Representation of spatial and conceptual knowledge in a Neural Network

Project of November 2021

1 Neural Network Model

The hypothesis of a common representation space between spatial and conceptual knowledge was then investigated on Neural Networks.
We trained our Neural Network on the same stimuli and stimulus-reward maps that the one presented to the experimental participants. The Network is composed of two hidden layers with a dynamic structure depending on the nature of the task (spatial or conceptual). Figure 1 shows the Network architecture.

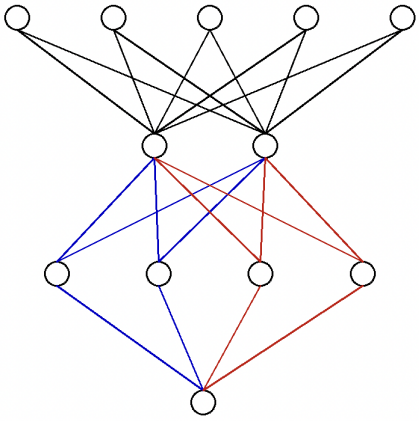

Figure 1: **Description of the Network architecture**. Sensory inputs corresponding to position (x,y) and weather condition (rain, sun) are mapped onto the first representational space in the first hidden layer. The seed value is added to the input vector as an indicator of the task and dictates the connections in the second and third layers. Connections in black are constant and define the shared representational subspace while connections in red (blue) are only activated for spatial tasks (conceptual tasks).

The network was exposed to N blocks of X trials. The nature of the task is constant within each block and switches from one block to another. Networks weights are updated after each block.

To compare the impact of a shared representational space on the training two independent Networks with an equivalent structure were implemented and described Figure 2.

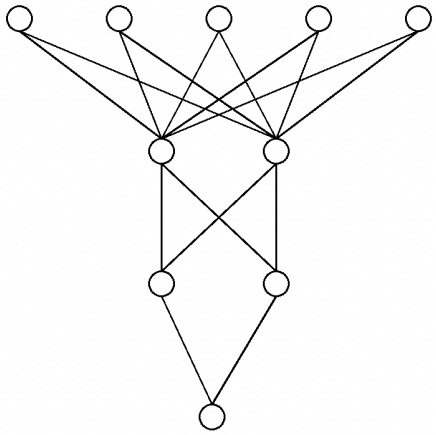

Figure 2: **Description of the Independent Network architecture.** Two independent models are trained respectively on the spatial and the conceptual task.

The dataset being the same, the two Networks were trained with the same number of epochs as for the two sub-networks.
If spatial and conceptual knowledge share a common representation in the first hidden layer, training one sub-network will automatically improve the other. As a result we will expect to see a faster convergence of the accuracy compared to the two independent Networks.