Experiments on the Consciousness Prior

Yoshua Bengio and William Fedus UNIVERSITÉ DE MONTRÉAL, MILA

Abstract

Experiments are proposed to explore a novel prior for representation learning, which can be combined with other priors in order to help disentangling abstract factors from each other. It is inspired by the phenomenon of consciousness seen as the formation of a low-dimensional combination of a few concepts constituting a conscious thought, i.e., consciousness as awareness at a particular time instant. This provides a powerful constraint on the representation in that such low-dimensional thought vectors can correspond to statements about reality which are true, highly probable, or very useful for taking decisions. Instead of making predictions in the sensory (e.g. pixel) space, the consciousness prior allows the agent to make predictions in the abstract space, with only a few dimensions of that space being involved in each of these predictions. Experiments on a synthetic dataset are proposed to validate some of the mechanisms proposed to implement the consciousness prior, in the simplest scenario where the consciousness mechanism is only used to make a prediction.

1. Introduction

This project explores the proposal for a new kind of prior for top-level abstract representations, inspired by our understanding of consciousness as a form of awareness (van Gulick, 2004), i.e., as defined by Locke, consciousness is "the perception of what passes in a man's own mind", or awareness of an external object or something within oneself (Wikipedia definition). This proposal is based on a regularization mechanism which encourages the top-level representation (meant to be at the most abstract level) to be such that when a sparse attention mechanism focuses on a few elements of the state representation (factors, variables or concepts, i.e. a few axes or dimensions in the representation space), that small set of variables of which the agent is aware at a given moment can be combined to make a useful statement about reality or usefully condition an action or policy.

2. Consciousness Prior Theory

See Bengio (2017) for a longer description of the conciousness prior. The following points can be derived from the basic idea introduced above, in the context of a learning agent, where we refer the reader to standard notions (Sutton and Barto, 1998) of reinforcement learning (RL).

2.1. Subnetworks

Let s_t be the **observed state** at time t and let h_t be the high-level representation derived from s_t (and from past observed values s_{t-k} in the partially observable case). For example, h_t could be the output of some kind of RNN (with whatever architecture is appropriate) that reads the sequence of s_t as input and produces an output h_t at each time step:

$$h_t = F(s_t, h_{t-1}) \tag{1}$$

where we call F the **representation RNN** or encoder and h_t the **representation state**. A core objective is to learn good representations in h_t, which disentangles abstract explanatory factors, in the sense that there exist a simple transformation of h_t which can select the information about a single factor (its value or uncertainty about it).

We can think of the representation RNN as the content of almost the whole brain at time t, i.e., the representation state h_t is a very high-dimensional vector (and probably sparse if we want to imitate biology),

which is an abstract representation of the full current information available to the agent (beyond what is stored in the weights), thus summarizing the current and recent past observations.

In contrast, we will define the **conscious state** c_t as a very low-dimensional vector which is derived from h_t by a form of attention mechanism applied on h_t, taking into account the previous conscious state as context.

$$c_t = C(h_t, c_{t-1}, z_t) \tag{2}$$

where z_t is a random noise source (only a notation to mean that C computes a distribution over the next thought given the previous thought and the current h, and C can be used to sample from that distribution). The cognitive interpretation is that the value of c_t corresponds to the content of a thought, a very small subset of all the information available to us unconsciously, but which has been brought to our awareness by a particular form of attention which picks several elements or projections from h_t. The function C is the **consciousness RNN** and because of its random noise inputs, produces a random choice of the elements on which the attention gets focused. This is useful if we think of the consciousness RNN as a tool for exploring interpretations or plans or to sample predictions about the future. We can also think of the consciousness RNN as the tool to isolate a particular high-level abstraction and extract the information about it (its value, uncertainty about it or even the fact that it is unobserved). This would happen if we think about a single factor, but in general C will aggregate a few (e.g. a handful) of such factors into a more complex and composed thought.

Below, we consider the special case of a conscious thought c_t which contains two different kinds of information: conditioning variables B_i with their value b_i, along with predicted variables A_j with their predicted value a_j, so that a verifier network (see below) can later verify the predictions and compute a loss to be optimized. The attention mechanism thus selects current (or hypothesized) variables (key-value pairs) as well as predicted variables. To make the actual predictions for the a_j's given the choice of A_j's and the conditioning set of pairs (B_i, b_i) one could use another network, a predictor or generator network $G(A_1, A_2, \ldots, B_1, b_1, B_2, b_2, \ldots)$ which outputs samples $(a_1, a_2, \ldots)$ associated with the predicted variables, or alternatively predicts parameters of a distribution over the A_j variables. So there is a single attention mechanism but when it selects a conditioning variable, the value of that variable can be read in the representation state h_t, while when it selects a predicted variable, the predicted value must be generated by G.

2.2. Training Objectives

To capture the assumption that a conscious thought can encapsulate a statement about the future, we introduce a **verifier network** which can match a current representation state h_t with a past conscious state c_{t-k}:

$$V(h_t, c_{t-k}) \in \mathbb{R} \tag{3}$$

which should be structured so that $V(h_t, c_{t-k})$ indicates the consistency of c_{t-k} with h_t, e.g., estimating the probability of the corresponding statement being true, given h_t. In the experiments proposed below, $k = 1$, i.e. we consider simple one-step predictions, although a more interesting application of the consciousness prior would be regarding longer-term predictions which are not attached to a particular time index but rather some trigger conditions.

For the proposed experiments, V will simply be the average negative log-likelihood of the selected variable A taking some predicted value a, where A corresponds to one of the dimensions of the representation state. The average is over the attention weights w_A given to every dimension A of h:

$$V = -\sum_A w_A \log P(h_{t,A} = a | c_{t-1}) \tag{4}$$

where $h_{t,A}$ is element A of h_t, the conditioning through c_{t-1} depends through the attention mechanism mostly on just a few elements of h_{t-1} and the weights w_A are computed also at $t-1$ as a function of c_{t-1}.

There are two distinct mechanisms at play which contribute to map the high-level state representation to the objective function: (1) the attention mechanism (e.g. the consciousness RNN) which selects and combines a few elements from the high-level state representation into a low-dimensional "conscious sub-state" object (the current content of our consciousness), and (2) the predictions or actions which are derived from the sequence of these conscious sub-states. The second mechanism is easy to grasp and frame in standard ML practice, either in deep learning or RL, e.g. for supervised or unsupervised or RL tasks.

For example, the attention mechanism could select elements B from the current representation state and choose to make a prediction about future elements A. Then to improve the quality of the prediction mechanism we may just want to maximize $logP(A|B)$ or some proxy for it, e.g., using a variational auto-encoder (Kingma and Welling, 2014) objective or a a conditional GAN (Mirza and Osindero, 2014) if one wants to sample accurately an A from B. Note again that such an objective function is not just used to learn the mapping from B to A (or to probabilities over the space of A values), but also drives the learning of the representation function itself, i.e., is back-propagated into the representation RNN). However, this part of the objective function (e.g. predictive value, computed by V above) is not sufficient and in fact is not appropriate to train the attention mechanism itself (which variables A and B should be selected?). Indeed, if that was the driving objective for attention, the learner would always pick a B which is trivially predictable (and there are such aspects of reality which are trivially predictable yet do not help us to further understand the world and make sense of it or achieve our goals). It remains an open question what other objectives would be appropriate for learning how to attend to the most useful elements, but ultimately we should be able to use the actual RL reward of the learning agent for that purpose (though some shorter-term proxy might be welcome). In the experiments below, we propose that a **policy network** be also trained, which is conditioned on c_t in order to take a decision. The conscious state thus becomes a form of planning (a more sophisticated form would involve a sequence of thoughts being formed before a decision is made, of course). The policy gradient acting on the policy network can then be back-propagated into the attention mechanism which selects the variables involved in the thought. In addition to an RL objective, some form of entropy or diversity may be needed so that the attention mechanism is stochastic and can choose a very diverse set of possible attended elements, so as to cover widely the possible variables A on which a prediction is made.

2.3. Naming Variables and Indirection

It would be very convenient for the consciousness attention mechanism and for the verifier network to be able to refer to the "names" of variables on which a prediction is made. In some models, we already distinguish keys and values in variations of memory augmented neural networks (Weston *et al.*, 2014; Graves *et al.*, 2014). The conscious state must indirectly *refer* to some of the aspects or dimensions computed in the representation h. Wether this should be done explicitly or implicitly remains to be determined. A key-value mechanism also makes it easier for the verifier network to do its job because it must match just the *key* of the predicted variable with its instances in a future representation state (with that variable becomes observed). If the key and value are mixed up and the predicted value differs substantially from the observed value, a simple associative process might miss the opportunity to match these and thus provide a strong training signal (to correct the predictor).

The simplest possible form of naming mechanism is possible when the representation state is made of grandmother cells (very unlikely in brains), i.e., with exactly one dimension per variable. In that case, the name of a variable is simply its index, and a set of attended variables can be represented by a binary vector with mostly zeros, and ones on the attended variables. To distinguish the conditioning variables from the predicted variables, a code could be used such as -1 for predicted, 0 for unused and 1 for conditioning variables.

3. Considerations for Experimenting with the Consciousness Prior

Because this is a novel theory which may be developped in many different ways, it is important to start with simple toy experiments allowing one to test and evaluate qualitatively different approaches, such that the turnaround time for each experiment is very short and the analysis of the representations learned very easy (because we already have a preconceived idea of what concepts would be the most appropriate to disentangle).

Although working with natural language input would be likely to help the agent learn better and more abstract representations, it would be better to start with experiments with no linguistic input, to make sure that it is the training objective and the training framework alone which are leading to the discovery of the appropriate high-level features. For example, learning some form of intuitive physics is done by babies without the need for linguistic guidance. Similarly, although the consciousness prior could be used in supervised learning or task-oriented RL, testing its ability alone to discover high-level abstractions would be best done in the context of unsupervised RL, e.g., us-

ing an intrinsic reward which favours the discovery of how the environment works.

It would be more interesting for the learning task to involve meaningful abstractions which have a high predictive power. For example, consider predicting whether a pile of blocks will fall on or off a table. It involves a high-level discrete outcome which can be predicted easily, even if the details of where the blocks will fall is very difficult even for humans to predict. In that case, predicting the future at the pixel level would be extremely difficult because future states have high entropy, with a highly multimodal distribution. However, some aspects of the future may have low entropy. If in addition, these aspects have a big impact on predicting what will come next (or on taking the right decisions now), then the consciousness prior should be very useful.

In terms of experimental comparisons, it would be good to compare systems based on the consciousness prior with systems based on more common RL approaches such as policy gradient deep RL on one hand, or model-based RL on the other hand (still with neural nets to learn the transition operator in sensory space, as well as the reward function). Even better, in toy problems we can compute the oracle solution, so we can get an upper bound on the best achievable performance.

3.1. Experiment's Desiderata

We propose to create a synthetic RL experiment with the following characteristics. To simplify the discussion, consider 3 time steps t_0, t_1 and t_2: the agent takes decisions a_0 and a_1 at t_0 and t_1 respectively with corresponding states s_0 and s_1 (and s_2 for the final state) and gets a reward R at t_2. Let us use upper case for random variables and lower case for their value. We design the synthetic task so that there exist a few abstract features of the state, in vector $h_t^* = f^*(s_t)$ with dimension(h_t^*) $\ll$dimension(s_t) (i.e. f^* throws away a lot of bits and returns a truly abstract aspect of the state) but dimension(h_t^*) is sufficiently large to create a combinatorial space (e.g. 10 dimensions), with the properties that (1) for $s_0 \sim P(s_0)$, $P(S_1|s_0, a_0)$ is very complex (e.g., has a large number of modes, which don't factorize in s-space but do in the space of h^*, i.e., most of the information in S_1 does not matter for the reward), (2) $P(H_t^*|h_{t-1}^*, a_{t-1})$ is simple but non-linear (logistic regression would not be enough to do a good job, but a very small MLP would do the job, and it should combine in such a non-linear way both the previous state and the previous action), (3) the reward R depends only on h_2^*, thus combining non-linearly a_0, a_1 and h_1^*.

As a consequence of (1) and (2) applied for $t = 1$, a traditional model-based RL approach (which tries to capture $P(S_1|s_0, a_0)$ explicitly) would require a lot of data and capacity, whereas the approach based on the consciousness prior would only need to model the simpler transition in h^* space (or its learned proxy). We also expect that this setup will mean trouble for policy-gradient (model-free) RL as a consequence of (2) and based on the experience observed by Gulcehre and Bengio (2016), plus the usual difficulties arising in policy gradient due to the non-stationarity of the learning environment, because as the policy changes, the target value or Q function changes.

Many experiments could be designed which satisfy the above requirements. As an example, consider using element-wise XOR and the modulo operation: $h_{t+1}^* = h_t^*$ XOR $f(a_t)$, where f appropriately maps the action a_t to a binary vector, and $R = \text{parity}(h_2)$, which should tangle the many bits of action in a somewhat confusing way. Because this transition function is element-wise, it means that one can predict one of the dimensions $h_{t+1,i}^*$ using only two bits of information, $h_{t,i}^*$ and $f_i(a_t)$, exactly the kind of assumption which matches well with the consciousness prior. In other words, we expect conscious states to encapsulate a statement about such three bits (the two conditioning bits and the predicted bit) while we expect the representation state to contain the same information as the full h^*.

What remains to be done is to design a visual representation for the state that makes it easy to read out the underlying h^* and also makes it non-trivial to map the state (an image) to the h^* space, as well as design a correspondingly intuitive action space with a non-revealed map to the same space, via the function f above.

References

Bengio, Y. (2017). The consciousness prior. *arXiv preprint arXiv:1709.08568*.

Graves, A., Wayne, G., and Danihelka, I. (2014). Neural turing machines. *arXiv preprint arXiv:1410.5401*.

Gulcehre, C. and Bengio, Y. (2016). Knowledge matters: Importance of prior information for optimization. *Journal of Machine Learning Research*, **17**(8), 1–32.

Kingma, D. P. and Welling, M. (2014). Auto-encoding variational bayes. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

Sutton, R. and Barto, A. (1998). *Reinforcement Learning: An Introduction*. MIT Press.

van Gulick, R. (2004). Consciousness. In *Stanford Encyclopedia of Philosophy*.

Weston, J., Chopra, S., and Bordes, A. (2014). Memory networks. *arXiv preprint arXiv:1410.3916*.