# Weak Inductive Biases for Discovering Composable Primitive Representations

**Wilka Carvalho**
University of Michigan

I argue that the ability to discover composable primitive representations will play a key role in enabling deep RL agents that can as our helpful companions in society. In particular, I argue that the ability to discover composable primitive representations will facilitate generalization and goal-completion for time-horizons spanning months or years. As a concrete example, consider a deep RL variant of a digital assistant (e.g., Apple's Siri or Samsung's Bixby) tasked with improving a user's health. The agent observes the user's data (e.g. heart rate, geographical data, etc.) and will need to learn a representation of user-state that enables a policy that predicts *personalized* user-recommendations (e.g. how to excercise, when to sleep, etc.). I argue that here (and in many real-world settings), an effective policy requires a state representation that spans and effectively summarizes long time-horizons over potentially months of data. This is challenging because real-world data is complex, continuously changing, and partially observable. Additionally, in order achieve a health goal, the agent will need to successfully predict actions over a time-horizon spanning months or years. Finally, this agent has the potential to learn from and predict actions for millions of users. This will require that it learn behavior that it can generalize to new users and new user-conditions.

My argument is that this type of long-horizon policy learning can be supported by the ability to discover composable primitive representations. Returning to our example, the agent benefits from discovering "activity-primitives" that summarize a user's exercise and body-recovery periods (e.g. the period around "walking"). Let's assume the agent has a mechanism for discovering compositional *sequence-primitive* representations (such as how convolutional filters can be composed hierarchically). This allows the agent to represent user-state with a sequence-primitive defined over (potentially months of) activity-primitives. This would facilitate modelling the user's fitness level and providing personalized exercise recommendations. Additionally, the agent can use the space of activity-sequences to define goals for activity-recommendation options. Activity-recommendation options may facilitate planning an exercise regimen that will help the user achieve their health goal over months or years. As the agent experiences more users, it can discover which activity-recommendation options are more effective for health-goals given starting user-states.

**In my view, weak inductive biases are the key to algorithms that can discover composable primitives.** CNNs and Transformers are two exciting examples showcasing the utility of discovering composable primitives during state-representation learning. With Transformers, one can learn hierarchically composable attention primitives. With CNNs, one can learn hierarchically composable spatial feature primitives. CNNs are a de facto modelling choice when learning state representations over images, and Transformers are beginning to see utility for learning state representations over longer time-horizons. Neither one of these methods was developed by or for deep RL. Instead, they were developed in other fields and applied to deep RL. **As a field, we should place a larger emphasis on developing weak inductive biases designed specifically for reinforcement learning.**

The ability of CNNs and Transformers to capture a broad range of composable primitives has become more apparent as the two have have been trained on massive amounts of image and language data. **Perhaps, this is what deep RL is missing: a large, rich environment where an agent has the ability to learn from massive amounts of diverse experiences in the world.** Large, rich environments might provide the right breeding ground for developing weak inductive biases focused on enabling agents to discover reinforcement-learning oriented composable primitives. For example, large, rich environments may enable agents that discover composable primitives for state-representation learning such as state abstractions or for policy learning such as options.

Preprint. Under review.