How We Learn Structured Knowledge with Limited Resources

Hanqi Zhou

hanqi.zhou@uni-tuebingen.de

Cluster of Excellence "Machine Learning for Science", Tübingen AI Center,
University of Tübingen

Program induction

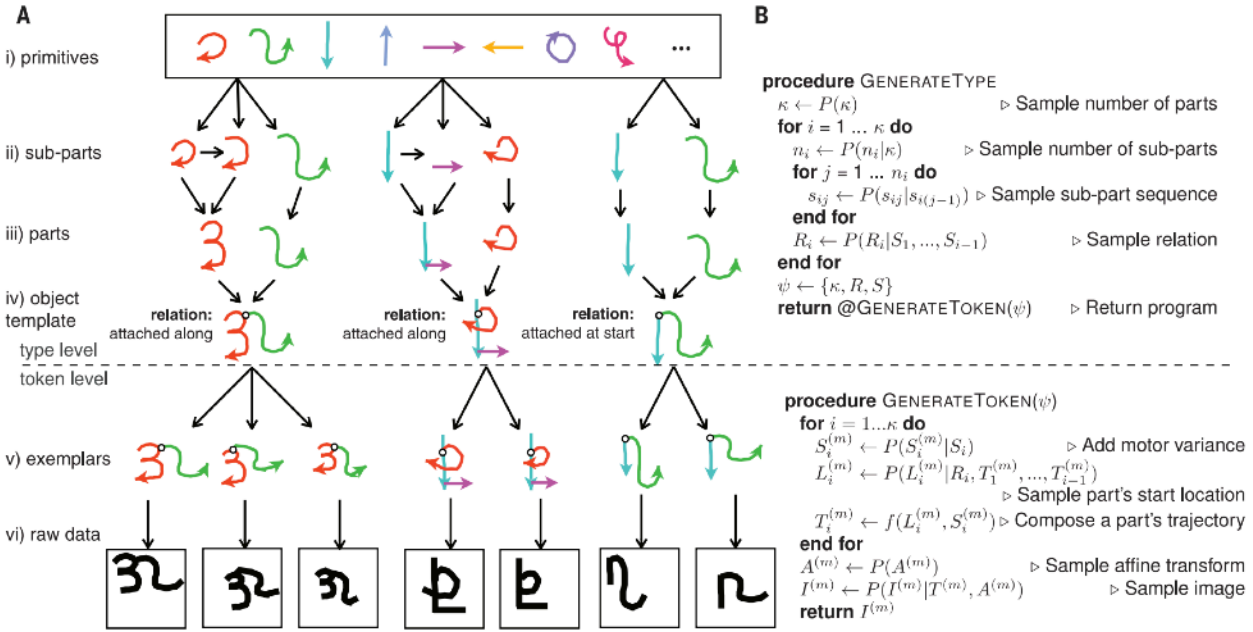


Fig. 3. A generative model of handwritten characters. (**A**) New types are generated by choosing primitive actions (color coded) from a library (i), combining these subparts (ii) to make parts (iii), and combining parts with relations to define simple programs (iv). New tokens are generated by running these programs (v), which are then rendered as raw data (vi). (**B**) Pseudocode for generating new types ψ and new token images $I^{(m)}$ for $m = 1, ..., M$. The function $f(\cdot, \cdot)$ transforms a subpart sequence and start location into a trajectory.

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. Science, 350(6266), 1332-1338.

Resource rationality

- **Resource:** we are resource constrained, e.g., attention, memory, …
- **Rationality:** we act to maximize our expected utility
- **Trade-off:** we make use of limited resources in an optimal or rational manner

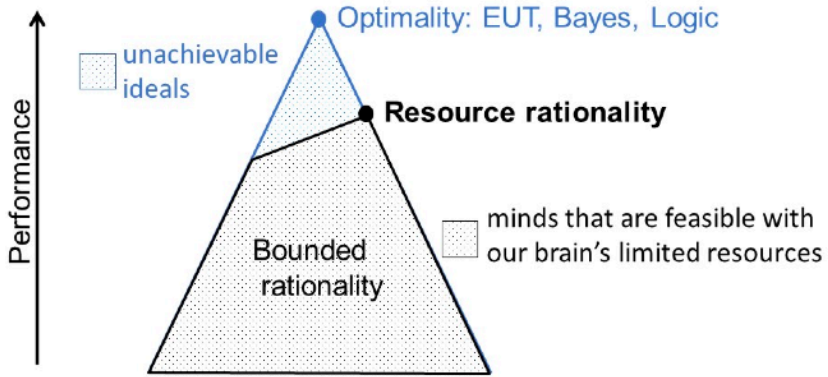


Figure: the best biologically feasible mind out of the infinite set of bounded-rational minds

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. Behavioral and brain sciences, 43, e1.

Resource rationality

Rate-distortion theory (RDT) is used as a computational model

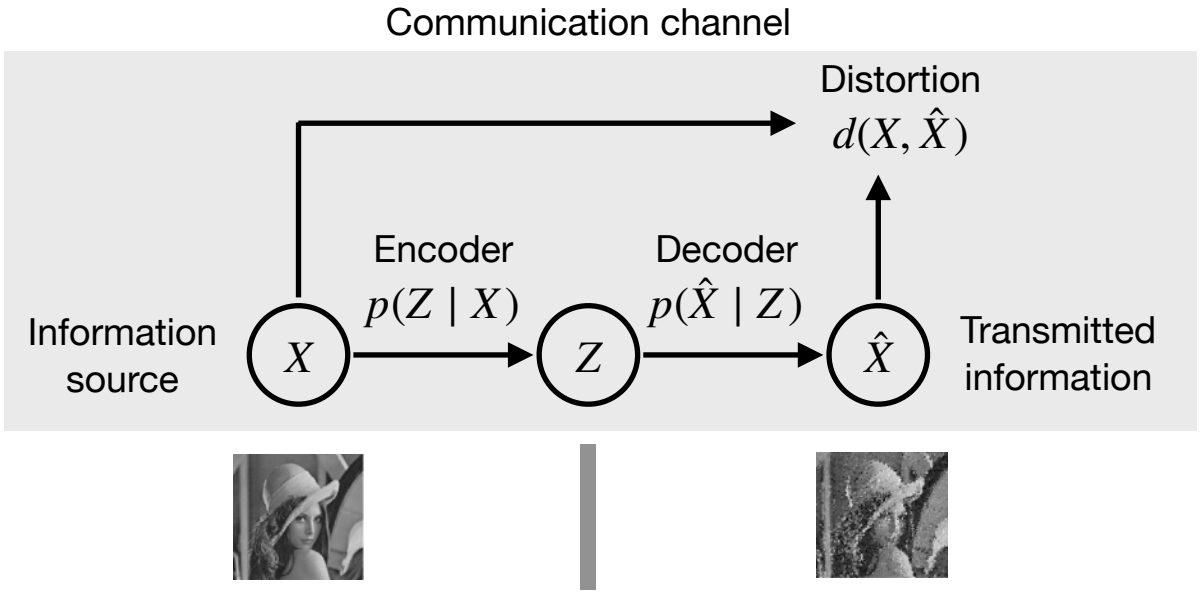
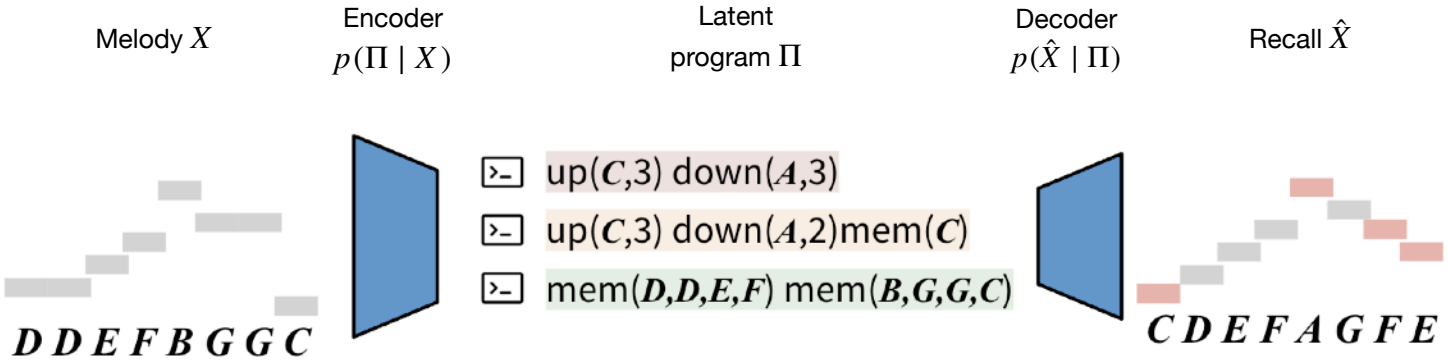
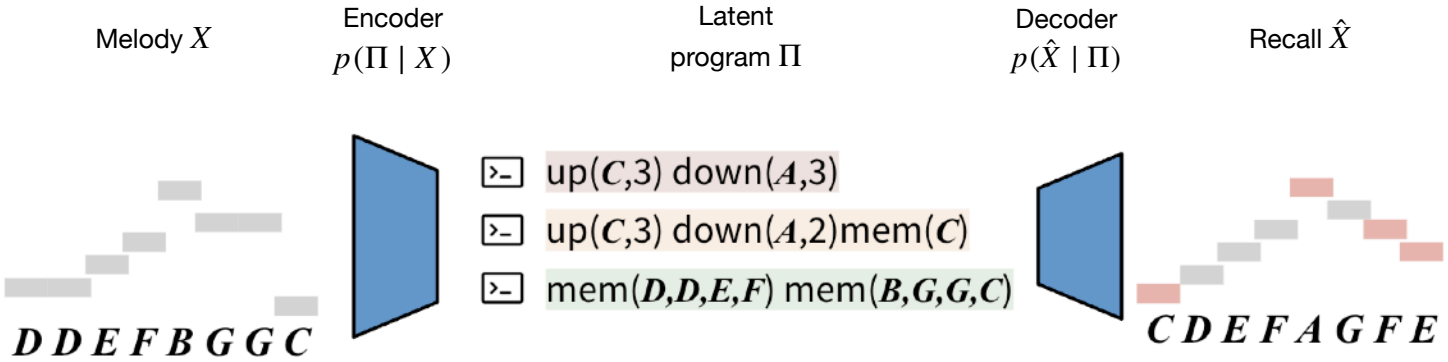


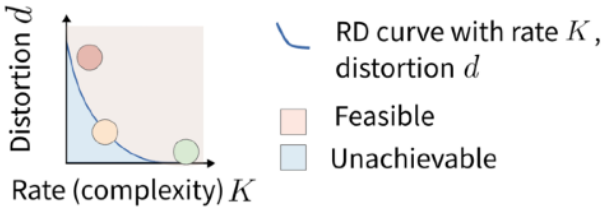
Figure adapted from Gershman, S. J. (2021). The rational analysis of memory. Oxford handbook of human memory.

Melody program induction under constraints



Melody program induction under constraints



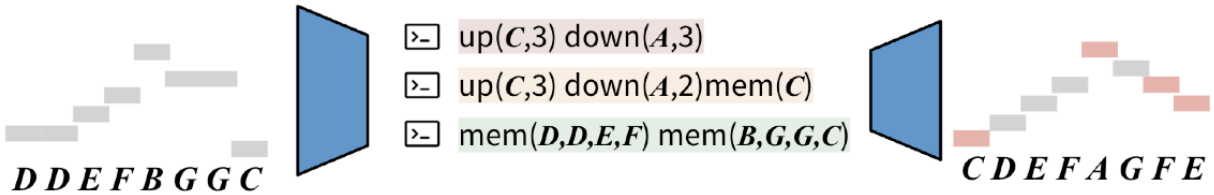


Melody program induction under constraints

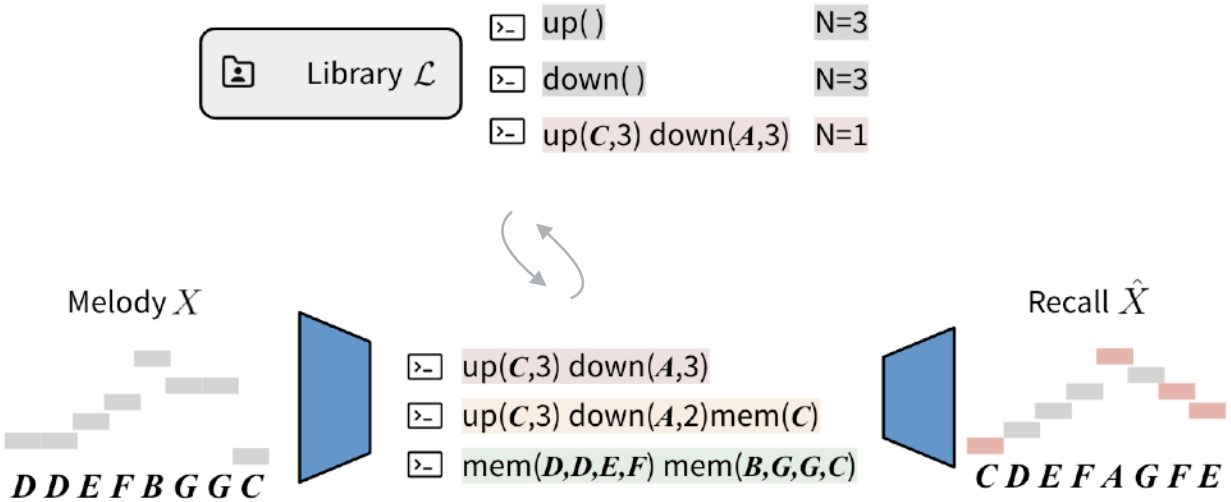
What is missing when modeling human learning with RDT?

- RDT: How can we transmit a message as efficiently as possible while minimizing information loss?

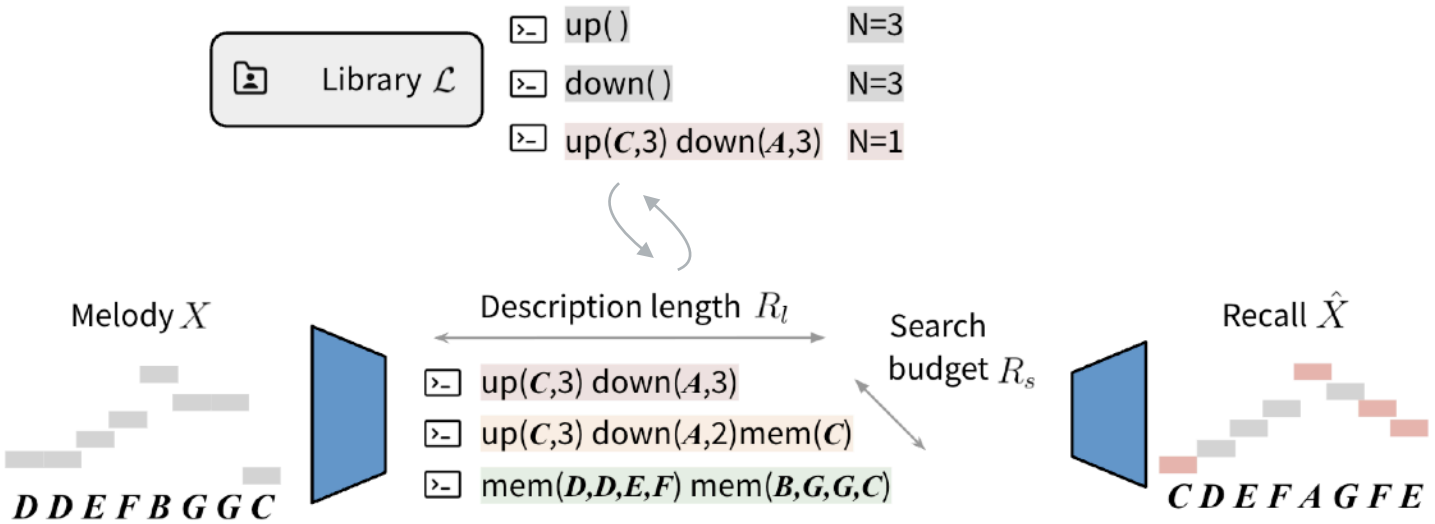
- Human: How do our internal compression models evolve as we learn? / How do we determine what constitutes the “optimal” encoding over time?



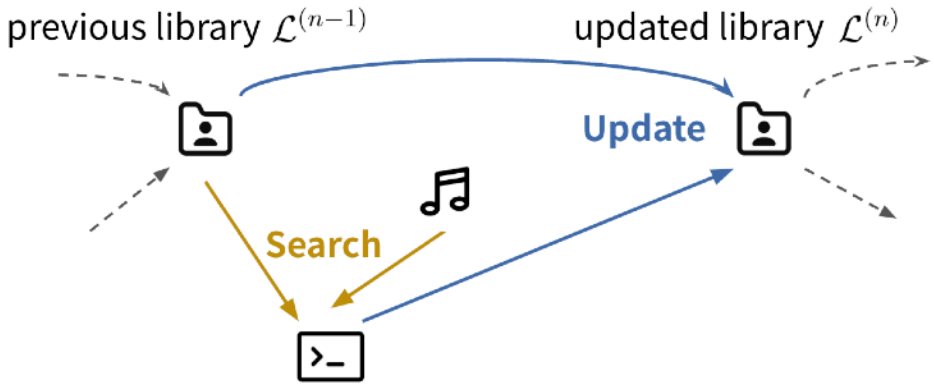
Melody program induction under constraints



Melody program induction under constraints



Melody learning empowered by libraries

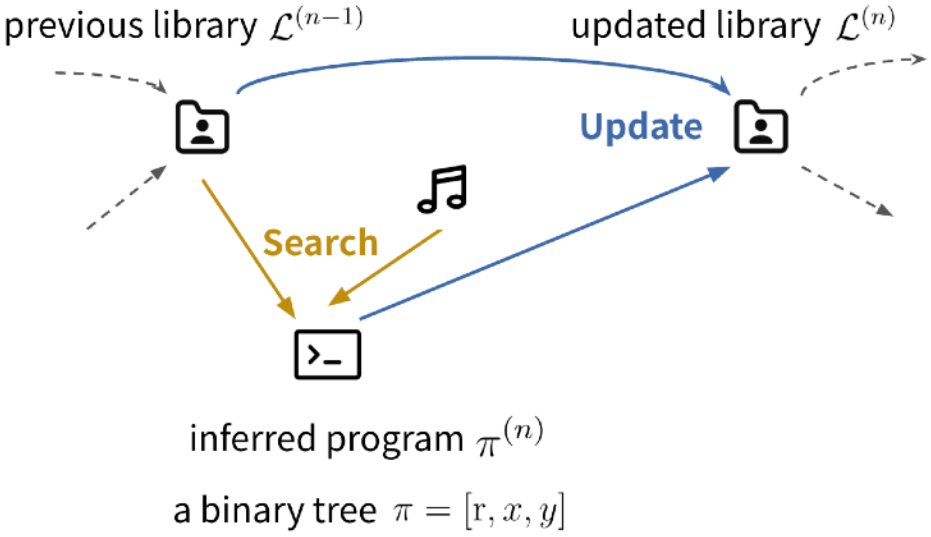


Melody learning empowered by libraries

Search for programs (within a melody):

- Description length R_l: An upper limit on the length of the programs
- Search budget R_s: An upper limit on how many programs are considered

$$p(\pi \mid X^{(n)}) \propto p(X^{(n)} \mid \pi)\, p(\pi)$$



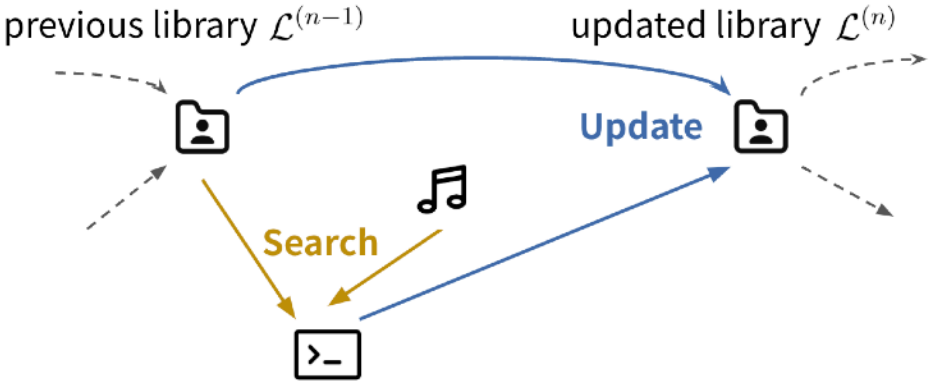
Melody learning empowered by libraries

Search for programs (within a melody):

- Description length R_l: An upper limit on the length of the programs
- Search budget R_s: An upper limit on how many programs are considered

$$p(\pi \mid X^{(n)}) \propto p(X^{(n)} \mid \pi)\, p(\pi)$$

Prior
$\propto$ Description length^{-1}



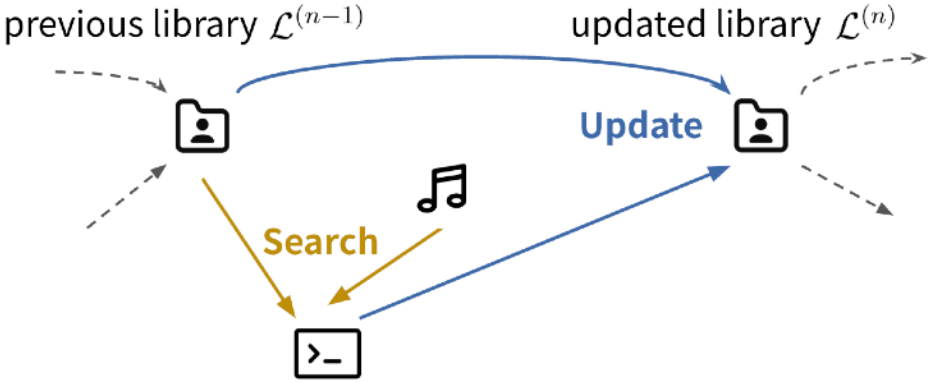
Melody learning empowered by libraries

Search for programs (within a melody):

- Description length R_l: An upper limit on the length of the programs
- Search budget R_s: An upper limit on how many programs are considered

$$p(\pi \mid X^{(n)}) \propto p(X^{(n)} \mid \pi)\, p(\pi)$$

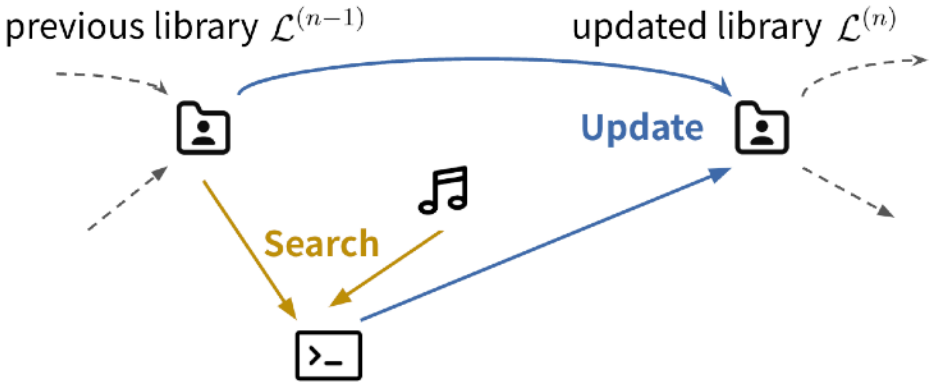
Likelihood $\propto$ Distortion^{-1}



inferred program $\pi^{(n)}$

a binary tree $\pi = [\mathrm{r}, x, y]$

Melody learning empowered by libraries



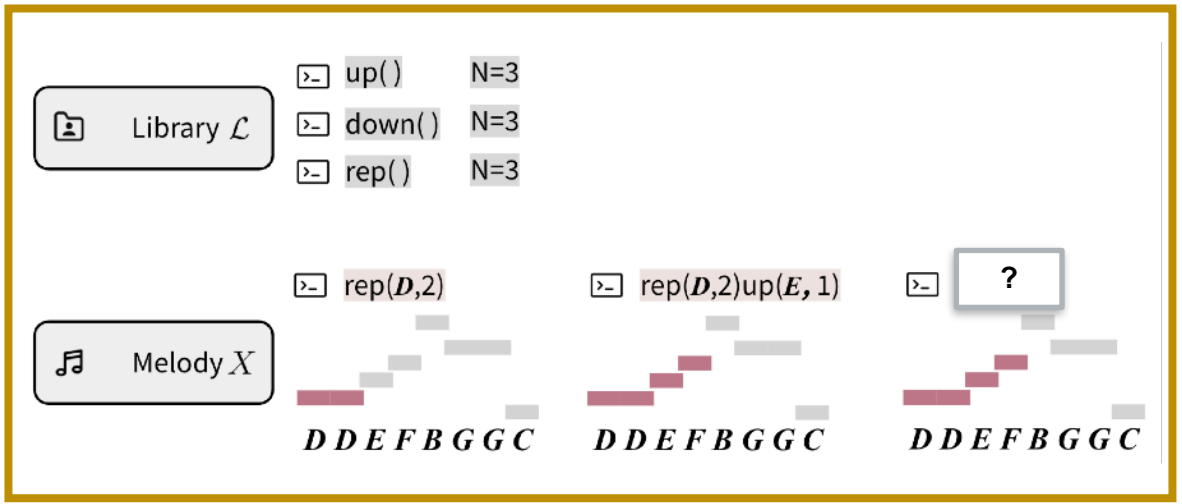
Update the library to prefer useful programs (across melodies):

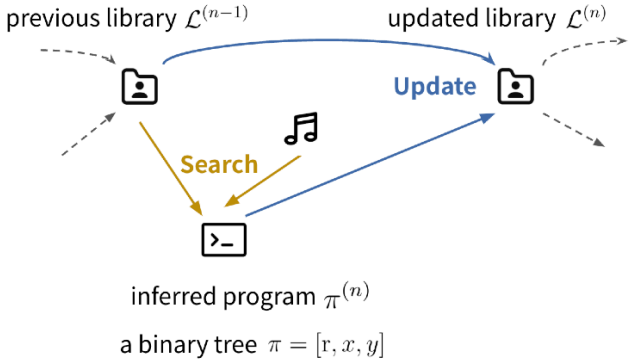
- Initial prior: proportional to simplicity
- Updated prior: proportional to usage (entropy coding)

$$p(\pi \mid X^{(n)}) \propto p(X^{(n)} \mid \pi)\, p(\pi \mid X^{(1:n-1)})$$

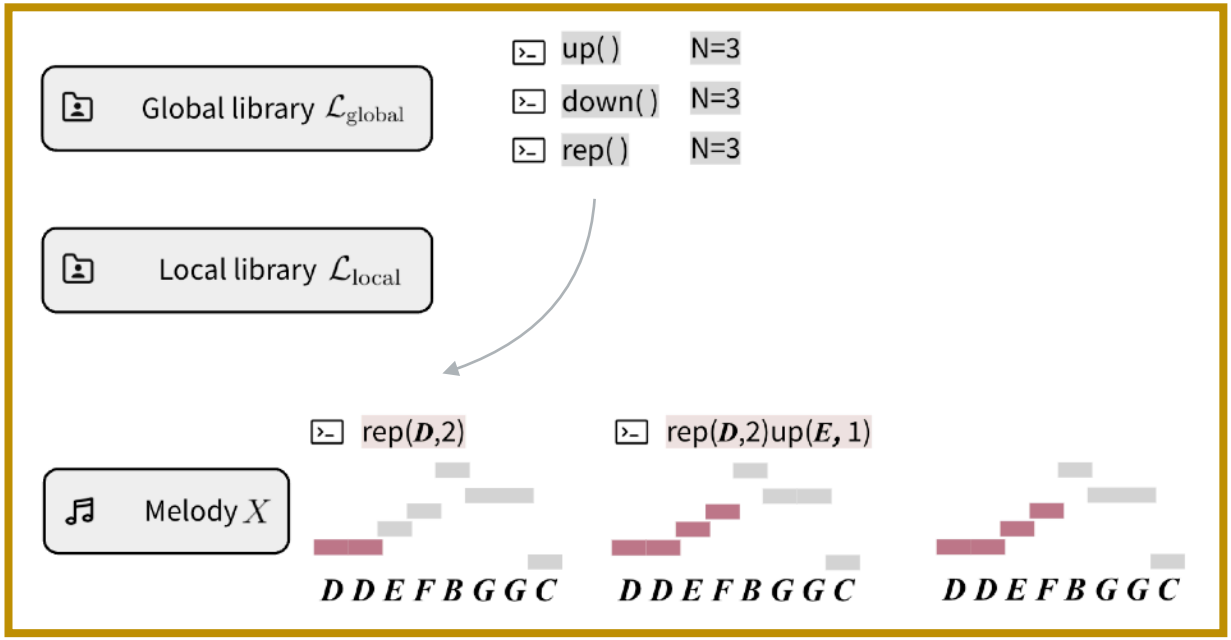
Updated prior
$\propto$ Usage frequency

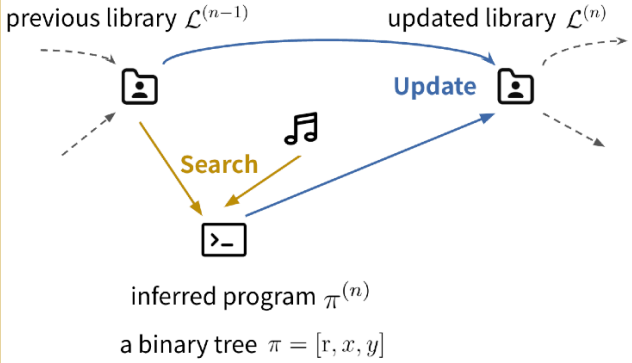
Augmented by 1) backtracking



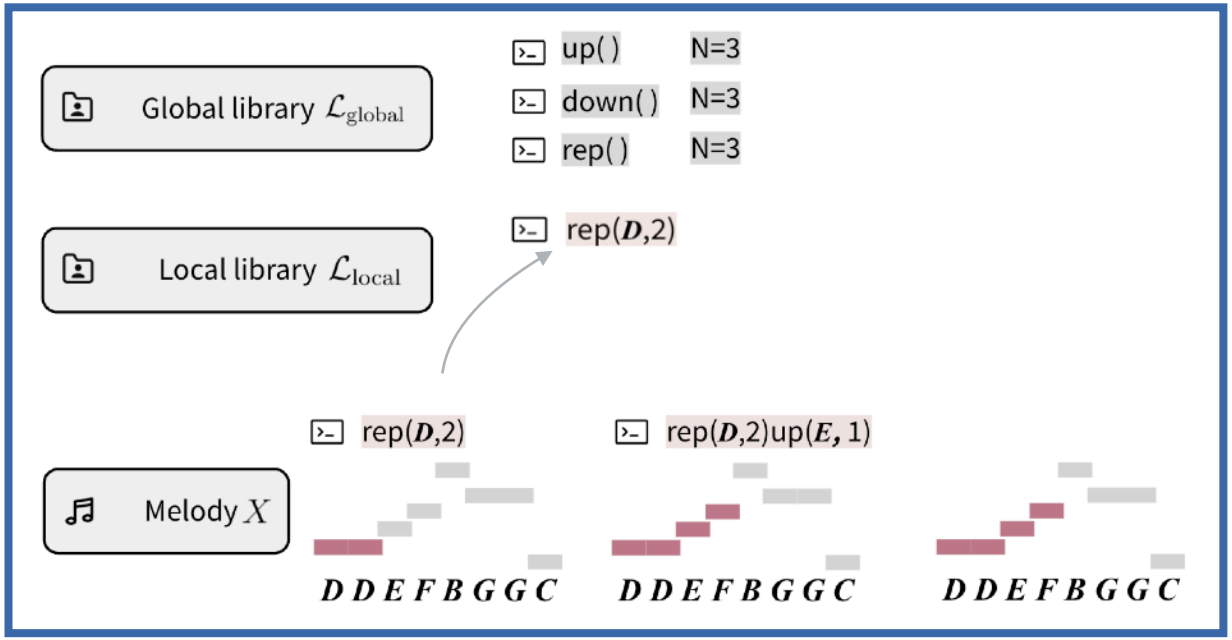


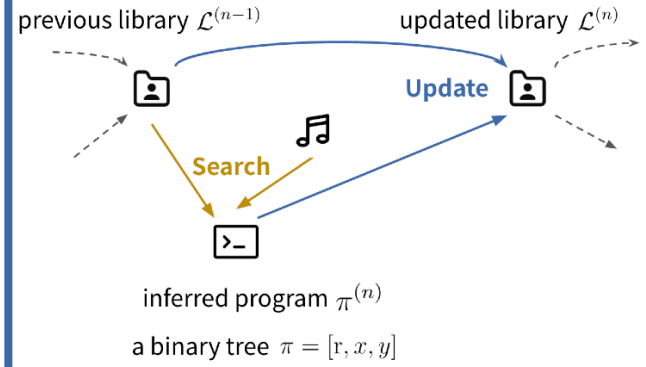
Augmented by 2) hierarchical library



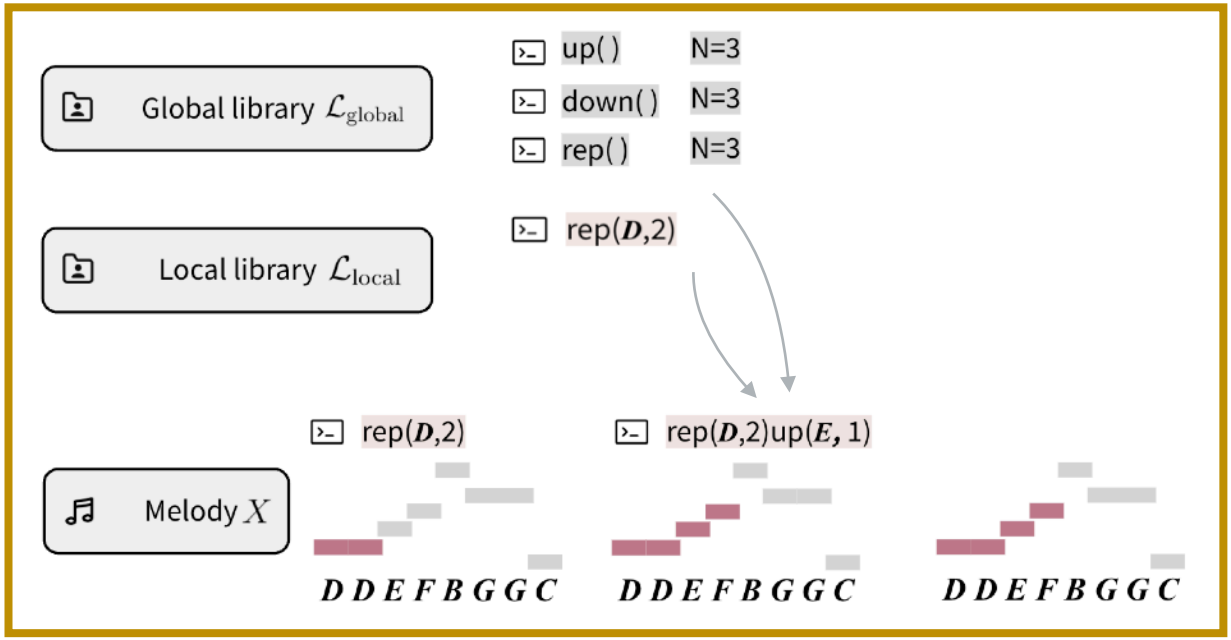


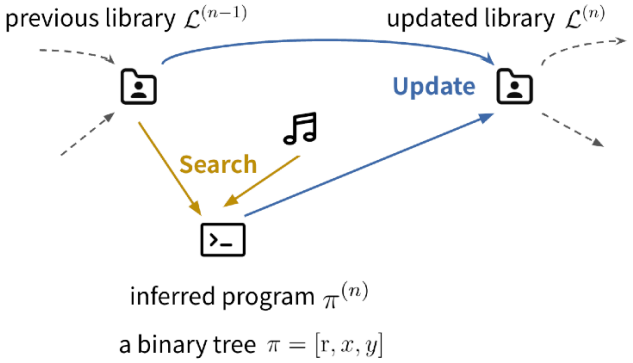
Augmented by 2) hierarchical library



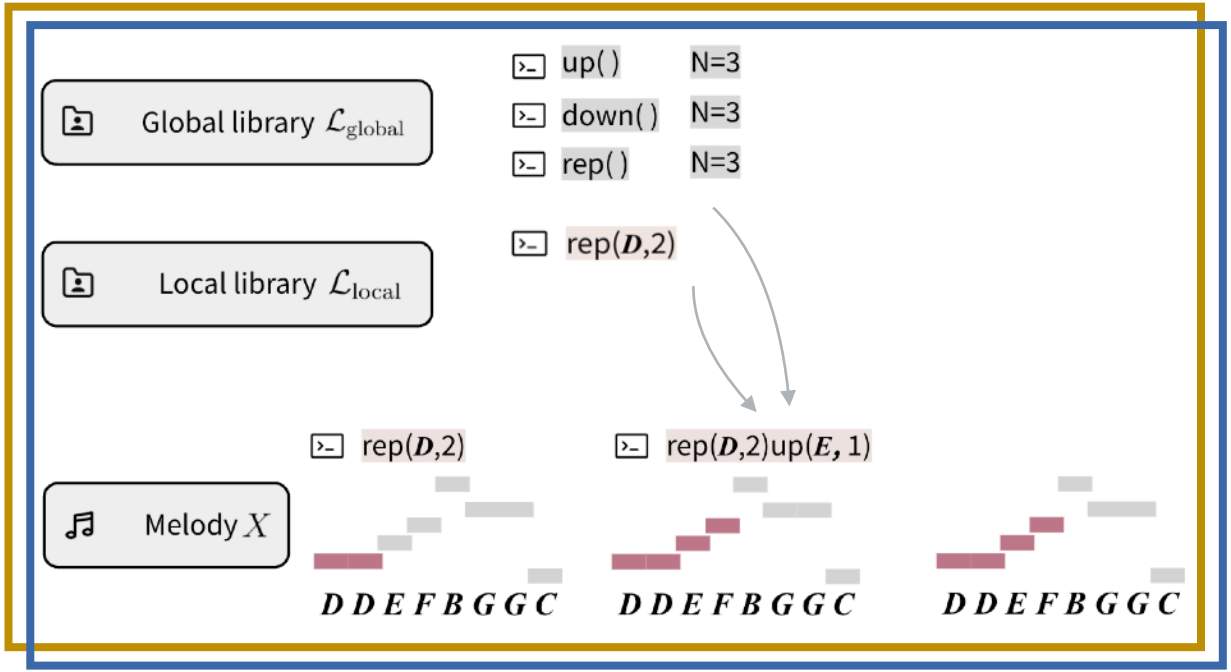


Augmented by 2) hierarchical library





Augmented by 2) hierarchical library



All melodic patterns

Pop, rock, classical, jazz

All English words

Fiction, history, autobiography

Blei, D. M., Griffiths, T. L., & Jordan, M. I. (2010). The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies. Journal of the ACM (JACM), 57(2), 1-30.

Experiment paradigm: ChirpChampion

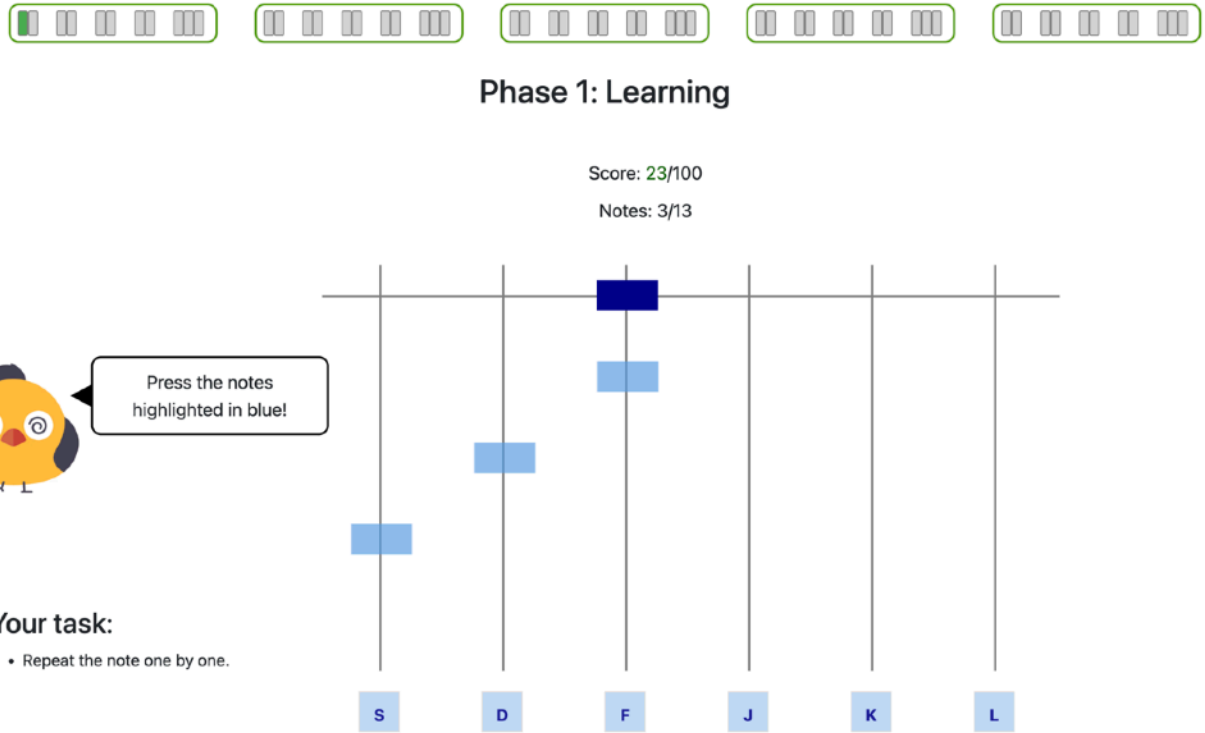
Learning - recalling - predicting

Welcome to Chirp Champions, where every bird has a song and every song tells a story!

You are a young bird in a vibrant, musical forest, eager to learn the enchanting songs of your community. Your bird teachers will guide you in learning the songs of the various birds species in your area.

Experiment paradigm: ChirpChampion

Learning - recalling - predicting



Experiment data

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

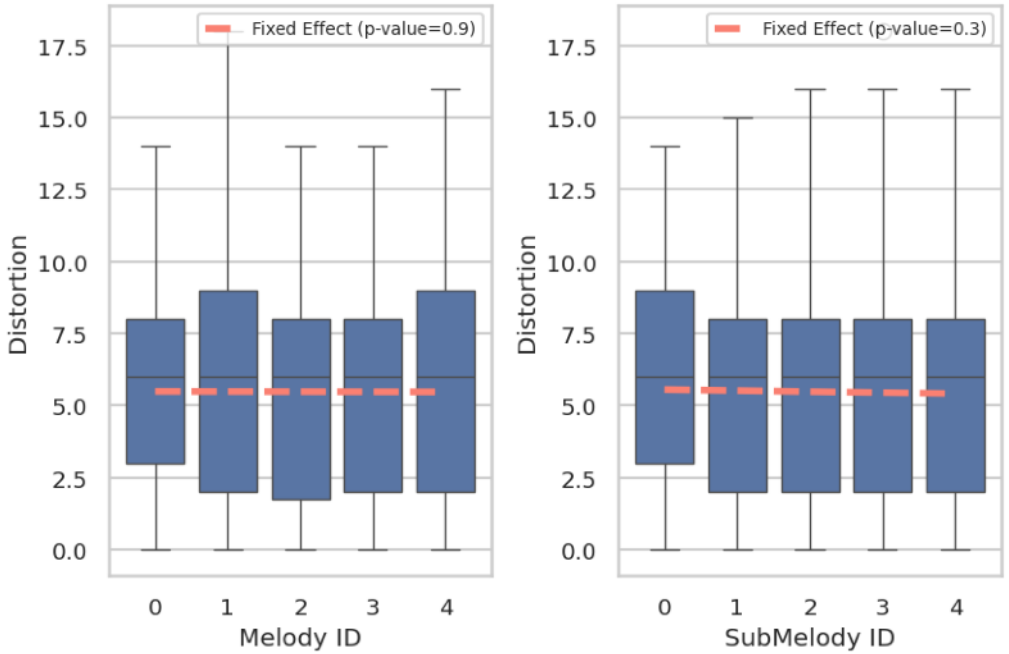
1. We do compression

- How can we explain performance (distortion in recall)?
- Hypothesis: Performance may improve with practice, but worsens as the melody becomes more ***complex***.

Experiment data

1. We do compression

- How can we explain performance (distortion in recall)?
- Hypothesis: Performance may improve with practice



Experiment data

1. We do compression

- Hypothesis: Performance worsens as the melody becomes more ***complex***.

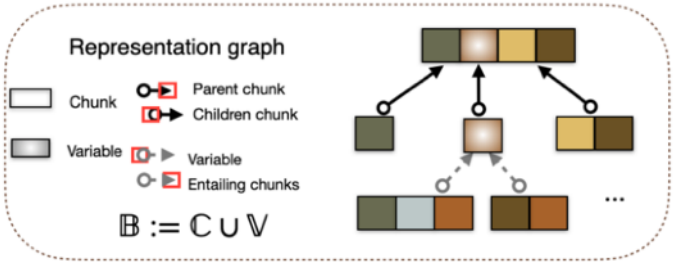
distortion ~

program complexity

+ chunking complexity

+ entropy

+ length



order-1 transition probabilities

Planton, S., van Kerkoerle, T., Abbih, L., Maheu, M., Meyniel, F., Sigman, M., ... & Dehaene, S. (2021). A theory of memory for binary sequences: Evidence for a mental compression algorithm in humans. PLoS computational biology, 17(1), e1008598.
Wu, S., Thalmann, M., Dayan, P., Akata, Z., & Schulz, E. (2024). Building, Reusing, and Generalizing Abstract Representations from Concrete Sequences. arXiv preprint arXiv:2410.21332.

Experiment data

1. We do compression

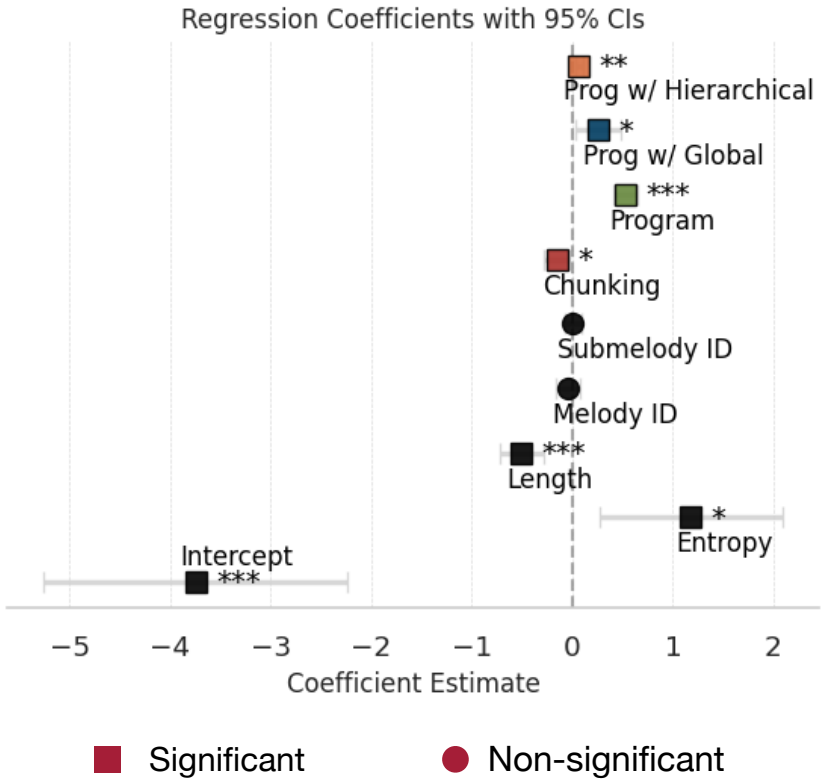
distortion ~

program complexity

+ chunking complexity

+ entropy

+ length



Experiment data

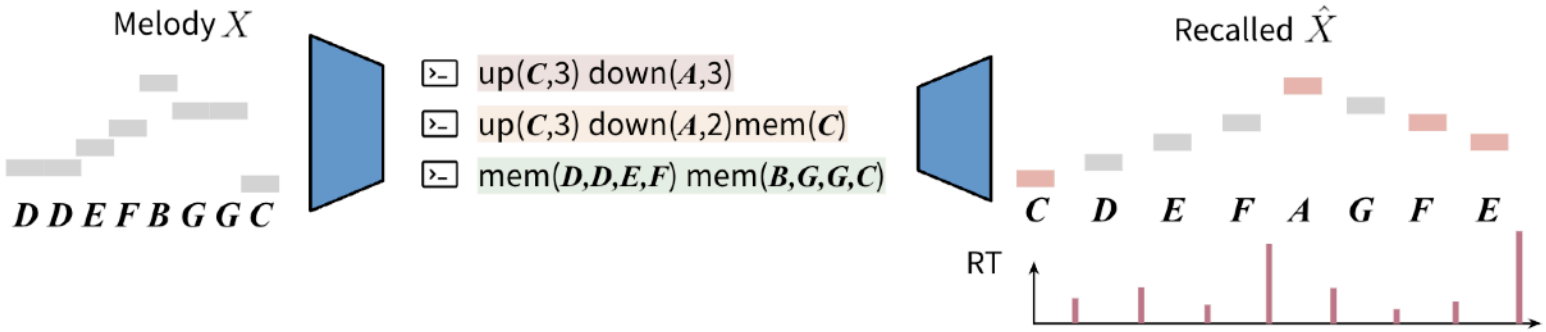
2. We do compression with programs

1) Reaction time: to identify ‘chunking’ boundaries, i.e., where participants naturally pause, indicating melody structure.

2) Error pattern: to identify whether participants make systematic errors.

Experiment data

2.1 Reaction time (human) vs. program boundary counts (model)



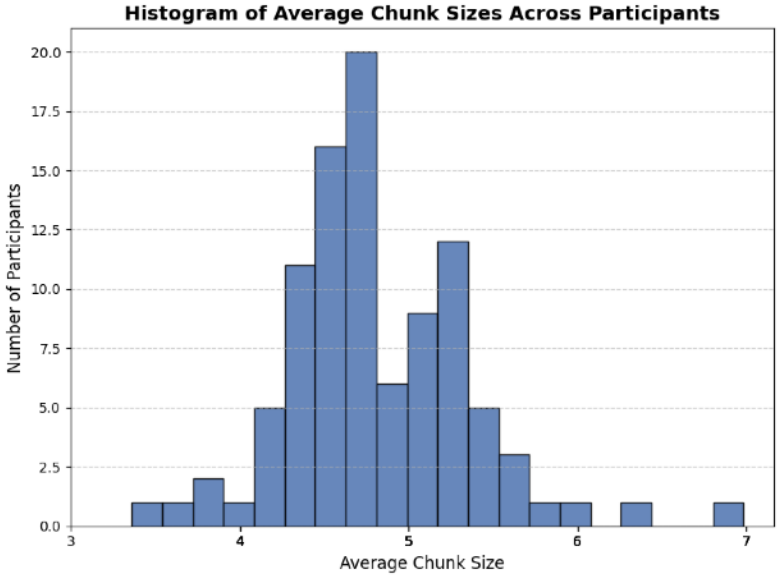
Experiment data

2.1 Reaction time (human) vs. program boundary counts (model)

Big changes in reaction time (RT) show when people are switching between mental “chunks”

- Compute first-order RT differences: $\Delta_i = RT_{i+1} - RT_i$
- Define a threshold to detect significant increases:

  Threshold $= \mu_\Delta + \lambda \cdot \sigma_\Delta$



Experiment data

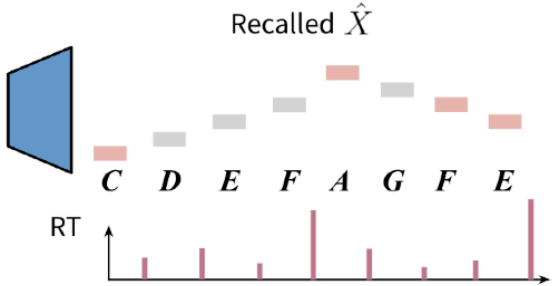
2.1 Reaction time (human) vs. program boundary counts (model)

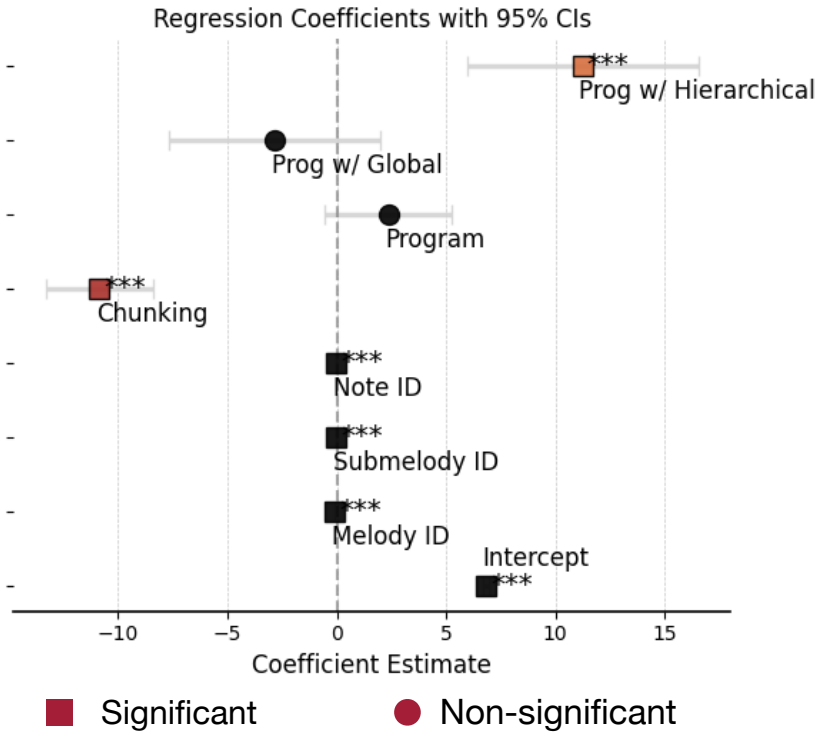
Reaction time (log space) ~

+ Program boundary

+ Chunking boundary

+ Note index

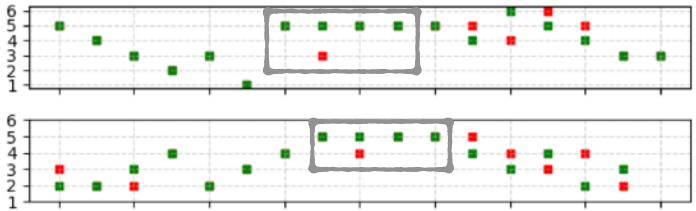




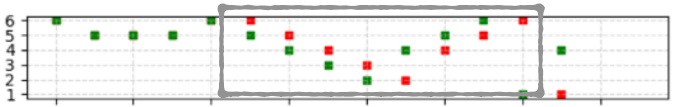
Experiment data

2.2 Error patterns

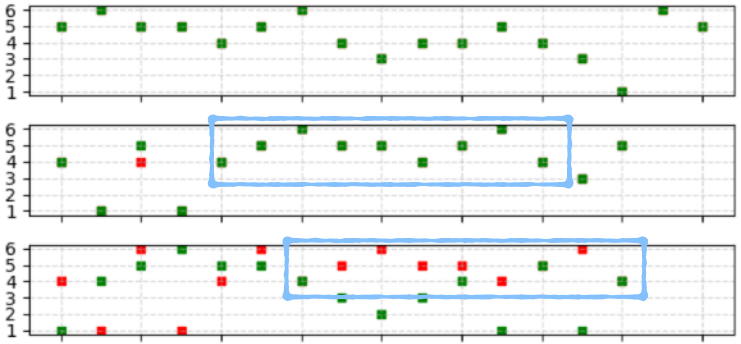
- Vertical shift



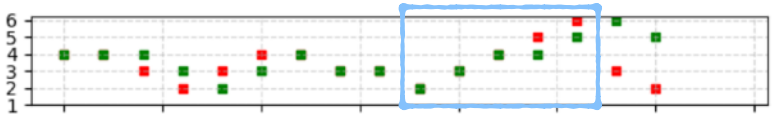
- Temporal shift



- Repeating



- Simple program



ground-truth

error

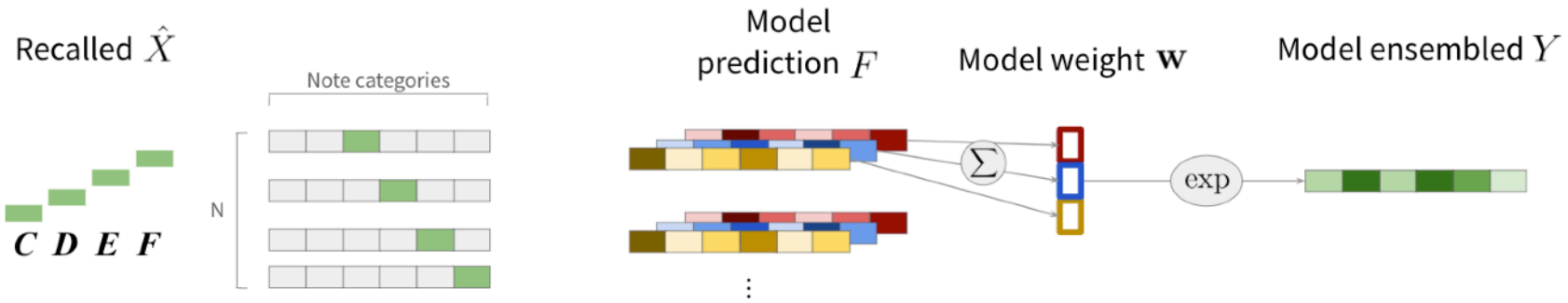
Experiment data

2.2 Recalled note sequence (human) vs. prediction (model)

Probability of pressing a key is based on a weighted sum of features:

$$\text{Choice}(x) \propto \exp(F \cdot \mathbf{w})$$

- F = feature matrix (what the model predicts)
- $\mathbf{w}$ = weights (learned from data)



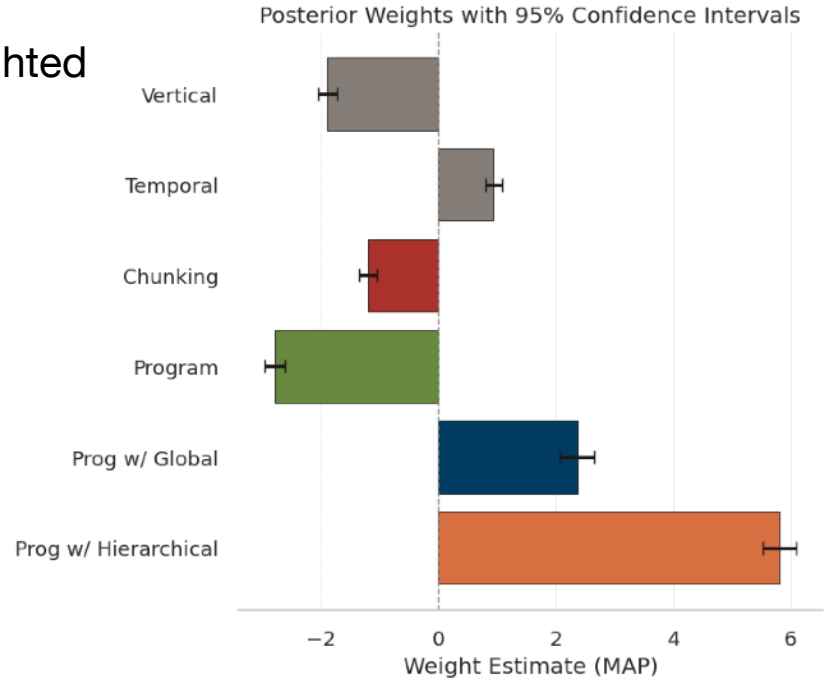
Experiment data

2.2 Recalled note sequence (human) vs. prediction (model)

Probability of pressing a key is based on a weighted sum of features:

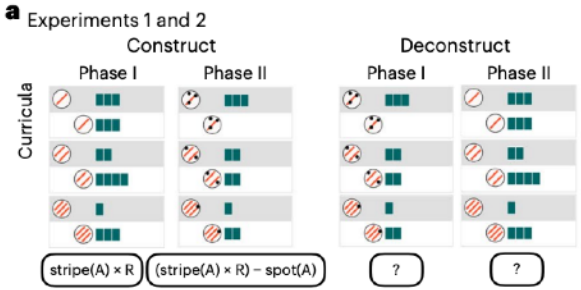
$$\text{Choice}(x) \propto \exp(F \cdot \mathbf{w})$$

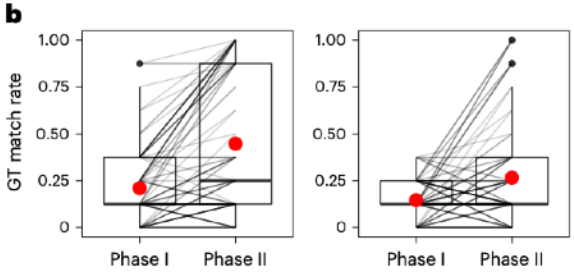
- F = feature matrix (what the model predicts)
- $\mathbf{w}$ = weights (learned from data)

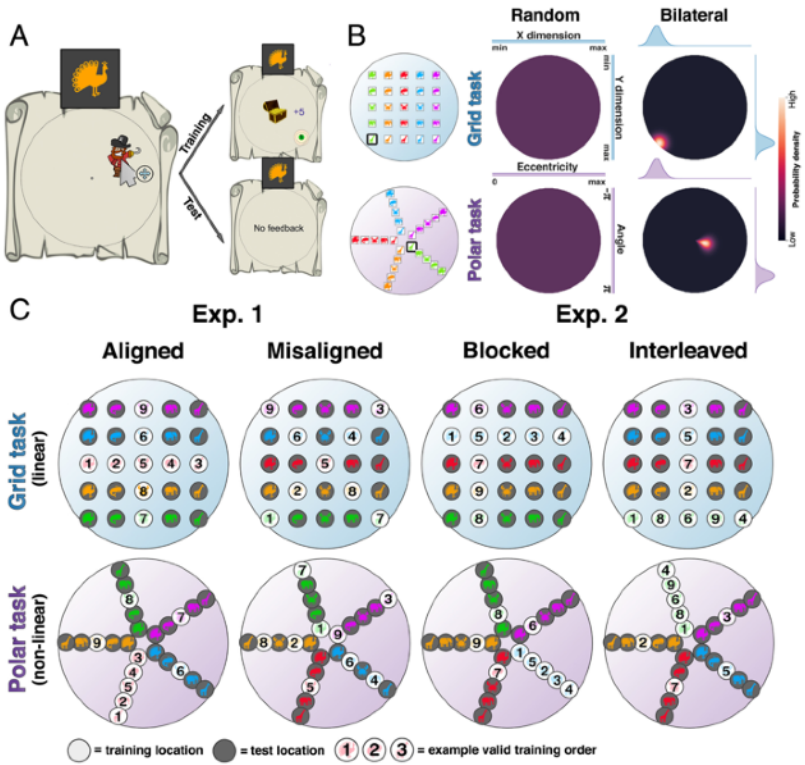


Simulation on curriculum

Curriculum in existing works







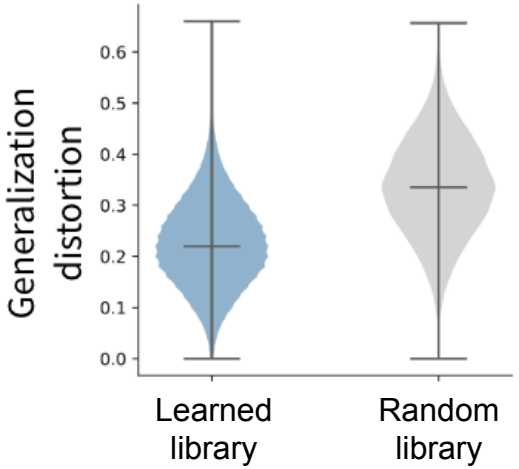
Zhao, B., Lucas, C. G., & Bramley, N. R. (2024). A model of conceptual bootstrapping in human cognition. Nature Human Behaviour, 8(1), 125-136.
Dekker, R. B., Otto, F., & Summerfield, C. (2022). Curriculum learning for human compositional generalization. Proceedings of the National Academy of Sciences, 119(41), e2205582119.

Simulation on curriculum

Sensitivity to curriculum

50 melodies under 100 different curricula:

- Only ~3.19% of programs are the same under different curricula
- This cannot be accounted for by stochasticity of training alone
- Humans also show path-dependence

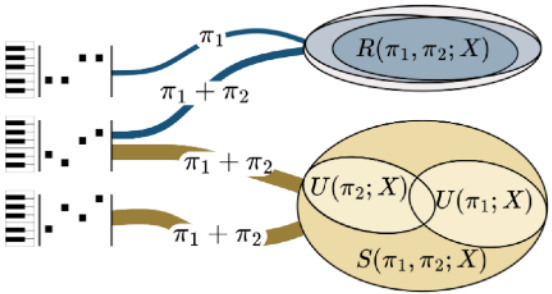


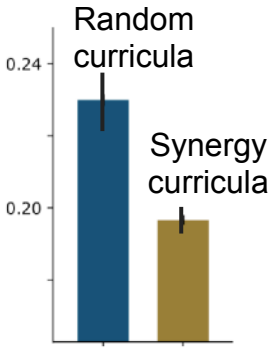
Simulation on curriculum

What is a good library?

A synergistic curriculum building method using principles of partial information decomposition

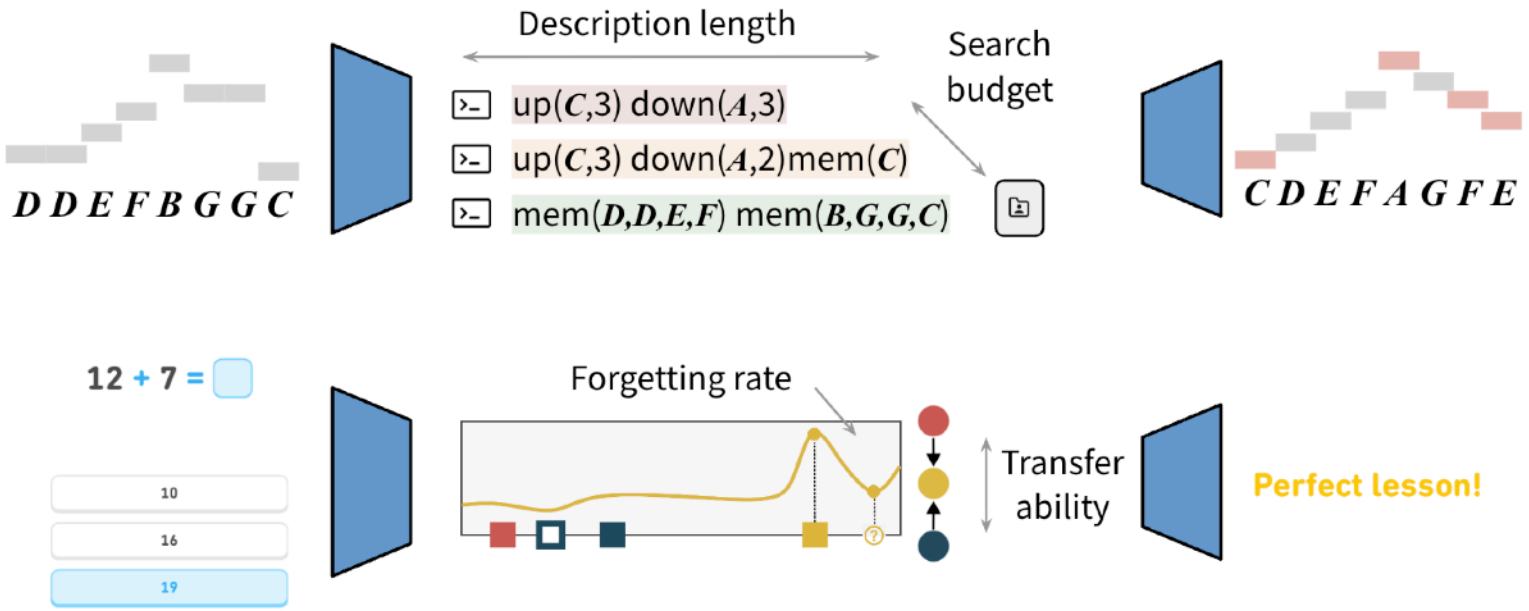
$$S(\mathscr{L}; X) = I\left(\pi_1, \pi_2; X\right) - R\left(\pi_1, \pi_2; X\right) - U\left(\pi_1; X\right) - U\left(\pi_2; X\right)$$





Proca, A. M., Rosas, F. E., Luppi, A. I., Bor, D., Crosby, M., & Mediano, P. A. (2024). Synergistic information supports modality integration and flexible learning in neural networks solving multiple tasks. PLOS Computational Biology, 20(6), e1012178.

Knowledge tracing in mathematics

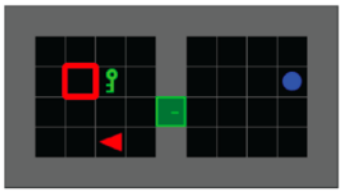


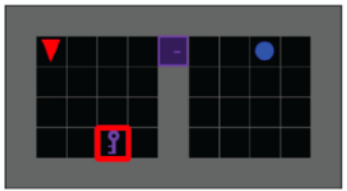
Zhou, H., Nagy, D. G., & Wu, C. M. (2024). Harmonizing program induction with rate-distortion theory. arXiv preprint arXiv:2405.05294.
Zhou, H., Bamler, R., Wu, C. M., & Tejero-Cantero, Á. (2024). Predictive, scalable and interpretable knowledge tracing on structured domains. arXiv preprint arXiv:2403.13179.

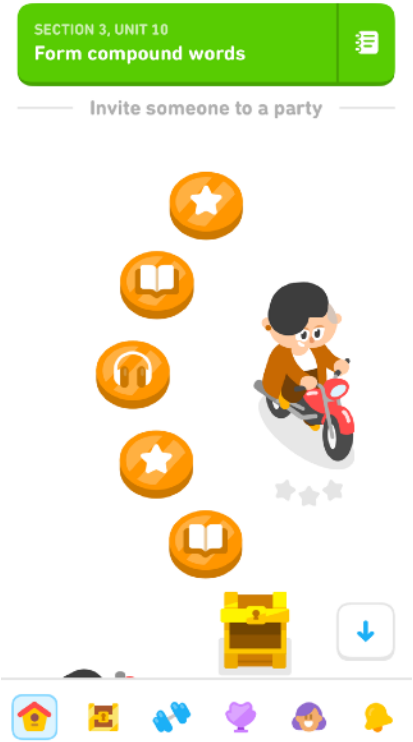
Explanations beyond curriculum

Curriculum in existing works

- Learning in the right order
- Given the target concept, learning with the right material







(Right) Figure source: (my own) Duolingo screenshot

(Left) Campero, A., Raileanu, R., Küttler, H., Tenenbaum, J. B., Rocktäschel, T., & Grefenstette, E. (2020). Learning with amigo: Adversarially motivated intrinsic goals. arXiv preprint arXiv:2006.12122.

Explanations beyond curriculum

Explanations in existing works

- What?

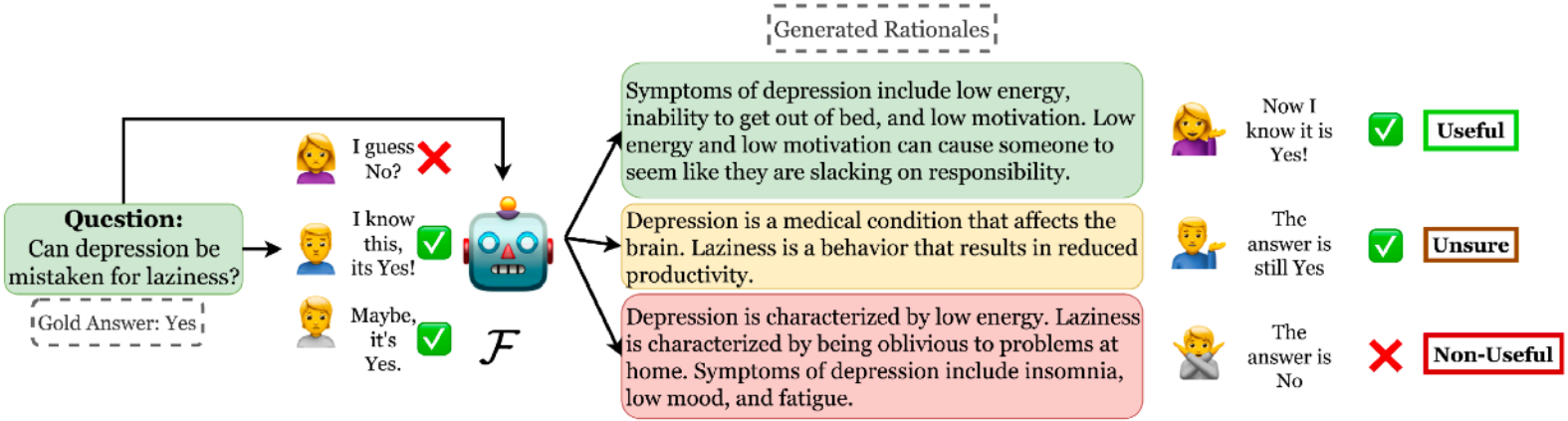


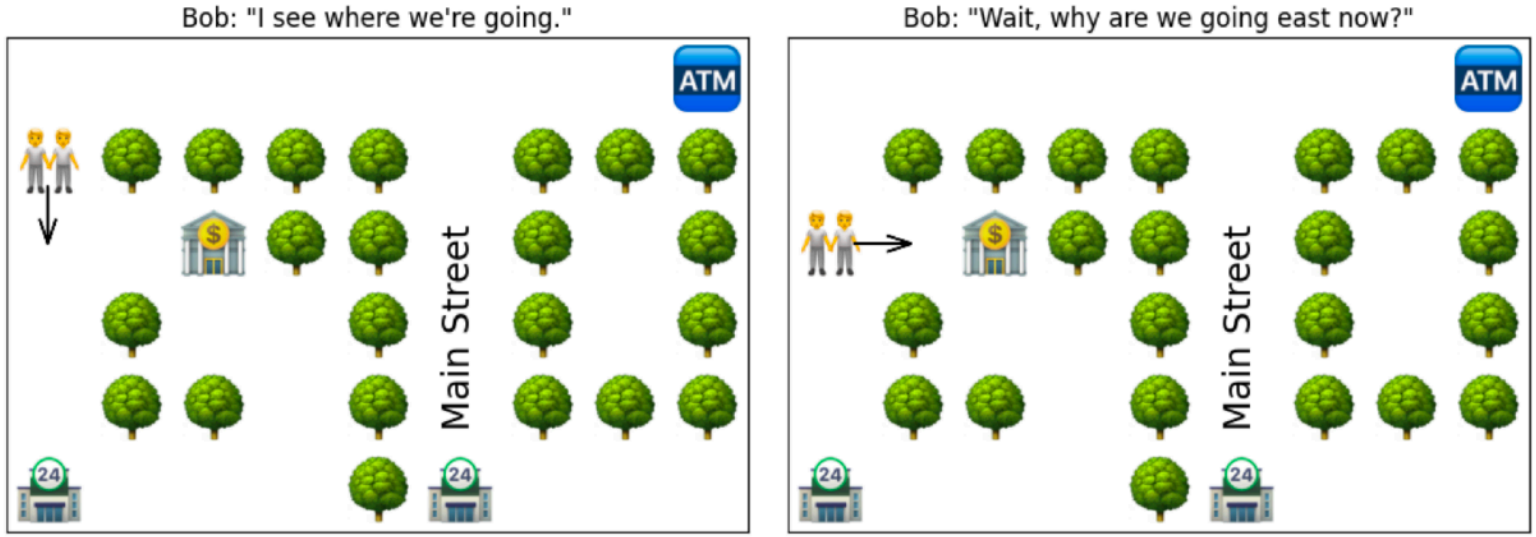
Figure 2: **An illustration of measuring human utility of machine rationales.** We evaluate whether a human's belief of the answer changes before and after seeing a rationale generated by an LM.

Joshi, B., Liu, Z., Ramnath, S., Chan, A., Tong, Z., Nie, S., ... & Ren, X. (2023). Are machine rationales (not) useful to humans? measuring and improving human utility of free-text rationales. arXiv preprint arXiv:2305.07095.

Explanations beyond curriculum

Explanations in existing works

- What & Why?

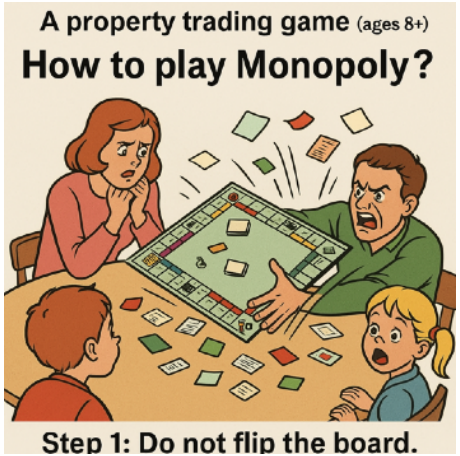


Chandra, K., Chen, T., Li, T. M., Ragan-Kelley, J., & Tenenbaum, J. (2024). Cooperative explanation as rational communication. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 46).

Explanations beyond curriculum

How…

- How to do it?
- How’s my performance?
- How do I improve on it?
- How do you know all of this?
- …



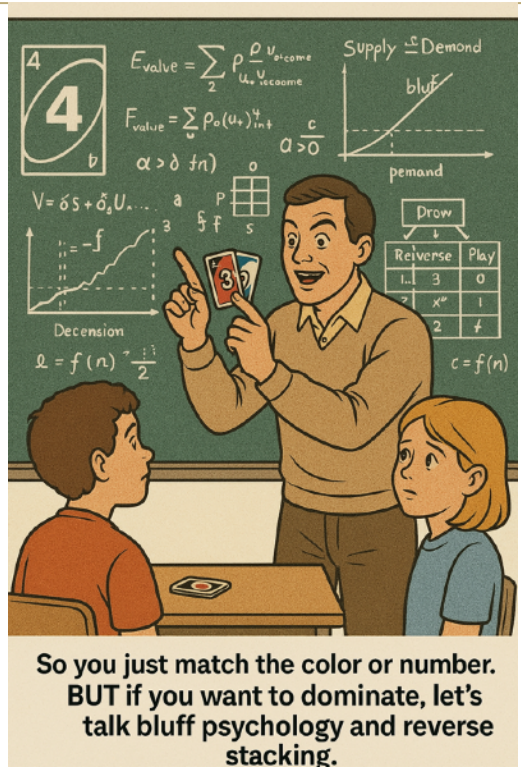


Figure source: GPT-4o (prompts are my own :)

What makes an explanation good, in context?

Generate and evaluate explanations that are tuned to

- the learner's current knowledge and abilities, and
- pedagogical goals (engagement, curiosity, and learning transfer).

The big question is:

- What makes an explanation good, in context?
- And how can we design tools (LLMs) that don't just explain well, but explain like a great teacher would?

Problem setting

Has a task $T^{(n)}$ with a reward function R

- **Knowledge**
  - Knows the task (and even the distribution $T \sim \mathcal{T}$)
  - A (more "complete") world model w_T
  - Learner's past experiences $T^{(1:n-1)}$
- **Goal**
  - Extrinsic: help the learner finish the task with explanation e
  - Intrinsic: pedagogy + unknown

- **Knowledge**
  - A (limited) world model w_L
- **Goal**
  - Extrinsic: finish the task and get the reward
  - Intrinsic: unknown

Problem setting

Has a task $T^{(n)}$ with a reward function R

- What can be explained?
  - Policy π: here's exactly what you should do next.
  - World model w: here's how the terrain works -- pits block you, keys open gates.
  - Value function v: it's important to explore early, worth taking a detour for long-term gain.
- How much should be explained?
  - Memory load
  - Inference load

Problem setting

Has a task $T^{(n)}$ with a reward function R

- Inference

$$w^{(n)}, \pi^{(n)} \sim P_L(w, \pi \mid e, w_L^{(n-1)})$$
$$= P_L(\pi \mid w, e) P_L(w \mid e, w_L^{(n-1)})$$

- Utility function

$$U_L(e; T^{(n)}) = R(\pi^{(n)}) + R_L \quad \text{Reward}$$
$$- C(P_L) \quad \text{Learning cost}$$

	$e = \pi_T^{(n)}$	$e = w_T$	$e = v_T$
Compute	L	M	H
Memory	H	M	L

Problem setting

Has a task $T^{(n)}$ with a reward function R

- Inference

$$\hat{w}_L^{(n-1)} \sim P_T(w \mid T^{(1:n-1)})$$

$$\hat{w}_L^{(n)}, \hat{\pi}_L^{(n)} \sim P_T\left(w, \pi \mid e, \hat{w}_L^{(n-1)}\right)$$

- Utility function

$$\begin{aligned} U_T(e; T^{(n)}) = R(\hat{\pi}^{(n)}) - C_T(P_L) & \quad \text{Learn's performance} \\ -C_T(e) & \quad \text{Communication cost} \\ +\gamma \cdot \mathbb{E}_{T^{(n+1)},\pi'}\left[R_{T^{(n+1)}}(\pi')\right] - C_T(P_L) & \quad \text{Long-term pedagogy reward} \\ -C_T(P_T) & \quad \text{Inference cost} \end{aligned}$$

- Inference

$$\begin{aligned} w^{(n)}, \pi^{(n)} &\sim P_L(w, \pi \mid e, w_L^{(n-1)}) \\ &= P_L(\pi \mid w, e) P_L(w \mid e, w_L^{(n-1)}) \end{aligned}$$

- Utility function

$$\begin{aligned} U_L(e; T^{(n)}) = R(\pi^{(n)}) + R_L \\ -C(P_L) \end{aligned}$$

Problem setting

How to play the game?

How to get to the trader?

Walk

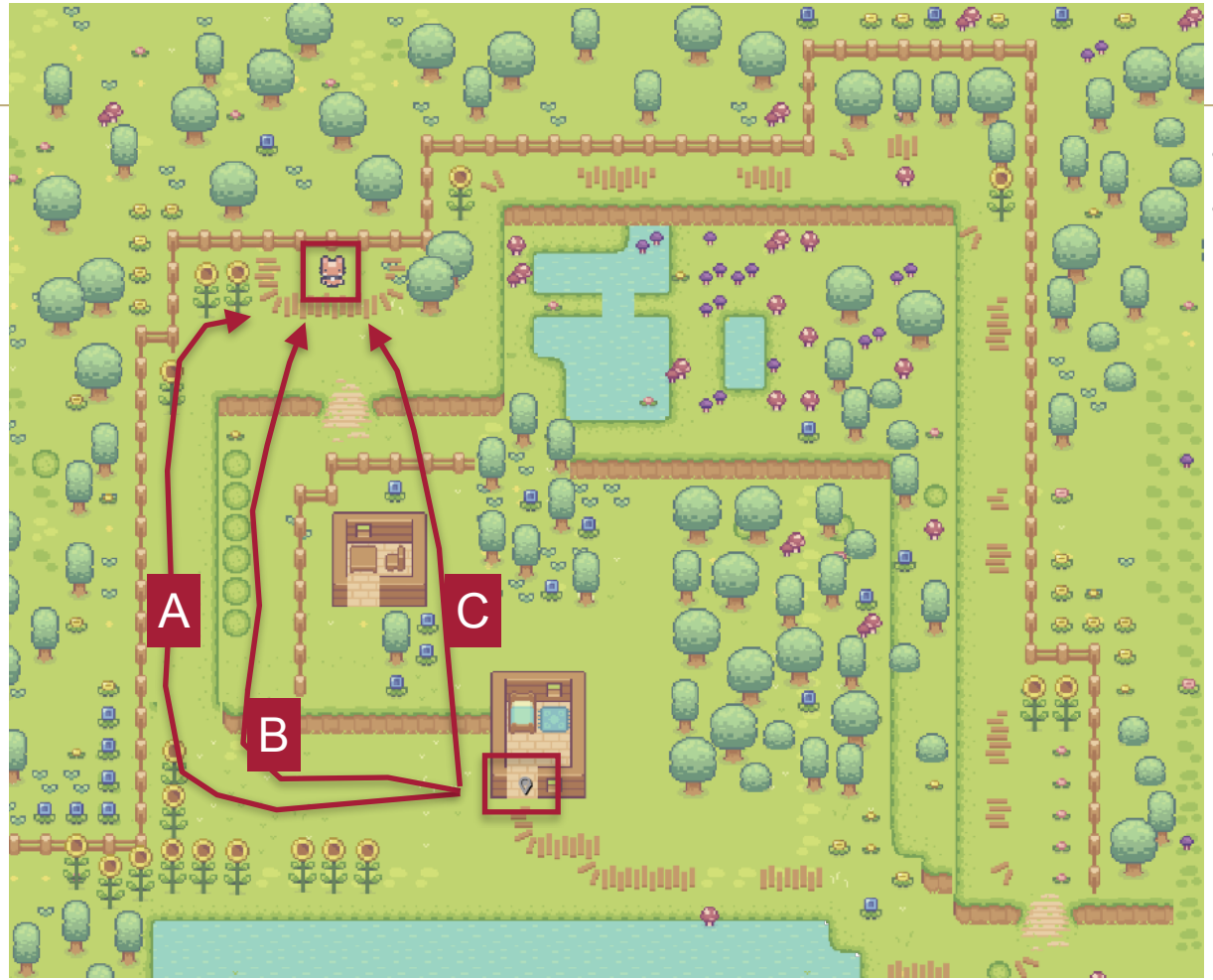
Locate

Clamber

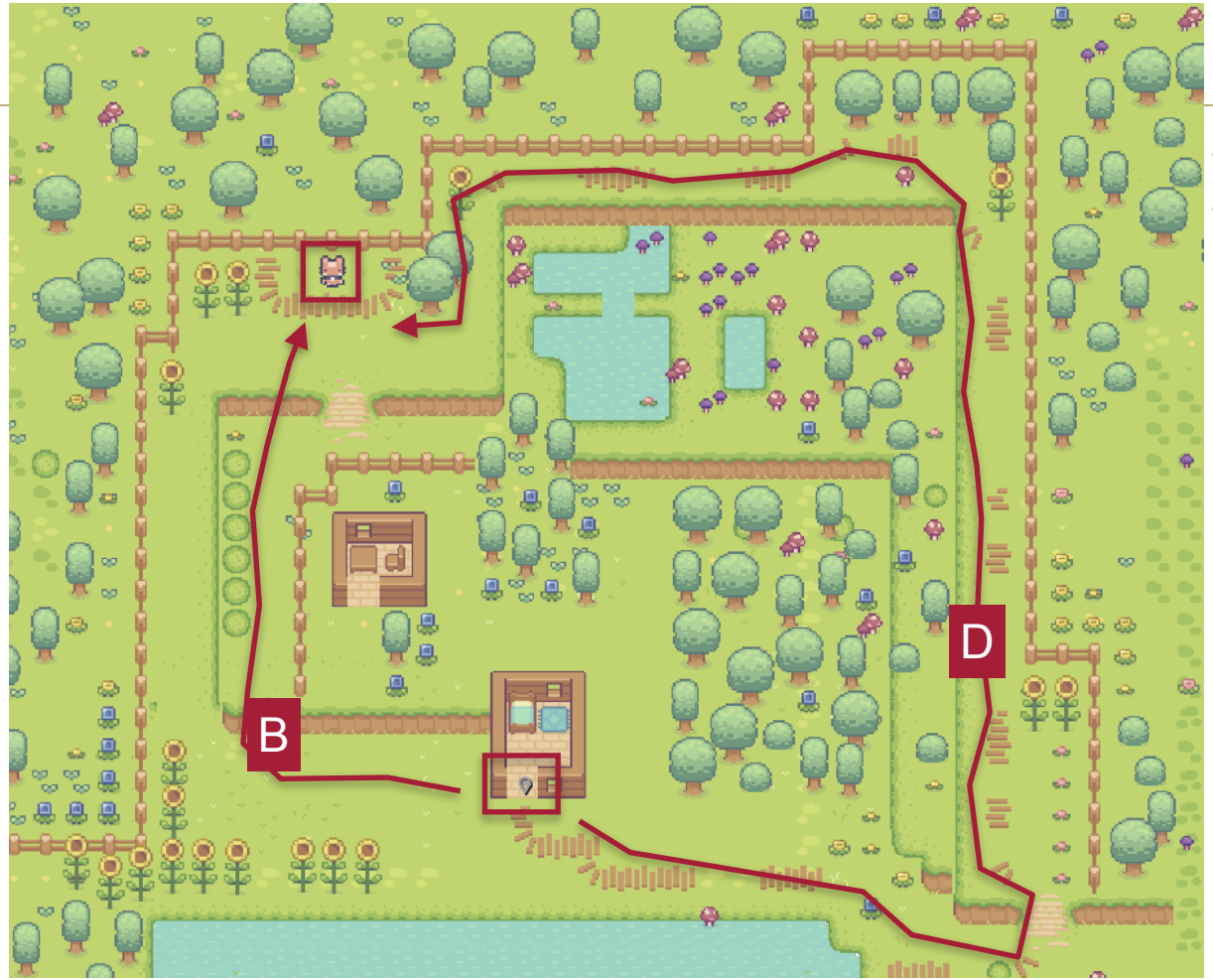
Vault

Hop

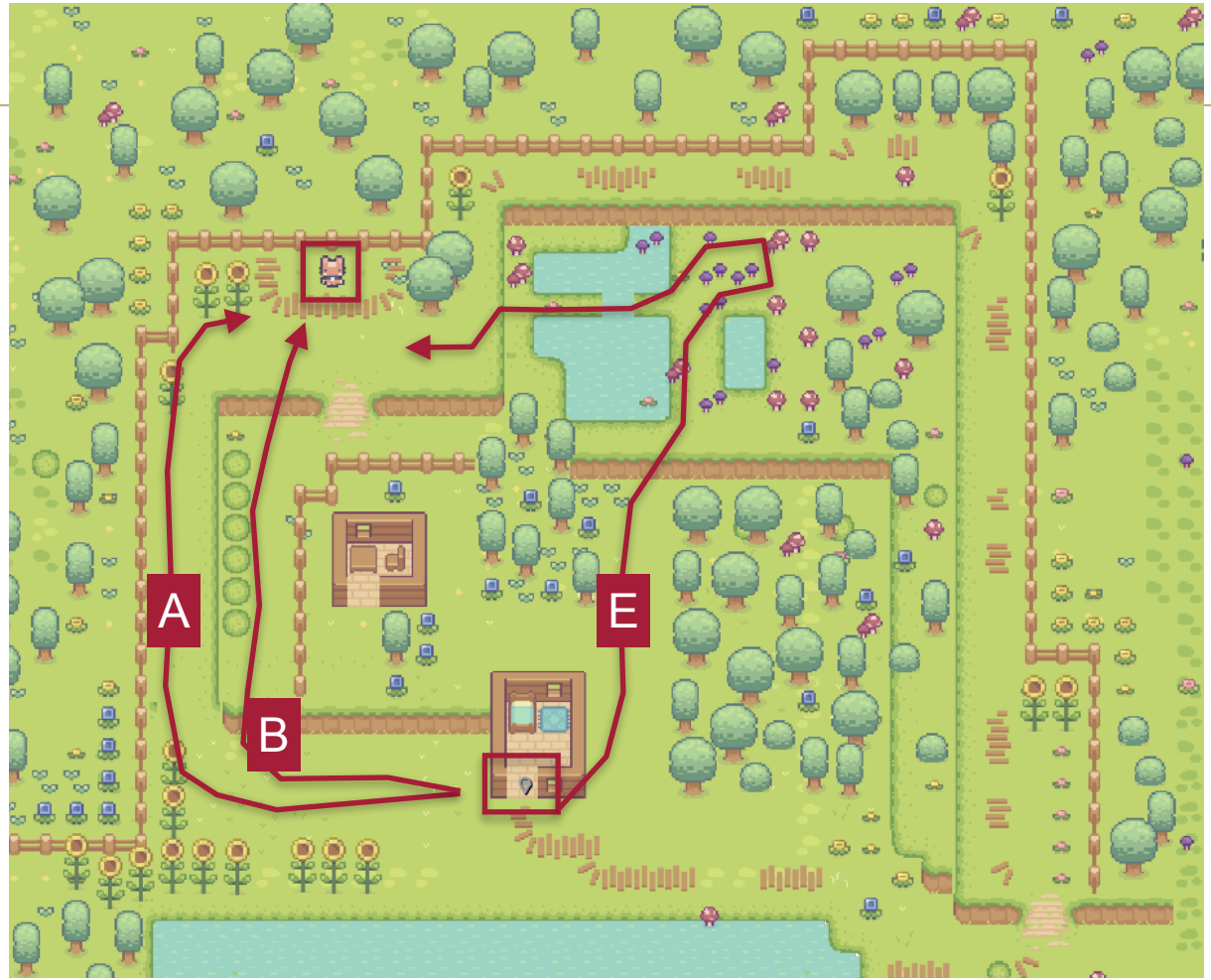
Forage



- Sensitive to learner's background
- A to B to C
  - Learner needs less effort (more reward)
  - Teacher needs more inference cost
  - Teacher needs more communication cost



- Sensitive to computation cost
- B and D:
  - Learner needs less effort on B (more reward)
  - Teacher needs more inference cost on B
  - Teacher needs more communication cost on B



- Sensitive to long-term goal
- A to B to E:
  - Learner needs more effort (more reward in the long run)
  - Teacher needs more inference cost
  - Teacher needs more communication cost

Thank you for listening!

Question?