Mental mapping

Gregor Schöner
Institut für Neuroinformatik
Ruhr-Universität Bochum, Germany
gregor.schoener@rub.de
Recall from last lecture: Concepts, relational thinking

Talking about objects entails bringing the targeted object into the attentional foreground

“red to the left of green”

[Lipinski, Sandamirskaya, Schöner 2009
... Richter, Lins, Schöner, Topics 2017]
into the reference and target field and enable these fields to track moving objects even if spatial attention is currently focused elsewhere.

3.2. Attention

The core of the attentional system consists of two three-dimensional attention fields. They are defined over the same dimensions as the two perception fields, but their activation remains below threshold unless additional input arrives from a feature attention field or a spatial attention field.

Fig. 2. Architecture with activation snapshots while it is generating a phrase about a video. Fields are shown as color-coded activation patterns; for three-dimensional fields, two-dimensional slices are shown. Node activation is denoted in opacity-coded circles. Spatial templates are illustrated as color-coded weight patterns (bottom left). Excitatory synaptic connections are denoted by lines with arrowheads, inhibitory connections by lines ending in circles. Transformations to and from polar coordinates are marked with a “T.” Steerable neural mappings are denoted as diamonds.
into the reference and target field and enable these fields to track moving objects even if spatial attention is currently focused elsewhere.

3.2. Attention

The core of the attentional system consists of two three-dimensional attention fields. They are defined over the same dimensions as the two perception fields, but their activation remains below threshold unless additional input arrives from a feature attention field or a spatial attention field.

Fig. 2. Architecture with activation snapshots while it is generating a phrase about a video. Fields are shown as color-coded activation patterns; for three-dimensional fields, two-dimensional slices are shown. Node activation is denoted in opacity-coded circles. Spatial templates are illustrated as color-coded weight patterns (bottom left). Excitatory synaptic connections are denoted by lines with arrowheads, inhibitory connections by lines ending in circles. Transformations to and from polar coordinates are marked with a "T." Steerable neural mappings are denoted as diamonds.
3.2. Attention

The core of the attentional system consists of two three-dimensional attention fields. They are defined over the same dimensions as the two perception fields, but their activation remains below threshold unless additional input arrives from a feature attention field or a spatial attention field.

Fig. 2. Architecture with activation snapshots while it is generating a phrase about a video. Fields are shown as color-coded activation patterns; for three-dimensional fields, two-dimensional slices are shown. Node activation is denoted in opacity-coded circles. Spatial templates are illustrated as color-coded weight patterns (bottom left). Excitatory synaptic connections are denoted by lines with arrowheads, inhibitory connections by lines ending in circles. Transformations to and from polar coordinates are marked with a “T.” Steerable neural mappings are denoted as diamonds.

into the reference and target field and enable these fields to track moving objects even if spatial attention is currently focused elsewhere.

3.2. Attention

The core of the attentional system consists of two three-dimensional attention fields. They are defined over the same dimensions as the two perception fields, but their activation remains below threshold unless additional input arrives from a feature attention field or a spatial attention field.

Fig. 2. Architecture with activation snapshots while it is generating a phrase about a video. Fields are shown as color-coded activation patterns; for three-dimensional fields, two-dimensional slices are shown. Node activation is denoted in opacity-coded circles. Spatial templates are illustrated as color-coded weight patterns (bottom left). Excitatory synaptic connections are denoted by lines with arrowheads, inhibitory connections by lines ending in circles. Transformations to and from polar coordinates are marked with a “T.” Steerable neural mappings are denoted as diamonds.

relational neural operators
3.2. Attention

The core of the attentional system consists of two three-dimensional attention fields. They are defined over the same dimensions as the two perception fields, but their activation remains below threshold unless additional input arrives from a feature attention field or a spatial attention field.

Fig. 2. Architecture with activation snapshots while it is generating a phrase about a video. Fields are shown as color-coded activation patterns; for three-dimensional fields, two-dimensional slices are shown. Node activation is denoted in opacity-coded circles. Spatial templates are illustrated as color-coded weight patterns (bottom left). Excitatory synaptic connections are denoted by lines with arrowheads, inhibitory connections by lines ending in circles. Transformations to and from polar coordinates are marked with a “T.” Steerable neural mappings are denoted as diamonds.
“red to the left of green”
Mental mapping and inference

- making sense of propositions (about spatial relations) purely mentally, without any perception to ground in
- and operating on such “sense” by drawing inferences...
Mental mapping and inference

- mental map formation from propositions
  - “There is a cyan object above a green object.”
  - “There is a red object to the left of the green object.”
  - “There is a blue object to the right of the red object.”
  - “There is an orange object to the left of the blue object.”

- inference
  - “Where is the blue object relative to the red object?”
Figure 1: Activation snapshot of the architecture as it forms a mental model consisting of five objects. For two-dimensional fields, activation is shown color-coded, where blue colors denote subtreshold and yellow colors denote suprathreshold activation. For three-dimensional fields, two-dimensional slices of activation are shown. Neural nodes are denoted by circles that are filled if the node is active and empty if inactive. Excitatory synaptic connections are shown by black lines with arrowheads, inhibitory connections by lines ending in black circles; patterned connections are marked with a star. Steerable neural mappings are denoted by blue diamonds. See text for details.

The spatial transformation system represents these three elements in dedicated dynamic neural fields, the target field, the relational field, and the reference field, respectively. The target field and reference field are defined over two-dimensional space and receive input from the attention field. Whenever there is a peak in the attention field, one of the fields may be brought into the dynamic regime to form peaks. The two-dimensional relational field represents the relative position of a target object with respect to the reference object. The field is defined such that the reference object would be in the center of the field. The relational field also receives input from the production nodes of all spatial relation concepts (e.g., TO THE LEFT OF, see Figure 1). Coordinate transformations between the absolute spatial positions in the target field and the relative positions in the relational field are based on steerable neural mappings (blue diamonds in Figure 1; Schöner, 2012), which are approximated by convolutions here. The architecture has three such coordinate transforms: the first (leftmost blue diamond) enables the position of an already existing target object to be transformed into the relational field. This enables the architecture to make inferences on an already established mental model. The second coordinate transform (middle diamond) enables the model to transform peaks in the relational field back into the target field. This path accounts for the creation of new objects in the scene: a peak is induced in the relational field from the spatial template that represents one of the spatial relations. The position in space where the peak forms determines where the new object is going to be placed in space. The third transformation (right diamond) has a crucial impact on the position where the peak forms in the relational field. It transforms the output of the spatial scene representation field and feeds inhibitorily into the relational field, introducing inhibition in positions that are already occupied by objects in the mental model. Due to this inhibition, peaks induced in the relational field tend to shift further outward, avoiding changes to the already established mental model. This is consistent with the preferred mental models that humans tend to build (Ragni & Knauff, 2013).
Figure 1: Activation snapshot of the architecture as it forms a mental model consisting of five objects. For two-dimensional fields, activation is shown color-coded, where blue colors denote subtreshold and yellow colors denote suprathreshold activation. For three-dimensional fields, two-dimensional slices of activation are shown. Neural nodes are denoted by circles that are filled if the node is active and empty if inactive. Excitatory synaptic connections are shown by black lines with arrowheads, inhibitory connections by lines ending in black circles; patterned connections are marked with a star. Steerable neural mappings are denoted by blue diamonds. See text for details.

The spatial transformation system represents these three elements in dedicated dynamic neural fields, the target field, the relational field, and the reference field, respectively. The target field and reference field are defined over two-dimensional space and receive input from the attention field. Whenever there is a peak in the attention field, one of the fields may be brought into the dynamic regime to form peaks. The two-dimensional relational field represents the relative position of a target object with respect to the reference object. The field is defined such that the reference object would be in the center of the field. The relational field also receives input from the production nodes of all spatial relation concepts (e.g., TO THE LEFT OF, see Figure 1). Coordinate transformations between the absolute spatial positions in the target field and the relative positions in the relational field are based on steerable neural mappings (blue diamonds in Figure 1; Schneegans & Schöner, 2012), which are approximated by convolutions here. The architecture has three such coordinate transforms: the first (leftmost blue diamond) enables the position of an already existing target object to be transformed into the relational field. This enables the architecture to make inferences on an already established mental model. The second coordinate transform (middle diamond) enables the model to transform peaks in the relational field back into the target field. This path accounts for the creation of new objects in the scene: a peak is induced in the relational field from the spatial template that represents one of the spatial relations. The position in space where the peak forms determines where the new object is going to be placed in space. The third transformation (right diamond) has a crucial impact on the position where the peak forms in the relational field. It transforms the output of the spatial scene representation field and feeds inhibitory into the relational field, introducing inhibition in positions that are already occupied by objects in the mental model. Due to this inhibition, peaks induced in the relational field tend to shift further outward, avoiding changes to the already established mental model. This is consistent with the preferred mental models that humans tend to build (Ragni & Knauff, 2013).
“blue right of red”
“where is blue relative to green?”

deep spatial relation right

target color memory

target color

target centered on reference

cyan

red

blue

green

red

blue

cyan

spatial relation memory

spatial relation

right

reference color memory

reference color

reference processes

attention (space)

spatial scene representation

red

blue

target

reference
How do the sequential processing steps emerge from neural dynamic systems?

- this is an issue in all integrated neural architecture....

- but also a broader issue... lecture on sequence generation