How do infants acquire knowledge?_ We believe that understanding spatial events is crucial. We investigate the important events like occlusion and launching in a unified framework.

Many of the early conceptions of events have been suggested to be innate [4, 5]. Here, we demonstrate that the conceptions of occlusion, object permanence and causality can be learned.

We model two habituation-dishabituation experiments on occlusion and causality perception. A simple recurrent network—the Elman network—is trained to predict its next inputs. The prediction error is taken as a model of the looking time in infant experiments.

Training

The network was trained in 3 phases: pre-training, habituation and test phase.

After pre-training, we “habituated” (i.e. trained) the network for additional 1000 time steps with the habituation stimuli a), b) or c). Then, we tested the networks performance, i.e. prediction capacity, with the test stimuli d) and e) or f) and e). The total prediction errors for each test stimulus were then compared to the prediction errors during the habituation. The difference, i.e. the dishabituation “error” is plotted in the figure above.

Experimental results:

1) Looking times declined significantly from first to last habituation trials.
2) The group shown the direct launching stimulus increased its looking time significantly more than the group shown delayed launching or launching without occlusion.
3) The no-prior-movement stimulus attracted significantly longer looking times than the no-reaction stimulus, regardless of the group.

The model explains all of these results.

Discussion

In his experiment, Leslie tested whether a temporal or spatial gap would disrupt the impression of “causality” in infants. Indeed, the results indicate that it does. Leslie concluded that a gap makes the infant perceive two separate movements instead of one launching movement.

Our model shows the same behavior as infants but there are no separate movements in the inputs. There are just sequences of input vectors. The model shows that one can account for Leslie’s data assuming that the infant just learns to predict spatiotemporal patterns in its visual input.

Similarly, the model accounts for the occlusion data. Here, the network learns to represent the occluded object. But it can only do so with time. Therefore, the network shows a similar developmental trajectory as infants seem to follow. Moreover, we can show that the network develops an actual representation of the occluded object which is difficult to do with infants. Thus, the model provides us with an explanation of how and why change occurs.

Methods

The Elman network consists of an input, hidden, context and output layers. Hidden unit activities, $Y_t$ and output unit activities, $Z_t'$, are defined as:

$$Y_t = 1 + \exp \left( -\sum_{j=0}^{N} w_{ij}X_t + \sum_{j=0}^{N} u_{ij}C_{t-1} \right)^{-1}$$

and

$$Z_t' = 1 + \exp \left( -\sum_{j=0}^{N} w_{ij}Y_t \right)^{-1}$$

Each time step, the hidden layer activities are copied into the context layer: $C_t = Y_t$. Training with error backpropagation with following objective function:

$$E(t) = \sum_{t=1}^{T} (Z(t) - X(t))^2$$

References