measuring and addressing world instability in generative environments. *Pattern Recognition*, 177: 113351, 2026. doi: 10.1016/j.patcog.2026.113351. URL https://doi.org/10.1016/j.patcog.2026.113351.

Xutao Ma, Yixiao Huang, Hanlin Zhu, and Somayeh Sojoudi. Breaking the reversal curse in autoregressive language models via identity bridge, 2026. URL https://arxiv.org/abs/2602.02470.

Alethea Power, Yuri Burda, Harri Edwards, Igor Babuschkin, and Vedant Misra. Grokking: Generalization beyond overfitting on small algorithmic datasets. *CoRR*, abs/2201.02177, 2022. URL https://arxiv.org/abs/2201.02177.

Keyon Vafa, Justin Y. Chen, Ashesh Rambachan, Jon Kleinberg, and Sendhil Mullainathan. Evaluating the world model implicit in a generative model. In *Advances in Neural Information Processing Systems*, volume 37, pp. 26941–26975, 2024. doi: 10.52202/079017-0846. URL https://papers.nips.cc/paper_files/paper/2024/hash/2f6a6317bada76b26a4f61bb70a7db59-Abstract-Conference.html.

A SCALING MODEL, WORLD, AND COMPUTE UNDER NEXT-STATE TRAINING DOES NOT ENABLE CONSTRAINT PROPAGATION

We explore whether scaling a transformer under next-state training can enable constraint propagation on the inverse world by individually scaling model width from 64 to 320 (0.33M to 5.6M parameters), world size from 1024 to 8192 states, and training time from 15k to 300k steps. We find that every next-state trained model fits the observed transitions but fails to recover held-out transitions (all at random floor). The compositionally trained control does generalize, achieving 97% on held-out inverse transitions.

Table 1: **The next-state objective does not propagate constraints even after scaling model, world, and compute.** Transformers trained on inverse world at $T = 2$. Every row is next-state training except the last, which trains the same world, training paths, and model but compositionally. Accuracies are average $\pm$ standard deviation over five seeds.

	Width	Parameters (M)	States (N)	Steps	Seen (%)	Unseen (%)
Model size	64	0.33	1024	150k	100.0	0.4 ± 0.5
	128	1.05	1024	150k	100.0	0.4 ± 0.5
	256	3.67	1024	150k	100.0	0.2 ± 0.4
	320	5.57	1024	150k	100.0	0.4 ± 0.5
World size	320	5.57	1024	150k	100.0	0.4 ± 0.5
	320	7.54	4096	150k	100.0	0.1 ± 0.1
	320	10.16	8192	150k	100.0	0.1 ± 0.1
Training time	320	10.16	8192	15k	100.0	0.1 ± 0.1
	320	10.16	8192	75k	100.0	0.1 ± 0.1
	320	10.16	8192	150k	100.0	0.1 ± 0.1
	320	10.16	8192	300k	100.0	0.1 ± 0.1
Compositional control	320	10.16	8192	150k	100.0	97.0 ± 0.5

B MODELS INCORRECTLY PROPAGATE COPY CONSTRAINTS IN CONFLICT WITH SEEN TRANSITIONS

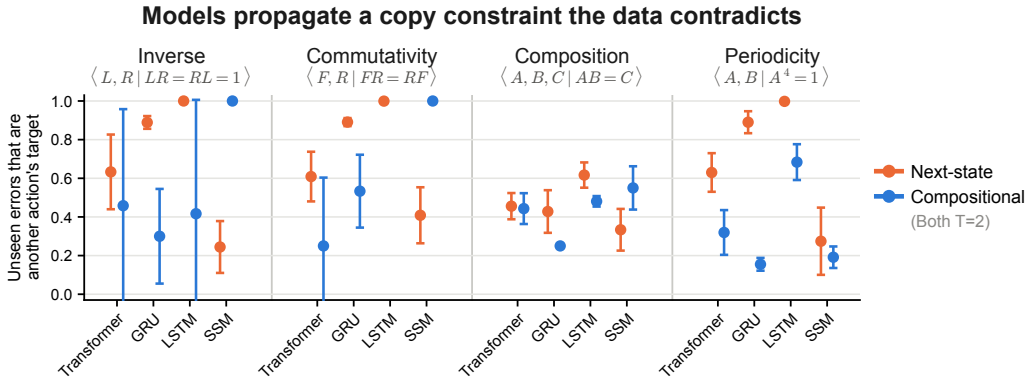


Figure 5: **Models systematically fail by recovering unseen transitions with another action's target.** When models fail to recover an unseen transition $(s, A, A(s))$, they instead often predict $B(s)$ for another action B. Points averaged over 5 seeds and bars show standard deviations.

When models trained with either the next-state or the compositional objective fail to propagate a constraint, they fail in a systematic way. For an incorrectly answered unseen transition $(s, A, A(s))$, we count a copy when the model predicts $B(s)$, the target of another action B from the same state.

Across inverse, commutative, composition, and periodicity worlds and all four architectures, 64% of next-state errors are copies, compared with 36% of compositional errors. Copy rate depends on architecture, with the highest rates (all next-state trained) being LSTMs at 90%, GRUs at 77%, and transformers at 58%.

C Compositionally trained models propagate constraints in proportion to observed evidence

We vary constraint evidence, the number of seen transitions satisfying the constraint, while fixing the world size. We train a transformer at $T = 2$ over the composition world $N = 1024$, varying the amount of held-out C-transitions from 10% to 97% (evidence levels of 922 and 31, respectively). Next-state trained models do not generalize, while compositionally trained models generalize monotonically with evidence: 2%, 4%, 11%, 24%, 62%, and 93% at evidence 31, 62, 128, 256, 512, and 922.

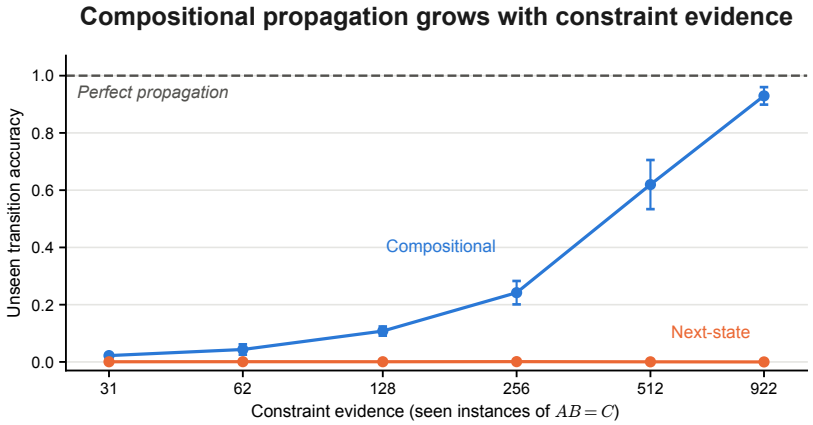


Figure 6: **Compositionally trained transformers recover unseen transitions in proportion to constraint evidence.** Transformers trained on the composition world $N = 1024$ with $T = 2$ paths. Points averaged over 5 seeds and bars show standard deviations.

D Composition length has minimal effect on recovery of immediately derivable (proof depth $d = 1$) transitions

We explore whether increasing composition length improves constraint propagation to unseen transitions that are immediately derivable from seen ones ($d = 1$). We train every compositional model of Figure 2 at $T = 4$ in addition to $T = 2$, keeping the worlds, held-out transitions, and 150k training steps fixed. Compositional path length is the only difference. All unseen transitions have proof depth $d = 1$. Compositionally trained models at $T = 4$ propagate inverse, commutative, and composition constraints with an average accuracy of 96%, the same as at $T = 2$, and copy at 100% (Figure 7). Neither $T = 2$ nor $T = 4$ propagates periodicity.

E Converse constraint propagation decreases sharply with branching

We explore how branching affects converse constraint propagation by building synthetic family trees that interpolate between purely inverse $\beta = 0$ and purely converse constraints $\beta = 1$. The tree is generated breadth-first with $N = 1024$, where each parent has one guaranteed child plus another with probability β (expected number of children is $1 + \beta$). We hold out 52 parent and 52 child relationships (5%) which are jointly entailed by observed transitions and the converse constraint. We train a transformer on $T = 2$ paths for 150k steps with the compositional or the next-state objective

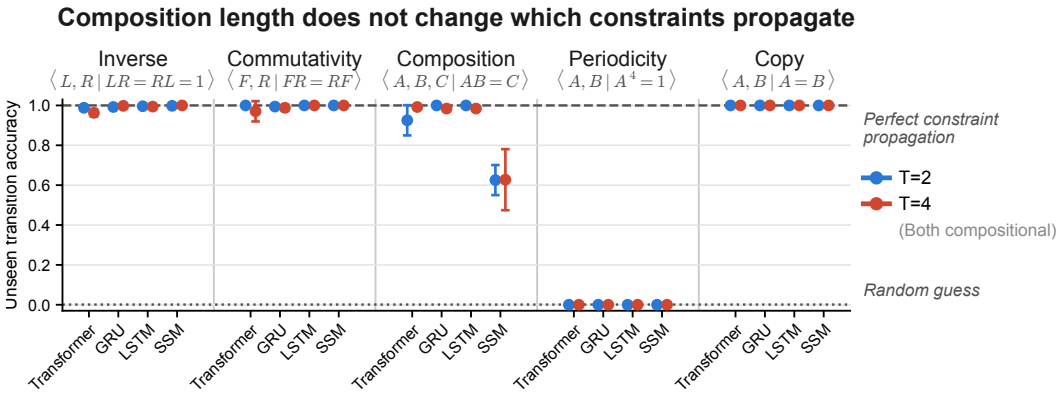


Figure 7: **Composition length has minimal effect on propagating immediately derivable transitions.** Compositional training at $T = 4$ propagates inverse, commutative, and composition constraints with an average accuracy of 96%, the same as at $T = 2$, and copy at 100%. Increasing composition length does not rescue periodicity. Points averaged over 5 seeds and bars show standard deviations.

on branching worlds of various β. Compositional training on the line world (pure inverse $\beta = 0$) generalizes at 99% accuracy, but generalization drops to 63% at $\beta = 0.1$, 35% at $\beta = 0.2$, 13% at $\beta = 0.5$, and $\leq 3\%$ for $\beta \geq 0.75$ (Figure 8). Next-state training fails for all β.

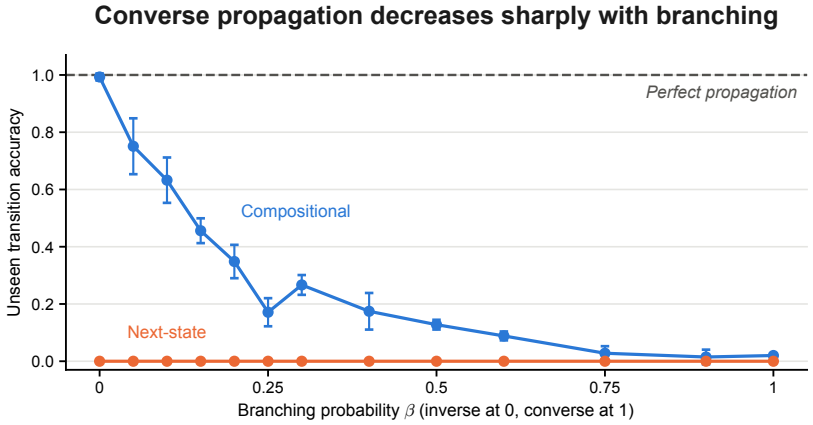


Figure 8: **Converse constraint propagation decreases sharply with branching.** Compositional training propagates the inverse constraint ($\beta = 0$) at 99%, but only a 10% chance of a second child ($\beta = 0.1$) drops generalization to 63%, and for $\beta \geq 0.75$ accuracy drops to $\leq 3\%$. Next-state training fails at every β. Points averaged over 3 seeds and bars show standard deviations.

F Composition propagation occurs without path shortcuts

We explore whether composition $AB = C$ propagates when shortcut paths $s \xrightarrow{A} A(s) \xrightarrow{B} C(s)$ for a held-out transition $(s, C, C(s))$ are removed from training. We consider the composition monoid world $N = 1024$ and hold out 10% of C-transitions. We train a transformer on $T = 2$ paths for 150k steps under the next-state or compositional objective with and without shortcut paths in the training dataset. Next-state trained models fail in both settings. Compositionally trained models generalize at 92% accuracy with shortcut paths and 50% without (Table 2).

Table 2: **Compositional training propagates composition without shortcut paths.** Transformers trained on the composition world $N = 1024$ at $T = 2$ with and without the shortcut path $s \xrightarrow{A} A(s) \xrightarrow{B} C(s)$ of each held-out transition. Accuracies are average $\pm$ standard deviation over five seeds.

Objective	Shortcut paths	Seen (%)	Unseen (%)
Next-state	with	100	0 ± 0
	without	100	0 ± 0
Compositional	with	100	92 ± 6
	without	100	50 ± 10

G PERIODICITY IS NOT PROPAGATED ACROSS COMPOSITION LENGTHS

Composition length is varied $1 \leq T \leq 8$ and $A^k = 1$ periodicity length $k = 3, 4, 5$ over $N = 1020$ periodicity worlds to see if compositional training can propagate the constraint. A training path contains a full wrap-around loop when $T \geq k$. All unseen transitions have proof depth $d = 1$. We train transformers with the compositional objective for 150k steps. Every model fits observed transitions but fails to propagate the periodic constraint (all at random floor) even with $T \geq k$ (Table 3).

Table 3: **Periodicity is not propagated across composition lengths.** All transformers trained compositionally on $N = 1020$ periodicity worlds with $A^k = 1$ for $k = 3, 4, 5$ and composition lengths $1 \leq T \leq 8$ fail to propagate the periodicity constraint. Accuracies are averaged over five seeds, with standard deviations of at most 0.5%.

	Unseen (%) at path length T							
Constraint	1	2	3	4	5	6	7	8
$A^3 = 1$	0.4 ± 0.5	0.4 ± 0.5	0.2 ± 0.4	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$A^4 = 1$	0.6 ± 0.5	0.2 ± 0.4	0.2 ± 0.4	0.2 ± 0.4	0.0 ± 0.0	0.2 ± 0.4	0.0 ± 0.0	0.2 ± 0.4
$A^5 = 1$	0.6 ± 0.5	0.2 ± 0.4	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0