Digital Intelligence vs Biological Intelligence

Ying Nian Wu, UCLA

On the occasion of Prof. Alan Yuille's 70th birthday

Moving Beyond Past Debates

- We won't focus on past debates (still meaningful and important though):
  - How the brain implements backpropagation
  - Spiking networks vs. artificial neurons
  - Neural symbolic
  - System 2 vs system 1
  - Scaling laws
- Instead, we'll examine a fundamental distinction:

  Learning for Reaction vs. **Learning for Planning**
- This perspective illuminates core differences between digital and biological intelligence

Two Learning Paradigms

Learning for Reaction

- Maps inputs directly to outputs/actions
- Relies on extensive training examples
- Essentially interpolative look-up table
- Muscle memory, reflex
- Struggles with novel scenarios
- Performance degrades when conditions differ from training

Learning for Planning

- Builds representations that support planning and reasoning
- Creates cognitive maps that can be traversed mentally
- Enables navigation through never-explored state spaces
- Adapts to new goals without retraining
- Generalizes robustly to novel scenarios

Hippocampal Place Cells: A Concrete Example

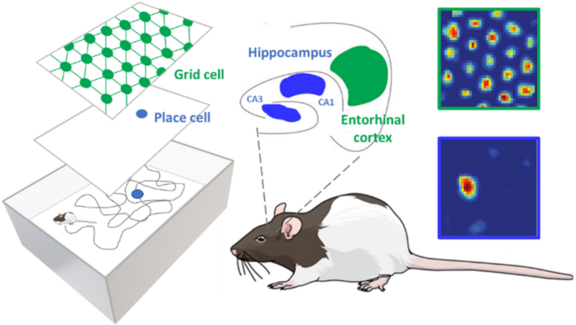


Figure: Place cells and grid cells (from internet).

Hippocampal Place Cells: A Concrete Example

- Place cells in the hippocampus fire at specific locations as animals navigate
- Traditional view: Individual cells encode specific locations
- Our approach: Population of place cells collectively encode transition probabilities
- We reconceptualize place cells as position embeddings of multi-time random walk kernels

Key mathematical formulation:

$$\langle h(x,t), h(y,t)\rangle = p(y|x,t) \tag{1}$$

where:

- $h(x,t) \in \mathbb{R}^n$ is the embedding at location x and scale t, $h_i(x,t) \geq 0$.
- $p(y|x,t)$ is the transition probability.
- Symmetric random walk $\rightarrow$ heat equation with reflective boundary condition
- $\sqrt{t}$ defines a spatial scale (like dorsoventral axis)
- Exploration or mapping policy, not a navigation policy

Matrix Squaring: From Local to Global

- We compute multi-time transition probabilities efficiently:

$$P_{2t} = P_t^2 \tag{2}$$

- This matrix squaring process:
  - Requires only local exploration (P_1) to build global knowledge
  - Needs no successful trajectories for learning
  - Implicitly encodes the recursions in dynamic programming
  - Enables hippocumpal preplay
- Learning reduces pairwise adjacency relationships to individual embeddings (**map**):

$$\mathcal{L} = \sum_{x,y} [p(y|x,t) - \langle h(x,t), h(y,t) \rangle]^2 \tag{3}$$

Path Planning via Adaptive Gradient Following

- Planning becomes straightforward gradient following

$$\nabla_x p(y|x,t) = \nabla_x \langle h(x,t), h(y,t) \rangle = -\nabla_x \| h(x,t) - h(y,t) \|^2 \tag{4}$$

- Adaptive scale selection for optimal guidance

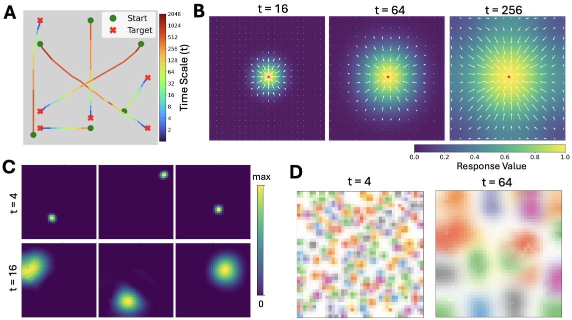
$$t^* = \arg\max_{t \in \mathcal{T}} \| \nabla_x \langle h(x,t), h(y,t) \rangle \| \tag{5}$$

$$= \arg\max_{t \in \mathcal{T}} \| \nabla_x \| h(x,t) - h(y,t) \|^2 \| \tag{6}$$

- Planning is "straightforward" (following "straight" path in embedding space).

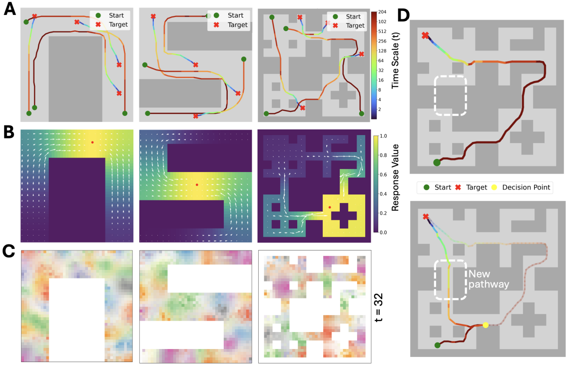
Multi-Scale Representation

- Time parameter $\sqrt{t}$ defines spatial scale
- Resembles dorsoventral organization in hippocampus
- Small t: geodesic distance
- Large t: topological connectivity
- Adaptive selection during navigation
- Non-negative $\rightarrow$ symbols-like



- Adaptive scale selection (colors) dynamically adjusts based on distance to goal
- Larger scales (red/orange) for distant planning, smaller scales (blue/green) for precision

Navigation in Complex Environments

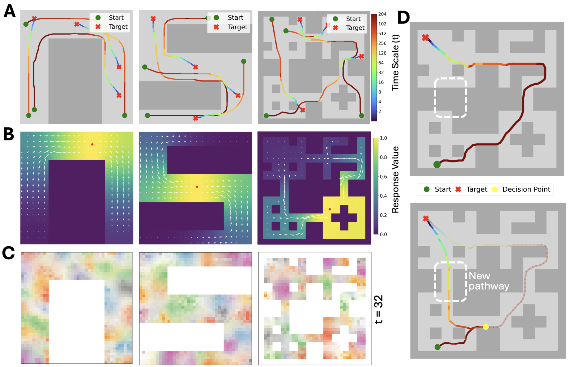


- Our model achieves 100% success rate in complex mazes
- Gradient fields naturally create diffraction-like patterns around obstacles
- Smooth, continuous paths through complex environments

Learning for Planning vs. Learning for Reaction

- **Key insight**: Learning for planning requires no successful trajectories at all
  - Random walks suffice to build planning representations
  - No goal-directed behavior needed during training
  - No reward functions or reward shaping required
- **Planning-centric learning advantages**:
  - Zero-shot adaptation to new goals
  - Local-to-global emergence through matrix squaring
  - Simple and Efficient planning through gradient following
  - Robust to environmental changes with fine-tuning

Shortcut Discovery Experiment



- After just 50 iterations of fine-tuning:
  - Model successfully identifies and utilizes the newly available shortcuts
  - Adapts planning based on new transition probabilities
  - Learning for reaction would require extensive retraining

Learning for Reaction

Evaluating the World Model Implicit in a Generative Model

Keyon Vafa
Harvard University

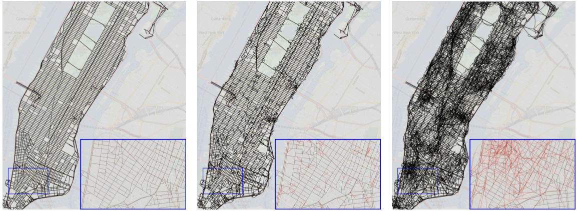
Justin Y. Chen
MIT

Ashesh Rambachan
MIT

Jon Kleinberg
Cornell University

Sendhil Mullainathan
MIT

We randomly split data into train and test splits, ensuring no origin-destination pair is in both train and test sets. We include all sequences containing less than 100 directions. Our training sets consist of 2.9M sequences (120M tokens) for shortest paths; 31M sequences (1.7B tokens) for noisy shortest paths; and 91M sequences (4.7B tokens) for random walks. We train two types of transformers [38] from scratch using next-token prediction for each dataset: an 89.3M parameter model consisting of 12 layers, 768 hidden dimensions, and 12 heads; and a 1.5B parameter model consisting of 48 layers, 1600 hidden dimensions, and 25 heads. We follow the architecture of GPT-2 for each model [29]. We train models on



(a) World model **(b)** World model with noise **(c)** Transformer

Why Digital Intelligence is Successful

- Not just about scaling laws or computational resources
- **The Power of the Residual Stream**:
  - Creates an implicit iterative algorithm
  - No explicit objective function or gradient calculation
  - Yet produces sophisticated multi-step computation
- This implicit algorithm may:
  - Replicate planning-like or reasoning-like processes
  - Discover computational shortcuts
  - Make "muscle memory" surprisingly comparable to planning
- But fundamental limitations remain:
  - Struggles with environmental changes
  - Limited generalization to out-of-distribution scenarios
  - Requires extensive training examples

Biological vs. Digital Intelligence

Biological Intelligence	**Digital Intelligence**
Learning for planning	Learning for reaction
Explicit planning through representations	Implicit algorithm through residual stream
Efficient from limited data	Data-hungry
Zero-shot adaptation to new environments	Requires extensive training examples
Robust to environmental changes	Brittle to distribution shifts
Balances exploration and exploitation	Often poor at exploration

"Understanding": enabling planning or reasoning with pre-defined algorithm

Digitalize biological intelligence

Latent Thought Language Models

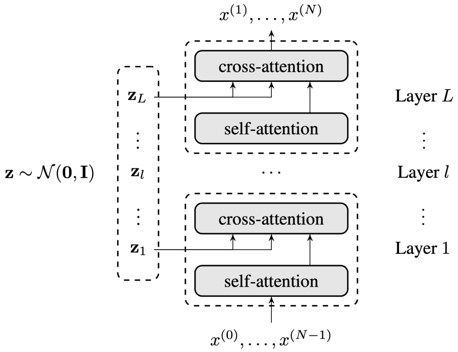


Figure: LTM architecture.

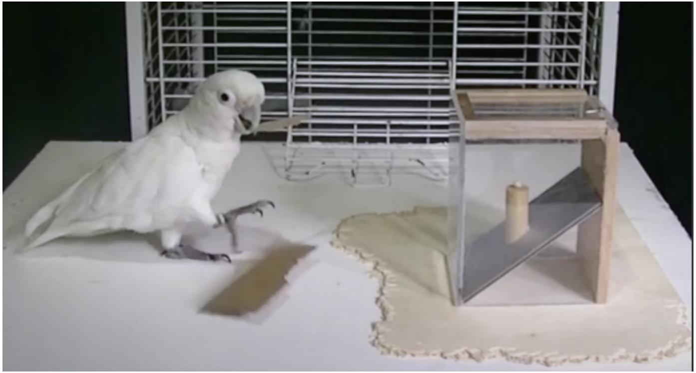
Reasoning = optimization in latent space

Latent Plan Models

Figure: Box catching by human and robot.

Planning = optimization in latent space

Stochastic Parrot?



Mimicking human speech: learning for reaction.
Finding food: learning for planning

- Zhao, M*; Xu, D*; Kong, Deqian*; Zhang, W; Wu, Y N, Place Cells as Position Embeddings of Multi-Time Random Walk Transition Kernels for Path Planning, 2025.
- Kong, Deqian*; Zhao, M*; Xu, D*; Pang, B; Wang, S; Honig, E; Si, Z; Li, C; Xie, J[†]; Xie, S[†]; Wu, Y N[†], Latent Thought Models with Variational Bayes Inference-Time Computation. ICML 2025.
- Noh, D*; Kong, Deqian*; Zhao, M; Lizzaraga, A.; Xie, J, Wu, Y N[†], Hong, D[†], Latent Adaptive Planner for Dynamic Manipulation, 2025.