Concept Learning with Energy-Based Models

Igor Mordatch (mordatch@openai.com)
OpenAI, San Francisco, CA

Keywords: concepts;meta-learning;imitation learning

Introduction

Many hallmarks of human intelligence, such as generalizing from limited experience, abstract reasoning and planning, analogical reasoning, creative problem solving, and capacity for language require the ability to consolidate experience into *concepts*, which act as basic building blocks of understanding and reasoning.

Examples of concepts include visual (*"red"* or *"square"*), spatial (*"inside"*, *"on top of"*), temporal (*"slow"*, *"after"*), social (*"aggressive"*, *"helpful"*) among many others (Rosch, Mervis, Gray, Johnson, & Boyes-braem, 1976; Lakoff & Johnson, 1980). These concepts can be either identified or generated - one can not only find a square in the scene, but also create a square, either physical or imaginary. Importantly, humans also have a largely unique ability to combine concepts compositionally (*"red square"*) and recursively (*"move inside moving square"*) - abilities reflected in the human language. This allows expressing an exponentially large number of concepts, and acquisition of new concepts in terms of others. We believe the operations of identification, generation, composition over concepts are the tools with which intelligent agents can understand and communicate existing experiences and reason about new ones.

Crucially, these operations must be performed on the fly throughout the agent's execution, rather than merely being a static product of an offline training process. Execution-time optimization, as in recent work on meta-learning (Finn, Abbeel, & Levine, 2017) plays a key role in this. We pose the problem of parsing experiences into an arrangement of concepts as well as the problems of identifying and generating concepts as optimizations performed during execution lifetime of the agent. The meta-level training is performed by taking into account such processes in the inner level.

Specifically, a concept in our work is defined by an energy function taking as input an event configuration (represented as trajectories of entities in the current work), as well as an attention mask over entities in the event. Zero-energy event and attention configurations imply that event entities selected by the attention mask satisfy the concept. Compositions of concepts can then be created by simply summing energies of constituent concepts. Given a particular event, optimization can be used to identify entities belonging to a concept by solving for attention mask that leads to zero-energy configuration. Similarly, an example of a concept can be generated by optimizing for a zero-energy event configuration. See Figure 1 for examples of these two processes.

The energy function defines a family of concepts, from which a particular concept is selected with a specific concept code. Encoding of event and attention configurations can be achieved by execution-time optimization over concept codes. Once an event is encoded, the resulting concept code structure can be used to re-enact the event under different initial configurations (task of imitation learning), recognize similar events, or concisely communicate the nature of the event. We believe there is a strong link between concept codes and language, but leave it unexplored in this work.

Description of events we consider and video results of our model learning on these events are available at: `sites.google.com/site/energyconceptmodels`

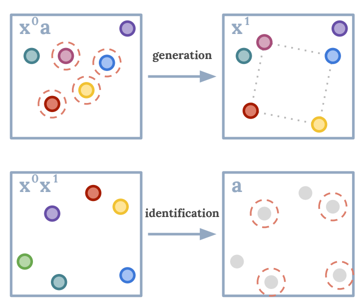

Figure 1: Examples of generation and identification processes for a *"square"* concept. a) Given initial state $\mathbf{x}^0$ and attention mask $\mathbf{a}$, square consisting of entities in $\mathbf{a}$ is formed via optimization over $\mathbf{x}^1$. b) Given states $\mathbf{x}$, entities comprising a square are found by optimization over attention mask $\mathbf{a}$.

Method

Existence of a particular concept is given by energy function $E(\mathbf{x},\mathbf{a},\mathbf{w}) \in \mathbb{R}^+$, where parameter vector $\mathbf{w}$ specifies a particular concept from a family. $E(\mathbf{x},\mathbf{a},\mathbf{w}) = 0$ when state trajectory $\mathbf{x}$ under attention mask $\mathbf{a}$ over entities satisfies the concept $\mathbf{w}$. Otherwise, $E(\mathbf{x},\mathbf{a},\mathbf{w}) > 0$. The conditional probabilities of a particular event configuration belonging to a concept and a particular attention mask identifying a concept are given by the Boltzmann distributions:

$$p(\mathbf{x}|\mathbf{a},\mathbf{w}) \propto \exp\{-E(\mathbf{x},\mathbf{a},\mathbf{w})\} \tag{1}$$

$$p(\mathbf{a}|\mathbf{x},\mathbf{w}) \propto \exp\{-E(\mathbf{x},\mathbf{a},\mathbf{w})\} \tag{2}$$

Given concept code $\mathbf{w}$, the energy function can be used for both generation and identification of a concept implicitly via optimization (see Figure 1):

$$\mathbf{x}(\mathbf{a}) = \operatorname*{argmin}_{\mathbf{x}} E(\mathbf{x},\mathbf{a},\mathbf{w}) \quad \mathbf{a}(\mathbf{x}) = \operatorname*{argmin}_{\mathbf{a}} E(\mathbf{x},\mathbf{a},\mathbf{w}) \tag{3}$$

To learn concepts from experience grounded in events, we pose a few-shot prediction task. Given a few demonstration example events and initial state for a new event, the task is to predict attention $\mathbf{a}$ and the future state $\mathbf{x}$ of the new event. We

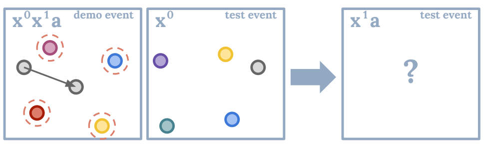

Figure 2: Example of a few-shot prediction task.

follow the maximum entropy inverse reinforcement learning formulation (Ziebart, Maas, Bagnell, & Dey, 2008) and assume demonstrations are samples from the distributions given by the energy function E and find energy function parameters θ via maximum likelihood estimation over future state and attention given initial state. The resulting loss for a particular dataset X is

$$\mathcal{L}_p^{\mathrm{ML}}(X,\mathbf{w}) = \mathbb{E}_{(\mathbf{x},\mathbf{a})\sim X}\left[-\log p\left(\mathbf{x}^1,\mathbf{a} \mid \mathbf{x}^0,\mathbf{w}\right)\right]$$

Where the joint probability can be decomposed in terms of probabilities in (1) and (2) as

$$\log p\left(\mathbf{x}^1,\mathbf{a} \mid \mathbf{x}^0,\mathbf{w}\right) = \log p\left(\mathbf{x}^1 \mid \mathbf{a},\mathbf{w}_x\right) + \log p\left(\mathbf{a} \mid \mathbf{x}^0,\mathbf{w}_a\right)$$

We use two concept codes, $\mathbf{w}_x$ and $\mathbf{w}_a$ to specify the joint probability. The interpretation is that $\mathbf{w}_x$ specifies the concept of the action that happens in the event (i.e. *"be in center of"*) while $\mathbf{w}_a$ specifies the argument the action happens over (i.e. *"square"*). This is a concept structure or syntax that describes the event. The concept codes are interchangeable and same concept code can be used either as action or as an argument because the energy function defining the concept can either be used for generation or identification. This importantly allows concepts to be understood from their usage under multiple contexts.

Experimental Results

We introduce a simulated environment and tasks for a two-dimensional scene consisting of a varying collection of entities, each processing position, color, and shape. We observe the following properties:

Concept inference in multiple contexts: An important property of our model is ability to learn from and apply it in both generation and identification contexts. We qualitatively observe that the model performs sensible behavior in both contexts. For example, we considered events with proximity relations *"closest"* and *"farthest"* and found model able to both attend to entities that are closest or furthest to another entity, and to move an entity to be closest or furthest to another entity.

Transfer between contexts: When our model trained on both contexts it shares experience between contexts. Knowing how to act out a concept should help in identifying it and vice versa. We perform an experiment where we train the energy model only in identification context and test the model's performance in generation context (and conversely). We observe that even without explicitly being trained on the appropriate context, the networks perform much better than baseline of two independently-trained networks.

Sharing codes across contexts: Another property of our model is that codes $\mathbf{w}_x$ and $\mathbf{w}_a$ for identifying concepts are interchangeable and can be shared between generation and identification contexts. For example, either turning an entity red would or identifying all red entities in the scene would ideally use the same concept of *"red"*. We indeed observe that events which involve recognizing entities of a particular color, the codes $\mathbf{w}_a$ match the codes $\mathbf{w}_x$ for setting entities to that color and find similar correlation in the other events as well.

Conclusion

We believe that execution-time optimization plays a crucial role in acquisition and generalization of knowledge, planning and abstract reasoning, and communication. In the current work we used a simple concept structure, but more complex structure with multiple arguments or recursion would be interesting to investigate in the future. It would also be interesting to test compositionality of concepts, which is very suited to our model as compositions corresponds to the summation of the constituent energy functions.

References

Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *arXiv preprint arXiv:1703.03400*.

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive science*, *4*(2), 195–208.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-braem, P. (1976). Basic objects in natural categories. *COGNITIVE PSYCHOLOGY*.

Ziebart, B. D., Maas, A. L., Bagnell, J. A., & Dey, A. K. (2008). Maximum entropy inverse reinforcement learning. In *Aaai* (Vol. 8, pp. 1433–1438).