Attention bias towards structure explained by an intrinsic reward for learning

Zhi-Wei Li (zhiwei.li@nyu.edu)
Center for Neural Science, New York University
New York, NY, USA

Todd M. Gureckis (todd.gureckis@nyu.edu)
Department of Psychology, New York University
New York, NY, USA

Abstract

Recent studies have found attention can be drawn and sustained toward particular types of sequential regularities in adults and young infants. We propose these results are naturally accommodated by an attention bias towards parts of environments with more "learning progress" (i.e., improvement in understanding and reduction of uncertainty about a regularity). Our theory provides an a-priori theoretical account of a variety of behavioral findings. Overall, this is the first step of a project evaluating the concept of "learning progress" to explain spontaneous attention allocation.

Keywords: attention; intrinsic motivation; learning progress;

Previous findings: Attention bias in sequence

Recent studies indicate that a spontaneous learning process happening in both adults and young infants during the exposure to a sequence of stimuli, even in the absence of any task or goal. In Zhao et al (2013), adults show enhanced attention to locations presenting a structured sequence (items organized into triplets in contrast to random sequences). Furthermore, infants have been shown to attend less to items either too surprising or too unsurprising (Kidd, Piantadosi, & Aslin, 2012), quantified by the predicted log probability (or information content, IC) of the item under a probabilistic learning model. Intermediate probability stimuli are proposed to be more "learning worthy" resulting in a "u-shaped" relationship between attention and element probabilities (Figure2A). While quite distinct, these studies indicate that the participants spontaneously learn and attempt to predict the structure while exposed to a sequence. Our goal here is to advance a novel theoretical account of these phenomena drawing from recent research on intrinsic motivation in AI and robotics.

Learning progress as an internal reward

It is well established that learning follows predictable patterns over time, reflected in the idealized learning curve such as exponential (Dubey & Griffiths, 2017) or the more general Gompertz growth curve (Pelz, Piantadosi, & Kidd, 2015). One natural definition of learning progress is then defined as the derivative of the learning curve(Figure 1A).

Such idealized curves, however, bear only a weak connection to the sequence learning experimental paradigms reviewed earlier. Thus it is useful to more directly specify the learning objective based on the specific learning models for
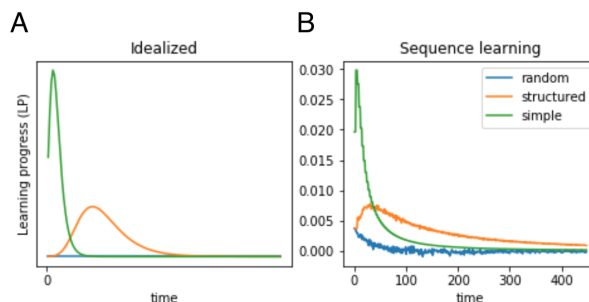

Figure 1: A. Learning progress with three parameter settings. B. Learning progress defined by Eq 1 in sequence learning. The "structured" (9 items organized into triplets) and "random" (9 items uniformly presented) sequence mimics the stimuli in Zhao et al.(2013). A "simple" sequence containing only 3 items is added to demonstrate different levels of difficulties.

sequence, such as counting the transition probability P as is done in Kidd et al.(2012). We use the Shannon entropy H to capture this intuition and define learning progress as the instantaneous entropy reduction:

$$LP(t) = H(P(t)) - H(P(t-1)) \qquad (1)$$

Figure 1B illustrates the example learning progress curves applied to actual sequential materials. They do resemble the shape derived from Golmpertz curve despite some differences. For example, for the total random sequence, the learning progress does not start strictly from zero due to some initial learning before the randomness is fully experienced.

Agreement with previous accounts

The learning progress curves in Figure 1B readily explains the attention bias towards regularity (Zhao et al., 2013): the learning progress for the structured sequence is almost always higher than the random sequence thus deserves more attention.

Regarding relation between learning progress and information content (IC), we found a U-shape curve relating LP and IC in simulated sequences, suggesting an a-priori account of the u-shaped "Goldlilocks" effect (Figure 2B).
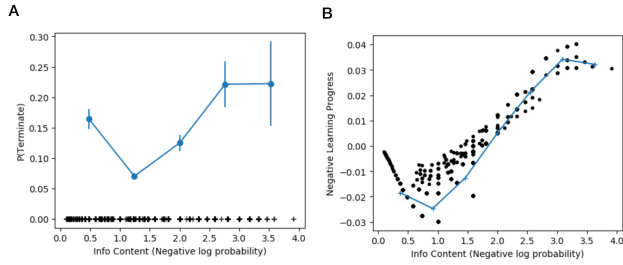
Figure 2: A. Empirical U-shape curve between probability of terminating fixation and information content. Reproduced using the data from (Piantadosi et al., 2014). Black "+"s marks the frequency of stimuli in that specific *IC* B. Negative learning progress also shows a U-shape relation with information content. Each dot is one stimuli presented and blue line is binned average.

Predicting attention dynamics in changing environment

More subtle attention biases have been observed in the variations of the Zhao et al. paradigm. For example, at the second half of the experiment if a structured sequence loses its transition regularity into random or vice versa, even though now both location presents same types of sequence, attention is still oriented towards the previously biased location(Yu & Zhao, 2015). Only when both sequences change to the opposite type will the attention difference become non significant (p>.05).

To explain quantitative attention difference, we formulate the attention allocation mechanism as maximizing the internal reward: expected learning progress. This idea has been implemented in AI and robotics literature but usually without a probabilistic world model(Oudeyer, Kaplan, & Hafner, 2007; Luciw, Graziano, Ring, & Schmidhuber, 2011; Pathak, Agrawal, Efros, & Darrell, 2017). There are different ways the learner might integrate historical *LP* to make the inference about future. Specifically we assume a delta update rule to calculate the expected learning progress:

$$\mathbb{E}[LP_{t+1}] \leftarrow \mathbb{E}[LP_t] + \alpha \cdot (LP_t - \mathbb{E}[LP_t]) \quad (2)$$

where α is a learning rate.

To decide which sequence to be engaged next, we feed the expected reward $\mathbb{E}[LP_{t+1}]$ into an softmax function to balance exploration and exploitation.

Our simulation is then able to replicate both the original regularity bias (the first column in Figure3), the "lingering effect" (the second and third column), and its elimination (the fourth column) where the effects become less significant (p>0.05). We arbitrarily set reaction time for attended versus non-attended stimuli to be drawn from $\mathcal{N}(1.3, 0.6)$ and $\mathcal{N}(1.55, 0.6)$ respectively to convert attention into time difference. More carefully model RT variances between and within
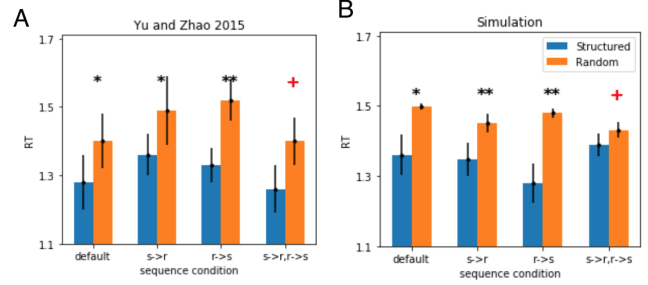
individuals is to be done.[1]

Figure 3: Replicating (Yu & Zhao, 2015) experiment 1-4 where the sequence type changes in the second half of the experiment. The RT advantage for structured sequence remains to be significant unless both sequences change (the fourth column where $p > 0.05$). The four columns indicate the four conditions of the second half, the first with no change, second and third with only one location changed, the fourth with both location changes. On x-axis labels the specific "s" means structured and "r" means random. "+" marks $p < 0.1$,"*" makrs $p < 0.05$,"**" makrs $p < 0.01$.

Discussion

We have shown that a number of previous phenomena where attention seemed automatically captured by particular types of sequential regularity might best be explained as "learning progress." While this is the first stages of this exploration, the link between research in AI and robotics and human psychology is intriguing. A number of important future questions remain including the nature and capacity of the learnign models people apply, and how people will respond to other aspects of the task such as novelty (e.g., Yu et al.(2015)).

References

Dubey, R., & Griffiths, T. L. (2017, May). A rational analysis of curiosity. *arXiv:1705.04351 [cs]*. (arXiv: 1705.04351)

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012, May). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLOS ONE*, *7*(5), e36399.

Luciw, M., Graziano, V., Ring, M., & Schmidhuber, J. (2011, August). Artificial curiosity with planning for autonomous perceptual and cognitive development. In *2011 IEEE International Conference on Development and Learning (ICDL)* (pp. 1–8). Frankfurt am Main, Germany: IEEE.

Oudeyer, P., Kaplan, F., & Hafner, V. V. (2007, April). Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation*, *11*(2), 265–286. doi: 10.1109/TEVC.2006.890271

[1]Simulation done with softmax noise parameter $\beta = 1000$, learning rate $\alpha = .001$.

Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. In *Proceedings of the ieee conference on computer vision and pattern recognition workshops* (pp. 16–17).

Pelz, M., Piantadosi, S. T., & Kidd, C. (2015, August). The dynamics of idealized attention in complex learning environments. In *2015 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 236–241). Providence, RI, USA: IEEE.

Piantadosi, S. T., Kidd, C., & Aslin, R. (2014). Rich analysis and rational models: Inferring individual behavior from infant looking data. *Developmental Science*, *17*(3), 321–337.

Yu, R. Q., & Zhao, J. (2015, October). The persistence of the attentional bias to regularities in a changing environment. *Attention, Perception, & Psychophysics*, *77*(7), 2217–2228.

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013, May). Attention Is Spontaneously Biased Toward Regularities. *Psychol Sci*, *24*(5), 667–677.