Ling 5801: Lecture Notes 21
Compositional Semantics

Following models of distributed associative memory from computational cognitive neuroscience (Marr, 1971; Anderson et al., 1977; Murdock, 1982; McClelland et al., 1995; Howard and Kahana, 2002), the broad-coverage sentence processing model used in this article is defined in terms of *referential states*, which generalize stimuli as characteristic patterns of neural activation in the brain, and *cued associations*, which associate referential states through potentiation of synapses between neurons that are active in a cue state and neurons that are active in a target state. In this article, referential states are notated with variables x, y, and z, and cued associations are notated as functions $\mathbf{f}$ from (cue) referential states to (target) referential states. Some referential states are then assumed to be *elementary predications* (Copestake et al., 2005). Elementary predications are referential states which have:

1. *predication types*, characteristic parts of activation patterns shared across elementary predication instances, notated here by $\mathbf{f_0}$ functions from (predication) referential states to type specifications, and

2. distinguished cued associations to *participant* referential states, notated here by numbered functions $\mathbf{f_1}$, $\mathbf{f_2}$, etc., from (predication) referential states to (participant) referential states.[1]

Collections of referential states connected by elementary predications form *cued association structures*, notated here using functions p and q from referential states to truth values, which are defined to be true if a particular structure holds at a particular referential state. These cued association structures are similar to semantic dependency structures (Kintsch, 1988; Mel'čuk, 1988; Kruijff, 2001; Baldridge and Kruijff, 2002; Copestake et al., 2005; White, 2006). For example, the cued association structure $p = \lambda_x \exists_e (\mathbf{f_0}\, e)=\mathbf{BeingOpen} \wedge (\mathbf{f_1}\, e)=x$ defines a structure at a referential state x that is the first participant of a 'being open' elementary predication e.

The sentence processing model used in this article also assumes referential states that represent narrower generalizations can inherit from referential states

[1] Reciprocal cued associations from participants to elementary predications may also be assumed, but the stronger direction, from elementary predications to unique participants, is notated.

that represent broader generalizations through the use of cued associations distinguished for *restriction inheritance*, *conjunction inheritance* and *extraction inheritance*, notated here as $\mathbf{f_{rin}}$, $\mathbf{f_{cin}}$ and $\mathbf{f_{ein}}$ functions from (narrower) referential states to (broader) referential states. For example, the cued association structure:

$$\lambda_y \, \exists_{x,e,e'} \, (\mathbf{f_0}\, e)\mathord{=}\mathbf{BeingADoor} \wedge (\mathbf{f_1}\, e)\mathord{=}x \wedge (\mathbf{f_0}\, e')\mathord{=}\mathbf{BeingOpen} \wedge (\mathbf{f_1}\, e')\mathord{=}y \wedge (\mathbf{f_{rin}}\, y)\mathord{=}x$$

defines a dependency structure at a referential state y that is the first participant of a 'being open' elementary predication e' and inherits from a referential state x the property of being the first participant of a 'being a door' elementary predication e. This inheritance may be used to distinguish argument constraints from modifier constraints, and to distinguish restrictor and nuclear scope arguments of generalized quantifiers, which allows cued association structures to be compiled into a logical form of expressions in typed lambda calculus (Schuler and Wheeler, 2014).

The sentence processing model described in this article operates on referential states for *signs* (de Saussure, 1916), which are elementary predications connected to signified referential states by cued associations distinguished for *signification*, notated here as $\mathbf{f_{sig}}$ functions from (sign) referential states to (signified) referential states. Predication types for these signs, here notated with variables α, β, γ, δ, and ε over domain S, may each contain a primitive clausal type τ or υ over domain T requiring zero or more syntactic arguments φ or ψ over domain $O \times S$, where each such argument may have a type-constructing operator (e.g. argument, modifier, conjunct, gap filler) in domain O followed by a sign type for the argument in domain S. A broad-coverage set of primitive clausal types and type-constructing operators for English is shown in Table 1.

The model described in this article assumes that cued association structures made of elementary predications are composed, stored and retrieved in associative memory according to operations of a left-corner parser (Aho and Ullman, 1972; Johnson-Laird, 1983; Abney and Johnson, 1991; Gibson, 1991; Resnik, 1992; Stabler, 1994; Lewis and Vasishth, 2005; van Schijndel et al., 2013) using a specific set of semantic processing functions R. These left-corner parser operations process sequences of observed word tokens of type $\omega, \omega', \omega''$, etc., by incrementally incorporating them into a cued association structure g. When adjacent words are not directly associated with each other, these cued association structures may consist of one or more sign fragments α/β, each a sign of type α lacking a sign of type β yet to come. For example, a sentence beginning with the words '*the very*,' may consist of a noun phrase lacking a common noun yet to come (for '*the*'), fol-

primitive clausal types in T		type-constructing operators in O
V finite verb	**T** top-level discourse	**-a** preceding argument
I infinitive	**S** top-level utterance	**-b** following argument
B base form	**Q** subject-auxiliary inverted	**-c** preceding conjunct
L participial	**C** complementized finite verb	**-d** following conjunct
A predicative	**F** complementized infinitive	**-g** gap filler
R adverbial	**E** complementized base form	**-h** heavy shift / extraposition
G gerund	**N** nominal clause / noun phrase	**-i** interrogative pronoun
P particle	**D** determiner / possessive	**-r** relative pronoun
	O non-possessive genitive	**-v** passive

Table 1: Primitive clausal types and type-constructing operators for English, adapted from Nguyen et al. (2012).

lowed by an adjective lacking an adjective yet to come (for '*very*'). Cued association structures that consist of multiple sign fragments can be represented as functions with arguments for the holes between these fragments. For example, a cued association structure with holes h', h between sign fragments $\alpha''/\beta'', \alpha'/\beta', \alpha/\beta$ can be represented as a function of type $\beta \rightarrow (\alpha \rightarrow \beta') \rightarrow (\alpha' \rightarrow \beta'') \rightarrow \alpha''$.

Sentence processing in this model starts with a top-level cued association structure (a function from syntactic type **T** to syntactic type **T**), followed by a sequence of word token units with types $\omega, \omega', \omega''$:

$$\lambda_{p:\mathbf{T}}\, \lambda_x\, (p\, x) : \mathbf{T} \;\cdot\; \mathbf{unit} : \omega \;\cdot\; \mathbf{unit} : \omega' \;\cdot\; \mathbf{unit} : \omega'' \cdots,$$

and proceeds by forking off and joining up sign fragments within this structure. At each word w, the sentence processing model considers whether to use that word to fork off a new complete sign fragment, using procedurally-learned lexical inference rules r to integrate semantic constraints from the word into the cued association structure. It may decide to fork, creating a new sign of type δ with a hole between δ and the bottom of the previous sign fragment β:

$$\frac{g : \beta \rightarrow \Gamma \;\cdot\; w : \omega}{(r\, g\, w) : (\delta \rightarrow \beta) \rightarrow \Delta}\; r : (\beta \rightarrow \Gamma) \rightarrow \omega \rightarrow (\delta \rightarrow \beta) \rightarrow \Delta \;\in\; R, \qquad \text{(+F)}$$

The sentence processing model may also decide not to fork, instead attaching word w at the bottom of the preceding sign fragment, using an identity function $(\lambda_p\, p)$ to fill in the hole between this new complete sign and the bottom of the previous sign:

$$\frac{g\!:\!\beta\rightarrow\Gamma\ \cdot\ w\!:\!\omega}{(r\,g\,w\,(\lambda_p\,p))\!:\!\Delta}\ r\!:(\beta\rightarrow\Gamma)\rightarrow\omega\rightarrow(\delta\rightarrow\beta)\rightarrow\Delta\ \in\ R, \tag{$-$F}$$

Lexical inference rules integrate semantic constraints from words of type ω into cued association structures g. For example, a lexical rule for a word of type **open** would define an elementary predication of type **BeingOpen** as the signified referential state of a sign x, with the first participant of that elementary predication as the first participant of x:

$$\lambda_{g:\beta\rightarrow\Gamma}\ \lambda_{w:\mathbf{open}}\ \lambda_{h:\mathbf{A\text{-}aN}\rightarrow\beta}$$
$$(g\circ h\,(\lambda_x\,(\mathbf{f_0}\circ\mathbf{f_{rin}}\circ\mathbf{f_{sig}}\,x)=\mathbf{BeingOpen},\,(\mathbf{f_1}\circ\mathbf{f_{rin}}\circ\mathbf{f_{sig}}\,x)=(\mathbf{f_1}\,x)))\!:\!\Gamma\ \in\ R$$

Each lexical inference rule requires the formation of only a small number of cued associations.

After each fork decision has been made, the sentence processing model considers whether to connect the complete sign fragment of type δ resulting from the previous fork decision to the bottom β of the previous disjoint incomplete sign fragment, using procedurally-learned grammatical inference rules r to compose left children of type δ and right children of type ε into parents of type γ. It may decide to join, using an identity function $(\lambda_p\,p)$ to fill in the hole between the new parent γ and the bottom of the previous sign:

$$\frac{g\!:(\delta\rightarrow\beta)\rightarrow\Gamma}{\lambda_{q:\varepsilon}\,(r\,g\,q\,(\lambda_p\,p))\!:\!\Delta}\ r\!:((\delta\rightarrow\beta)\rightarrow\Gamma)\rightarrow\varepsilon\rightarrow(\gamma\rightarrow\beta)\rightarrow\Delta\ \in\ R, \tag{+J}$$

The sentence processing model may also decide not to join, instead maintaining a separate sign fragment of type γ/ε with a hole between γ and the bottom β of the previous sign fragment:

$$\frac{g\!:(\delta\rightarrow\beta)\rightarrow\Gamma}{\lambda_{q:\varepsilon}\,(r\,g\,q)\!:(\gamma\rightarrow\beta)\rightarrow\Delta}\ r\!:((\delta\rightarrow\beta)\rightarrow\Gamma)\rightarrow\varepsilon\rightarrow(\gamma\rightarrow\beta)\rightarrow\Delta\ \in\ R, \tag{$-$J}$$

Finally, the sentence processing model can remove a non-local dependency which no longer appears in any sign fragment following it:

$$\frac{g\!:(\alpha\rightarrow\beta)\rightarrow(\alpha'\rightarrow\beta')\rightarrow(\alpha''\rightarrow\beta'')\rightarrow\ \dots\ \rightarrow\psi\rightarrow\Gamma}{\lambda_{h:\alpha\rightarrow\beta,\ h':\alpha'\rightarrow\beta',\ h'':\alpha''\rightarrow\beta'',\dots}\ \exists_z\,(g\,h\,h'\,h''\dots z)\!:\!\Gamma}\ \ \psi\notin\alpha,\beta,\alpha',\beta',\dots \tag{N}$$

A broad-coverage set of grammatical inference rules for English is shown in Tables 2 and 3. Each rule requires the formation of only a small number of cued associations. In the model described in this article, cued associations to older referential states incur integration cost as defined in the DLT.

$$\lambda_{g\,:\,(\gamma\psi_{1..\ell}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\tau\varphi_{1..n-1}\textbf{-a}\gamma\psi_{\ell+1..m}}\;\; \lambda_{h\,:\,\tau\varphi_{1..n-1}\psi_{1..m}\rightarrow\beta}$$
$$(g\,(\lambda_p\, h\,(\lambda_y\, \exists_x\,(p\,x), (q\,y), (\mathbf{f}_n\,y){=}(\mathbf{f}_{\mathbf{sig}}\,x)))) : \Gamma \;\in\; R \tag{Aa}$$

$$\lambda_{g\,:\,(\tau\varphi_{1..n-1}\textbf{-b}\gamma\psi_{1..\ell}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\gamma\psi_{\ell+1..m}}\;\; \lambda_{h\,:\,\tau\varphi_{1..n-1}\psi_{1..m}\rightarrow\beta}$$
$$(g\,(\lambda_p\, h\,(\lambda_x\, \exists_y\,(p\,x), (q\,y), (\mathbf{f}_n\,x){=}(\mathbf{f}_{\mathbf{sig}}\,y)))) : \Gamma \;\in\; R \tag{Ab}$$

$$\lambda_{g\,:\,(\tau\textbf{-a}\upsilon\psi_{1..\ell}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\gamma\psi_{\ell+1..m}}\;\; \lambda_{h\,:\,\gamma\psi_{1..m}\rightarrow\beta}$$
$$(g\,(\lambda_p\, h\,(\lambda_y\, \exists_x\,(p\,x), (q\,y), (\mathbf{f_1}\,x){=}(\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,y)))) : \Gamma \;\in\; R \tag{Ma}$$

$$\lambda_{g\,:\,(\gamma\psi_{1..\ell}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\tau\textbf{-a}\upsilon\psi_{\ell+1..m}}\;\; \lambda_{h\,:\,\gamma\psi_{1..m}\rightarrow\beta}$$
$$(g\,(\lambda_p\, h\,(\lambda_x\, \exists_y\,(p\,x), (q\,y), (\mathbf{f_1}\,y){=}(\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,x)))) : \Gamma \;\in\; R \tag{Mb}$$

$$\lambda_{g\,:\,(\gamma\psi_{1..\ell}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\tau\varphi_{1..n-1}\textbf{-a}\gamma\psi_{\ell+1..m}}\;\; \lambda_{h\,:\,\tau\varphi_{1..n-1}\psi_{1..m}\rightarrow\beta}\;\; (g\,(\lambda_p\, h\,(\lambda_x\, \exists_y\,(p\,x), (q\,y),$$
$$(\mathbf{f_1}\,x){=}(\mathbf{f_1}\,y), ..., (\mathbf{f}_{n-1}\,x){=}(\mathbf{f}_{n-1}\,y), (\mathbf{f}_n\,y){=}(\mathbf{f}_{\mathbf{sig}}\,x)))) : \Gamma \;\in\; R \tag{Ua}$$

$$\lambda_{g\,:\,(\tau\varphi_{1..n-1}\textbf{-b}\gamma\psi_{1..\ell}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\gamma\psi_{\ell+1..m}}\;\; \lambda_{h\,:\,\tau\varphi_{1..n-1}\psi_{1..m}\rightarrow\beta}\;\; (g\,(\lambda_p\, h\,(\lambda_y\, \exists_x\,(p\,x), (q\,y),$$
$$(\mathbf{f_1}\,x){=}(\mathbf{f_1}\,y), ..., (\mathbf{f}_{n-1}\,x){=}(\mathbf{f}_{n-1}\,y), (\mathbf{f}_n\,x){=}(\mathbf{f}_{\mathbf{sig}}\,y)))) : \Gamma \;\in\; R \tag{Ub}$$

$$\lambda_{g\,:\,(\tau\varphi_{1..n}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\gamma\textbf{-c}\tau\varphi_{1..n}}\;\; \lambda_{h\,:\,\gamma\rightarrow\beta}$$
$$(g\,(\lambda_p\, h\,(\lambda_x\, \exists_y\,(p\,x), (q\,y), (\mathbf{f_1}\,x){=}(\mathbf{f_1}\,y), ..., (\mathbf{f}_n\,x){=}(\mathbf{f}_n\,y),$$
$$(\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{cin}}{\circ}\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,y), (\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{cin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,y)))) : \Gamma \;\in\; R \tag{Ca}$$

$$\lambda_{g\,:\,(\tau\varphi_{1..n}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\gamma\textbf{-c}\tau\varphi_{1..n}}\;\; \lambda_{h\,:\,\gamma\textbf{-c}\tau\varphi_{1..n}\rightarrow\beta}$$
$$(g\,(\lambda_p\, h\,(\lambda_x\, \exists_y\,(p\,x), (q\,y), (\mathbf{f_1}\,x){=}(\mathbf{f_1}\,y), ..., (\mathbf{f}_n\,x){=}(\mathbf{f}_n\,y),$$
$$(\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{cin}}{\circ}\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,y), (\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{cin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,y)))) : \Gamma \;\in\; R \tag{Cb}$$

$$\lambda_{g\,:\,(\gamma\textbf{-d}\tau\varphi_{1..n}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\tau\varphi_{1..n}}\;\; \lambda_{h\,:\,\gamma\rightarrow\beta}$$
$$(g\,(\lambda_p\, h\,(\lambda_y\, \exists_x\,(p\,x), (q\,y), (\mathbf{f_1}\,x){=}(\mathbf{f_1}\,y), ..., (\mathbf{f}_n\,x){=}(\mathbf{f}_n\,y),$$
$$(\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,y){=}(\mathbf{f}_{\mathbf{cin}}{\circ}\mathbf{f}_{\mathbf{rin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,x), (\mathbf{f}_{\mathbf{sig}}\,y){=}(\mathbf{f}_{\mathbf{cin}}{\circ}\mathbf{f}_{\mathbf{sig}}\,x)))) : \Gamma \;\in\; R \tag{Cc}$$

$$\lambda_{g\,:\,(\delta\psi_{1..m}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\gamma\textbf{-g}\delta}\;\; \lambda_{h\,:\,\gamma\psi_{1..m}\rightarrow\beta}\;\; \lambda_{z\,:\,\textbf{-g}\delta}\;\; (g\,(\lambda_p\, h\,(\lambda_y\,(p\,z), (q\,y)))) : \Gamma \;\in\; R \tag{G}$$

$$\lambda_{g\,:\,(\gamma\textbf{-h}\delta\rightarrow\beta)\rightarrow(\alpha'\rightarrow\beta')\rightarrow(\alpha''\rightarrow\beta'')\rightarrow...\rightarrow\textbf{-h}\delta\rightarrow\Gamma}\;\; \lambda_{q\,:\,\varepsilon}\;\; \lambda_{h\,:\,\gamma\rightarrow\beta,\; h'\,:\,\alpha'\rightarrow\beta',\; h''\,:\,\alpha''\rightarrow\beta'',\,...}$$
$$\exists_z\,(g\,(\lambda_p\, h\,(\lambda_x\,(p\,x), (q\,z)))\, h'\, h''\, \ldots\, z) : \Gamma \;\in\; R \tag{H}$$

$$\lambda_{g\,:\,(\tau\varphi_{1..n-1}\textbf{-b}(\gamma\textbf{-i}\delta)\psi_{1..m}\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\gamma\textbf{-i}\delta}\;\; \lambda_{h\,:\,\tau\varphi_{1..n-1}\psi_{1..m}\rightarrow\beta}\;\; \lambda_{z\,:\,\textbf{-i}\delta}$$
$$(g\,(\lambda_p\, h\,(\lambda_x\, \exists_y\,(p\,x), (q\,y), (\mathbf{f}_n\,x){=}(\mathbf{f}_{\mathbf{sig}}\,z)))) : \Gamma \;\in\; R \tag{I}$$

$$\lambda_{g\,:\,(\gamma\rightarrow\beta)\rightarrow\Gamma}\;\; \lambda_{q\,:\,\delta\textbf{-r}\varepsilon}\;\; \lambda_{h\,:\,\gamma\rightarrow\beta}\;\; \lambda_{z\,:\,\textbf{-r}\varepsilon}\;\; (g\,(\lambda_p\, h\,(\lambda_y\,(p\,z), (q\,y)))) : \Gamma \;\in\; R \tag{R}$$

Table 2: Binary broad-coverage grammatical inference rules in R for argument attachment (Aa, Ab), modifier attachment (Ma, Mb), auxiliary attachment (Ua, Ub), conjunct attachment (Ca, Cb, Cc), gap filler attachment (G), extraposition or heavy shift attachment (H), interrogative pronoun antecedent attachment (I), and relative pronoun antecedent attachment (R), adapted from Nguyen et al. (2012).

$$\frac{\lambda_{g:(\tau\varphi_{1..n}\rightarrow\beta)\rightarrow(\alpha'\rightarrow\beta')\rightarrow(\alpha''\rightarrow\beta'')\rightarrow\ldots\rightarrow\psi\rightarrow\Gamma}\ \ \lambda_{h:\tau\varphi_{1..n-1}\psi\rightarrow\beta,\ h':\alpha'\rightarrow\beta',\ h'':\alpha'\rightarrow\beta'',\ldots,z:\psi}}{(g\,(\lambda_p\,h\,(\lambda_x\,\exists_y\,(p\,x),(\mathbf{f}_{\mathbf{ein}}\,y){=}z,(\mathbf{f}_n\,x){=}(\mathbf{f}_{\mathbf{sig}}\,y)))\,h'\,h''\ldots z):\Gamma\ \in\ R} \quad \text{(Ea)}$$

$$\frac{\lambda_{g:(\alpha\rightarrow\beta)\rightarrow(\alpha'\rightarrow\beta')\rightarrow(\alpha''\rightarrow\beta'')\rightarrow\ldots\rightarrow\psi\rightarrow\Gamma}\ \ \lambda_{h:\alpha\psi\rightarrow\beta,\ h':\alpha'\rightarrow\beta',\ h'':\alpha'\rightarrow\beta'',\ldots,z:\psi}}{(g\,(\lambda_p\,h\,(\lambda_x\,\exists_y\,(p\,x),(\mathbf{f}_{\mathbf{ein}}\,y){=}z,(\mathbf{f}_{\mathbf{1}}\,y){=}(\mathbf{f}_{\mathbf{rin}}\circ\mathbf{f}_{\mathbf{sig}}\,x)))\,h'\,h''\ldots z):\Gamma\ \in\ R} \quad \text{(Eb)}$$

$$\frac{\lambda_{g:(\gamma\varphi_1\varphi_2\psi_{1..m}\rightarrow\beta)\rightarrow\Gamma}\ \ \lambda_{h:\gamma\varphi_2\varphi_1\psi_{1..m}\rightarrow\beta}}{(g\,(\lambda_p\,h\,(\lambda_x\,\exists_y\,(p\,y),(\mathbf{f}_{\mathbf{2}}\,x){=}(\mathbf{f}_{\mathbf{1}}\,y),(\mathbf{f}_{\mathbf{1}}\,x){=}(\mathbf{f}_{\mathbf{2}}\,y),(\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{sig}}\,y)))):\Gamma\ \in\ R} \quad \text{(Q)}$$

$$\frac{\lambda_{g:(\tau\varphi_{1..n}\psi_{1..m}\rightarrow\beta)\rightarrow\ldots\rightarrow\psi_m\rightarrow\ldots\rightarrow\psi_1\rightarrow\Gamma}}{g:(\tau'\varphi'_{1..n}\psi'_{1..m}\rightarrow\beta)\rightarrow\ldots\rightarrow\psi'_m\rightarrow\ldots\rightarrow\psi_1\rightarrow\Gamma\ \in\ R} \quad \text{(T)}$$

$$\frac{\lambda_{g:(\mathbf{A\text{-}aN}\rightarrow\beta)\rightarrow\Gamma}\ \ \lambda_{h:\mathbf{L\text{-}aN\text{-}vN}\rightarrow\beta}\ \ \lambda_{z:\mathbf{\text{-}vN}}}{\exists_y\,(g\,(\lambda_p\,h\,(\lambda_x\,(p\,y),(\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{sig}}\,y),(\mathbf{f}_{\mathbf{1}}\,x){=}(\mathbf{f}_{\mathbf{sig}}\,z)))):\Gamma\ \in\ R} \quad \text{(V)}$$

$$\lambda_{g:(\mathbf{N}\rightarrow\beta)\rightarrow\Gamma}\ \ \lambda_{h:\mathbf{A\text{-}aN}\rightarrow\beta}\ \ (g\,(\lambda_p\,h\,(\lambda_x\,\exists_y\,(p\,y),(\mathbf{f}_{\mathbf{1}}\,x){=}(\mathbf{f}_{\mathbf{rin}}\circ\mathbf{f}_{\mathbf{sig}}\,y)))):\Gamma\ \in\ R \quad \text{(Za)}$$

$$\lambda_{g:(\mathbf{N}\rightarrow\beta)\rightarrow\Gamma}\ \ \lambda_{h:\tau\text{-}\mathbf{a}\upsilon\rightarrow\beta}\ \ (g\,(\lambda_p\,h\,(\lambda_x\,\exists_y\,(p\,y),(\mathbf{f}_{\mathbf{0}}\circ\mathbf{f}_{\mathbf{rin}}\circ\mathbf{f}_{\mathbf{sig}}\,x){=}\mathbf{BeingDuring},$$
$$(\mathbf{f}_{\mathbf{1}}\circ\mathbf{f}_{\mathbf{rin}}\circ\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{1}}\,x),(\mathbf{f}_{\mathbf{2}}\circ\mathbf{f}_{\mathbf{rin}}\circ\mathbf{f}_{\mathbf{sig}}\,x){=}(\mathbf{f}_{\mathbf{sig}}\,y)))):\Gamma\ \in\ R \quad \text{(Zb)}$$

$$\text{for } r:(\beta\rightarrow\Gamma)\rightarrow\omega\rightarrow(\delta'\rightarrow\beta)\rightarrow\Delta',\ \ r':((\delta'\rightarrow\beta)\rightarrow\Delta')\rightarrow(\delta\rightarrow\beta)\rightarrow\Delta\in R,$$
$$\lambda_{g:\beta\rightarrow\Gamma}\,r'\circ(r\,g):\omega\rightarrow(\delta\rightarrow\beta)\rightarrow\Delta\ \in\ R \quad (1)$$

$$\text{for } r:((\delta\rightarrow\beta)\rightarrow\Gamma)\rightarrow\varepsilon\rightarrow(\gamma'\rightarrow\beta)\rightarrow\Delta',\ \ r':((\gamma'\rightarrow\beta)\rightarrow\Delta')\rightarrow(\gamma\rightarrow\beta)\rightarrow\Delta\in R,$$
$$\lambda_{g:(\delta\rightarrow\beta)\rightarrow\Gamma}\,r'\circ(r\,g):\varepsilon\rightarrow(\gamma\rightarrow\beta)\rightarrow\Delta\ \in\ R \quad (2)$$

Table 3: Unary broad-coverage grammatical inference rules in *R* for non-local extraction (Ea, Eb), subject-auxiliary inversion (Q), type-changing (T), passive voice (V), and zero-head introduction (Za, Zb), adapted from Nguyen et al. (2012). These unary rules are combined with other lexical and grammatical inference rules using the recurrences in Equation 1 and 2.

References

Abney, S. P. and Johnson, M. (1991). Memory requirements and local ambiguities of parsing strategies. *J. Psycholinguistic Research*, 20(3):233–250.

Aho, A. V. and Ullman, J. D. (1972). *The Theory of Parsing, Translation and Compiling, Vol. 1: Parsing*. Prentice-Hall, Englewood Cliffs, New Jersey.

Anderson, J. A., Silverstein, J. W., Ritz, S. A., and Jones, R. S. (1977). Distinctive features, categorical perception and probability learning: Some applications of a neural model. *Psychological Review*, 84:413–451.

Baldridge, J. and Kruijff, G.-J. M. (2002). Coupling CCG and hybrid logic dependency semantics. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002)*, Philadelphia, Pennsylvania.

Copestake, A., Flickinger, D., Pollard, C., and Sag, I. (2005). Minimal recursion semantics: An introduction. *Research on Language and Computation*, pages 281–332.

de Saussure, F. (1916). *Cours de Linguistique Générale*. Payot.

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. PhD thesis, Carnegie Mellon.

Howard, M. W. and Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, 45:269–299.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press, Cambridge, MA, USA.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review*, 95(2):163–182.

Kruijff, G.-J. M. (2001). *A Categorial-Modal Architecture of Informativity: Dependency Grammar Logic and Information Structure*. PhD thesis, Charles University.

Lewis, R. L. and Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3):375–419.

Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society (London) B*, 262:23–81.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102:419–457.

Mel'čuk, I. (1988). *Dependency syntax: theory and practice*. State University of NY Press, Albany.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, 89:609–626.

Nguyen, L., van Schijndel, M., and Schuler, W. (2012). Accurate unbounded dependency recovery using generalized categorial grammars. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING '12)*, pages 2125–2140, Mumbai, India.

Resnik, P. (1992). Left-corner parsing and psychological plausibility. In *Proceedings of COLING*, pages 191–197, Nantes, France.

Schuler, W. and Wheeler, A. (2014). Cognitive compositional semantics using continuation dependencies. In *Third Joint Conference on Lexical and Computational Semantics (*SEM'14)*.

Stabler, E. (1994). The finite connectivity of linguistic structure. In *Perspectives on Sentence Processing*, pages 303–336. Lawrence Erlbaum.

van Schijndel, M., Exley, A., and Schuler, W. (2013). A model of language processing as hierarchic sequential prediction. *Topics in Cognitive Science*, 5(3):522–540.

White, M. (2006). Efficient realization of coordinate structures in combinatory categorial grammar. *Research on Language and Computation*, 4.