Compositional Models of Vector-based Semantics, ESSLLI 2022

Background reading for this exercise is Moortgat & Wijnholds (2017) url which gives a vector-based semantics for Dutch relative clauses. The exercises below deal with relative clauses in English.

1 Syntax

We return to an example from an earlier exercise

$$\textsf{games that kids love but parents hate}$$

for which you arrived at the lexicon below, where X and X' are s/np with appropriate $\Diamond, \Box$ decoration.

$$\begin{array}{rcl} \textsf{games} & :: & n \\ \textsf{kids, parents} & :: & np \\ \textsf{love, hate} & :: & (np\backslash s)/np \\ \textsf{that} & :: & (n\backslash n)/(s/\Diamond\Box np) \\ \textsf{but} & :: & (X\backslash X'/X) \end{array}$$

To show that games that kids love (or games that parents hate) is a well-formed phrase of type n, you can make a derivation for

$$n_0 \cdot ((n_1\backslash n_2)/(s_3/\Diamond\Box np_4) \cdot (np_5 \cdot (np_6\backslash s_7)/np_8)) \vdash n_9$$

using the categorical presentation of $\mathbf{NL}_{\Diamond,\Box,/,\otimes,\backslash}$ (Appendix A.1 of the Part 1 exercises set). Work with the Residuation+Monotonicity rules, plus the rule version of controlled semi-associativity $P1$. Atomic formula occurrences have been given a unique identifier that will be useful for the exercises that follow.

2 Interpretation

For the interpretation homomorphism, we set $\lceil s \rceil = \mathsf{S}$ and $\lceil n \rceil = \lceil np \rceil = \mathsf{N}$, in other words, common nouns and full noun phrases are mapped to the same vector space. We furthermore assume that $dim(\mathsf{S}) = dim(\mathsf{N}) = 3$ and that coefficients take values in $\mathbb{R}$. Syntactic phrases of type s, np, n then are all sent to interpretations in $\mathbb{R}^3$, but we allow for the N and S spaces to have different bases.

For concreteness, let's take the following interpretation for lexical items with type np or n.

$$\begin{array}{rcl} \mathbf{games} & = & (\ -11 \quad 13 \quad -17\ ) \\ \mathbf{kids} & = & (\ 5 \quad -2 \quad 7\ ) \\ \mathbf{parents} & = & (\ -6 \quad 4 \quad 3\ ) \end{array}$$

2.1

The interpretation of a transitive verb with type $(np\backslash s)/np$ is a $3 \times 3 \times 3$ tensor in the rank-3 tensor space $\mathsf{N} \otimes \mathsf{S} \otimes \mathsf{N}$. This tensor represents a multilinear (here: bilinear) map $f : \mathsf{N} \times \mathsf{N} \to \mathsf{S}$ sending two vectors in N (subject, object) to a vector in S.

To give a concrete toy interpretation for the transitive verbs **love** and **hate**, we use the *cross product* of the subject and object vectors. The cross product of vectors **a** and **b** (notation: $\mathbf{a} \times \mathbf{b}$) is a vector perpendicular to the given vectors, with a direction given by the right-hand corkscrew rule and magnitude equal to the area of the parallelogram spanned by vectors **a** and **b**.

The Levi-Civita function ϵ below gives the elements of a $3 \times 3 \times 3$ tensor T which computes the cross product of vectors **a** and **b** via the tensor contraction $\mathbf{a}_i T_{ijk} \mathbf{b}_k$. Even permuations of 123 are the cyclic permutations: $123, 231, 312$; uneven permutations are the anti-clockwise $321, 213, 132$.

$$\epsilon_{ijk} = \begin{cases} 0 & \text{of } i = j \text{ or } j = k \text{ or } k = i \\ 1 & \text{if } ijk \text{ is an even permutation of } 123 \\ -1 & \text{if } ijk \text{ is an uneven permutation of } 123 \end{cases}$$

We can now extend our lexicon with interpretations for love and hate.

$$\begin{array}{rcl} \mathbf{love}_{ijk} & = & \epsilon_{ijk} \\ \mathbf{hate}_{ijk} & = & -\epsilon_{ijk} \end{array}$$

Your turn Write out the $3 \times 3 \times 3$ arrays for **love** and **hate**. In other words, fill the 27 cells of these tensors with the values prescribed by $\pm\epsilon_{ijk}$. Index i is for the layers of the cube, j for rows, k for columns.

2.2

Consider now the example games that kids love. Determine, on the basis of the axiom linking for your derivation of §1, which indices have to be indentified. You can do this with a relabeling, using the Einstein summation convention, or by giving the (generalized) Kronecker δ function.

$$\mathbf{games}_i \otimes \mathbf{that}_{jklm} \otimes \mathbf{kids}_n \otimes \mathbf{love}_{opq}$$

2.3

Compute the semantic value for kids love, in other words the interpretation for the following subproof of the sequent derivation in §1:

$$np \cdot (np\backslash s)/np \Rightarrow s/\Diamond\Box np$$

Hint The goal formula $s/\Diamond\Box np$ finds its interpretation in the tensor space $\mathsf{S} \otimes \mathsf{N}$. This is a $dim(\mathsf{S}) \times dim(\mathsf{N}) = 3 \times 3$ matrix with elements

$$z_{jk} = \sum_{i=1}^{dim(\mathsf{N})} x_i \, y_{ijk}$$

computed by means of tensor contraction of the subject vector $\mathbf{x} \in \mathsf{N}$ with the first component of the cube $\mathbf{y} \in \mathsf{N} \otimes \mathsf{S} \otimes \mathsf{N}$ for the transitive verb.

2.4

The next step is the interpretation of the relative pronoun that. As discussed in class, we want to see the combination of a noun, e.g. games, and a relative clause, e.g. that kids love, as the elementwise multiplication of a vector **games** $\in \mathsf{N}$ and a vector $\mathbf{w} \in \mathsf{N}$ extracted from the interpretation of kids love, We obtain the desired semantic effect in two steps:

1. The interpretation of kids love, as computed in §2.3, is a 3×3 matrix $M \in \mathsf{S} \otimes \mathsf{N}$. We can reduce M to a vector $\mathbf{w} \in \mathsf{N}$ by multiplying on the left with $\mathbf{1}$:

   $$(1\ 1\ 1) \cdot M = \mathbf{w}$$

   Compute $\mathbf{w}$ with the value for M from your solution for §2.3.

2. For the second step, we need a bilinear map $f : \mathsf{N} \times \mathsf{N} \to \mathsf{N}$ sending the vectors **games** and $\mathbf{w}$ from (1.) to their elementwise multiplication **games** $\odot\, \mathbf{w}$.

   The desired map can take the form of a tensor $\mathbf{c} \in \mathsf{N} \otimes \mathsf{N} \otimes \mathsf{N}$ with elements $c_{ijk} = 1$ if $i = j = k$ and zero otherwise. Tensor contraction of $\mathbf{c}$ with vectors $\mathbf{x}, \mathbf{y} \in \mathsf{N}$ then effectively computes the vector $\mathbf{z} = \mathbf{x} \odot \mathbf{y} \in \mathsf{N}$ with elements

$$z_j = \sum_{i,k=1}^{dim(\mathsf{N})} x_i \, c_{ijk} \, y_k \qquad (1 \leq j \leq 3)$$

Substitute for $\mathbf{x}$ the semantic value for **games** given above and for $\mathbf{y}$ the vector $\mathbf{w}$ you found for (1.) and calculate the result.

2.5

Consider now the example we started with

games that kids love but parents hate

To interpret this example, we need to update our lexicon with a semantic value for the conjunction but. Below we repeat its syntactic type (where X, X' are your modally decorated versions of s/np) and the corresponding tensor space.

	syntactic type	semantic type
but	$(X \backslash X')/X$	$(\mathsf{S} \otimes \mathsf{N}) \otimes (\mathsf{S} \otimes \mathsf{N}) \otimes (\mathsf{S} \otimes \mathsf{N})$

As with §2.4 we want to capture the 'intersective' semantics of conjunction in terms of elementwise multiplication. In §2.4 this was elementwise multiplication of simple vectors $\mathbf{x}, \mathbf{y} \in \mathsf{N}$, computed by means of a $3 \times 3 \times 3$ tensor $\mathbf{c} \in \mathsf{N} \otimes \mathsf{N} \otimes \mathsf{N}$ with elements

$$c_{ijk} = \begin{cases} 1 & \text{if } i = j = k \\ 0 & \text{otherwise} \end{cases}$$

For but we are looking for a tensor

$$\overline{\mathbf{c}} \in (\mathsf{S} \otimes \mathsf{N}) \otimes (\mathsf{S} \otimes \mathsf{N}) \otimes (\mathsf{S} \otimes \mathsf{N})$$

for the conjunction of matrices $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathsf{S} \otimes \mathsf{N}$ interpreting **kids love** and **parents hate**.

Your tasks:

1. What are the elements of the tensor $\overline{\mathbf{c}}$? Enumerating the $3^6 = 729$ elements is not an attractive option. Give the schema
$$\overline{c}_{ijklmn} = \begin{cases} 1 & \text{if} \ldots \\ 0 & \text{otherwise} \end{cases}$$
2. Compute the interpretation of parents hate, as you did for kids love in §2.3.
3. Compute the final result: the interpretation of kids love but parents hate with
$$\overline{z}_{kl} = \sum_{i,j,m,n=1}^{3} \overline{x}_{ij} \, \overline{c}_{ijklmn} \, \overline{y}_{mn} \qquad (1 \leq k, l \leq 3)$$
where $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are the matrices for the interpretation of kids love en parents hate and $\overline{\mathbf{z}} = \overline{\mathbf{x}} \odot \overline{\mathbf{y}}$ the matrix for the conjunction kids love but parents hate.

2.6 Bonus

What are the elements of the $3 \times 3 \times 3 \times 3$ tensor **that** $\in \mathsf{N} \otimes \mathsf{N} \otimes \mathsf{S} \otimes \mathsf{N}$? This is the tensor that results from steps (1.) and (2.) of the previous exercise. Complete the index formula below

$$\mathbf{that}_{ijkl} = \begin{cases} 1 & \text{if} \ldots \\ 0 & \text{if} \ldots \end{cases}$$

rather than enumerating the 81 elements explicitly ...

□