Incremental Semantic Dependency Parsing

Marten van Schijndel
Department of Linguistics
The Ohio State University

April 9, 2012

Introduction

The person who officials say ___ stole millions ...

Goal

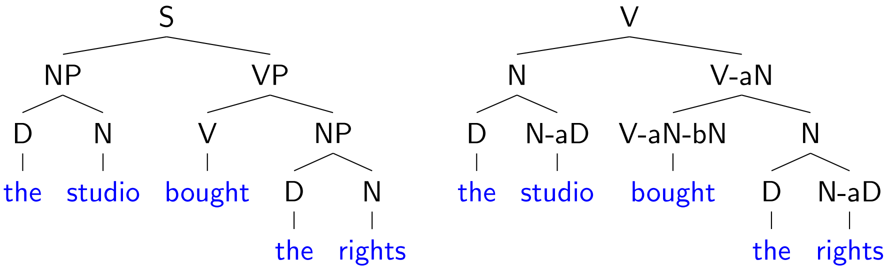
- Incrementally obtain correct parse despite filler-gap
- Test claims about human processing

Introduction

Previous studies [Gibson, 2000, Chen et al., 2005] have found filler-gap dependencies incur processing costs.
These studies conflate processing center embeddings with processing filler-gap constructions.

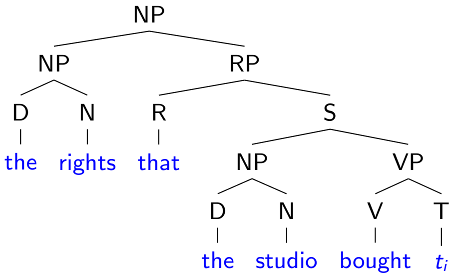
Training

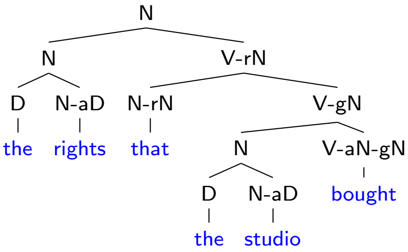
Reannotate WSJ [Nguyen et al., 2012]



Training

We can also keep the WSJ traces around.





Interpretation: Reannotation Rules

$$\frac{g{:}d \quad h{:}\, c\text{-}\mathbf{a}d}{(f_{c\text{-}\mathbf{ad}}\, g\, h){:}\, c} \quad \frac{g{:}d\psi \quad h{:}\, c\text{-}\mathbf{a}d}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, (g\, k)\, h){:}\, c\psi} \quad \frac{g{:}d \quad h{:}\, c\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, g\, (h\, k)){:}\, c\psi} \quad \frac{g{:}d\psi \quad h{:}\, c\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, (g\, k)\, (h\, k)){:}\, c\psi}$$

$$\frac{g{:}\, c\text{-}\mathbf{b}d \quad h{:}d}{(f_{c\text{-}\mathbf{bd}}\, g\, h){:}\, c} \quad \frac{g{:}\, c\text{-}\mathbf{b}d\psi \quad h{:}d}{\lambda_k\, (f_{c\text{-}\mathbf{bd}}\, (g\, k)\, h){:}\, c\psi} \quad \frac{g{:}\, c\text{-}\mathbf{b}d \quad h{:}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{bd}}\, g\, (h\, k)){:}\, c\psi} \quad \frac{g{:}\, c\text{-}\mathbf{b}d\psi \quad h{:}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{bd}}\, (g\, k)\, (h\, k)){:}\, c\psi} \tag{Aa–h}$$

$$\frac{g{:}\, u\text{-}\mathbf{a}d \quad h{:}c}{(f_{\mathrm{IM}}\, g\, h){:}c} \quad \frac{g{:}\, u\text{-}\mathbf{a}d\psi \quad h{:}c}{\lambda_k\, (f_{\mathrm{IM}}\, (g\, k)\, h){:}c\psi} \quad \frac{g{:}\, u\text{-}\mathbf{a}d \quad h{:}c\psi}{\lambda_k\, (f_{\mathrm{IM}}\, g\, (h\, k)){:}c\psi} \quad \frac{g{:}\, u\text{-}\mathbf{a}d\psi \quad h{:}c\psi}{\lambda_k\, (f_{\mathrm{IM}}\, (g\, k)\, (h\, k)){:}c\psi}$$

$$\frac{g{:}c \quad h{:}\, u\text{-}\mathbf{a}d}{(f_{\mathrm{FM}}\, g\, h){:}c} \quad \frac{g{:}c\psi \quad h{:}\, u\text{-}\mathbf{a}d}{\lambda_k\, (f_{\mathrm{FM}}\, (g\, k)\, h){:}c\psi} \quad \frac{g{:}c \quad h{:}\, u\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{\mathrm{FM}}\, g\, (h\, k)){:}c\psi} \quad \frac{g{:}c\psi \quad h{:}\, u\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{\mathrm{FM}}\, (g\, k)\, (h\, k)){:}c\psi} \tag{Ma–h}$$

$$\frac{g{:}\, c\text{-}\mathbf{a}d}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, \{k\}\, g){:}\, c\text{-}\mathbf{g}d} \quad \frac{g{:}\, c\text{-}\mathbf{b}d}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, \{k\}\, g){:}\, c\text{-}\mathbf{g}d} \quad \frac{g{:}c}{\lambda_k\, (f_{\mathrm{IM}}\, \{k\}\, g){:}c\text{-}\mathbf{g}d} \tag{Ga–c}$$

$$\frac{g{:}e \quad h{:}\, c\text{-}\mathbf{g}d}{\lambda_i\, \exists_j\, (g\, i) \wedge (h\, i\, j){:}e} \quad \frac{g{:}d\text{-}\mathbf{r}e \quad h{:}\, c\text{-}\mathbf{g}d}{\lambda_{k\,j}\, \exists_i\, (g\, k\, i) \wedge (h\, i\, j){:}\, c\text{-}\mathbf{r}e} \quad \frac{g{:}d\text{-}\mathbf{i}e \quad h{:}\, c\text{-}\mathbf{g}d}{\lambda_{k\,j}\, \exists_i\, (g\, k\, i) \wedge (h\, i\, j){:}\, c\text{-}\mathbf{i}e} \tag{Fa–c}$$

$$\frac{g{:}e \quad h{:}c\text{-}\mathbf{r}d}{\lambda_i\, \exists_j\, (g\, i) \wedge (h\, i\, j){:}e} \tag{R}$$

Interpretation: Reannotation Rules

$$\frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d}{(f_{c\text{-}\mathbf{ad}}\; g\; h){:}\,c} \qquad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\;k)\;h){:}\,c\psi} \qquad \frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\;g\;(h\;k)){:}\,c\psi} \qquad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\;k)\;(h\;k)){:}\,c\psi}$$

$$\frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d}{(f_{c\text{-}\mathbf{bd}}\; g\; h){:}\,c} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\;k)\;h){:}\,c\psi} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\;g\;(h\;k)){:}\,c\psi} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\;k)\;(h\;k)){:}\,c\psi} \tag{Aa–h}$$

$$\frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c}{(f_{\mathrm{IM}}\; g\; h){:}c} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,(g\;k)\;h){:}c\psi} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\;g\;(h\;k)){:}c\psi} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\,(g\;k)\;(h\;k)){:}c\psi}$$

$$\frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d}{(f_{\mathrm{FM}}\; g\; h){:}c} \qquad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d}{\lambda_k\,(f_{\mathrm{FM}}\,(g\;k)\;h){:}c\psi} \qquad \frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\;g\;(h\;k)){:}c\psi} \qquad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\,(g\;k)\;(h\;k)){:}c\psi} \tag{Ma–h}$$

$$\frac{g{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,\{k\}\;g){:}\,c\text{-}\mathbf{g}d} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,\{k\}\;g){:}\,c\text{-}\mathbf{g}d} \qquad \frac{g{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,\{k\}\;g){:}c\text{-}\mathbf{g}d} \tag{Ga–c}$$

$$\frac{g{:}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_i\,\exists_j\,(g\;i) \wedge (h\;i\;j){:}e} \qquad \frac{g{:}d\text{-}\mathbf{r}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\;k\;i) \wedge (h\;i\;j){:}\,c\text{-}\mathbf{r}e} \qquad \frac{g{:}d\text{-}\mathbf{i}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\;k\;i) \wedge (h\;i\;j){:}\,c\text{-}\mathbf{i}e} \tag{Fa–c}$$

$$\frac{g{:}e \quad h{:}c\text{-}\mathbf{r}d}{\lambda_i\,\exists_j\,(g\;i) \wedge (h\;i\;j){:}e} \tag{R}$$

Interpretation: Reannotation Rules

$$\frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d}{(f_{c\text{-}\mathbf{ad}}\,g\,h){:}\,c} \quad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\,k)\,h){:}\,c\psi} \quad \frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,g\,(h\,k)){:}\,c\psi} \quad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\,k)\,(h\,k)){:}\,c\psi}$$

$$\frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d}{(f_{c\text{-}\mathbf{bd}}\,g\,h){:}\,c} \quad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\,k)\,h){:}\,c\psi} \quad \frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,g\,(h\,k)){:}\,c\psi} \quad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\,k)\,(h\,k)){:}\,c\psi} \tag{Aa–h}$$

$$\frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c}{(f_{\mathrm{IM}}\,g\,h){:}c} \quad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,(g\,k)\,h){:}c\psi} \quad \frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\,g\,(h\,k)){:}c\psi} \quad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\,(g\,k)\,(h\,k)){:}c\psi}$$

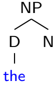
$$\frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d}{(f_{\mathrm{FM}}\,g\,h){:}c} \quad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d}{\lambda_k\,(f_{\mathrm{FM}}\,(g\,k)\,h){:}c\psi} \quad \frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\,g\,(h\,k)){:}c\psi} \quad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\,(g\,k)\,(h\,k)){:}c\psi} \tag{Ma–h}$$

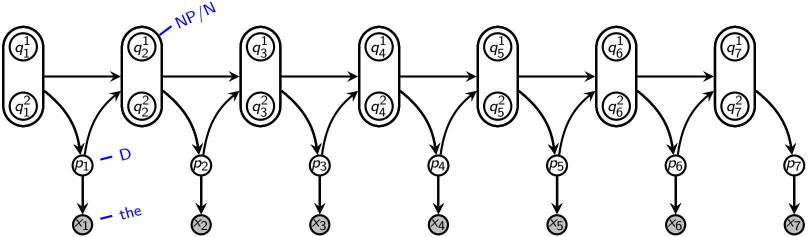
$$\frac{g{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,\{k\}\;g){:}\,c\text{-}\mathbf{g}d} \quad \frac{g{:}\,c\text{-}\mathbf{b}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,\{k\}\,g){:}\,c\text{-}\mathbf{g}d} \quad \frac{g{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,\{k\}\,g){:}c\text{-}\mathbf{g}d} \tag{Ga–c}$$

$$\frac{g{:}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_i\,\exists_j\,(g\,i) \wedge (h\,i\,j){:}e} \quad \frac{g{:}d\text{-}\mathbf{r}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\,k\,i) \wedge (h\,i\,j){:}\,c\text{-}\mathbf{r}e} \quad \frac{g{:}d\text{-}\mathbf{i}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\,k\,i) \wedge (h\,i\,j){:}\,c\text{-}\mathbf{i}e} \tag{Fa–c}$$

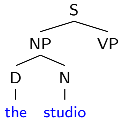
$$\frac{g{:}e \quad h{:}c\text{-}\mathbf{r}d}{\lambda_i\,\exists_j\,(g\,i) \wedge (h\,i\,j){:}e} \tag{R}$$

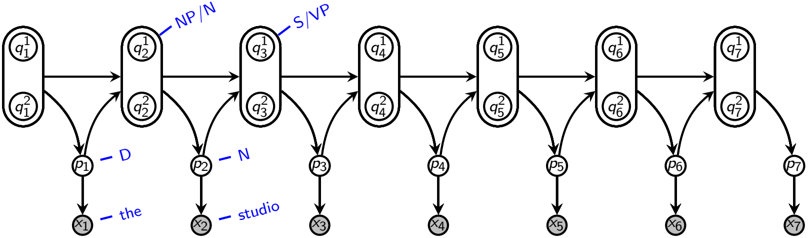
Practice Parse #1364



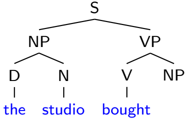


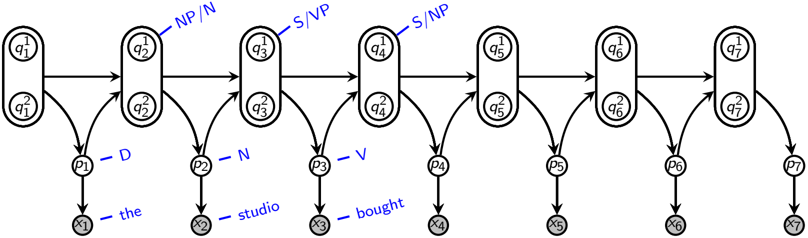
Practice Parse #1364



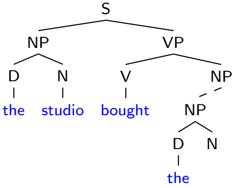


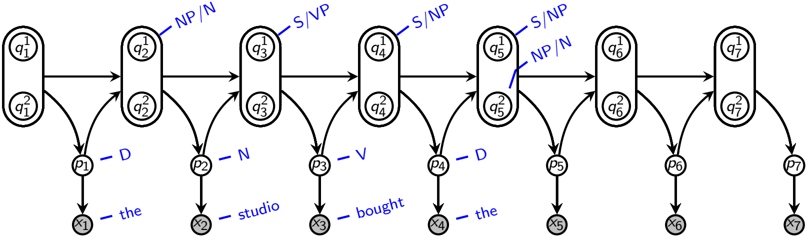
Practice Parse #1364



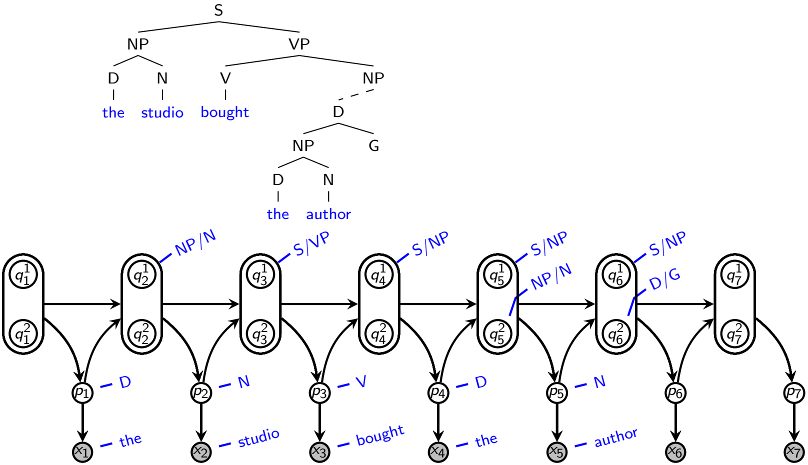


Practice Parse #1364

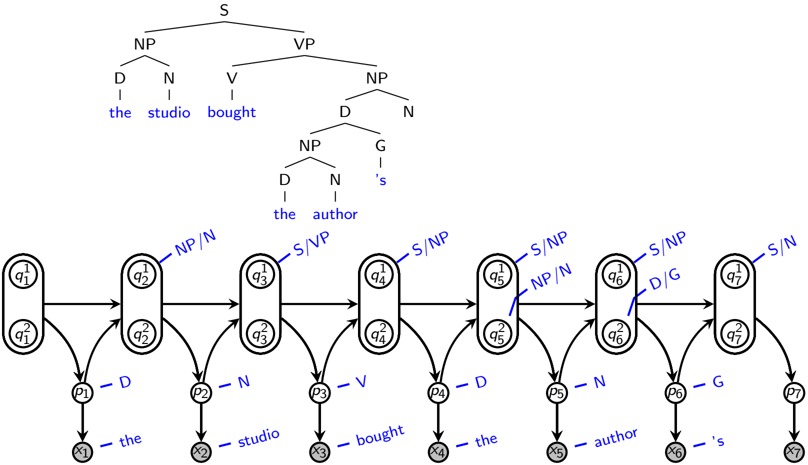




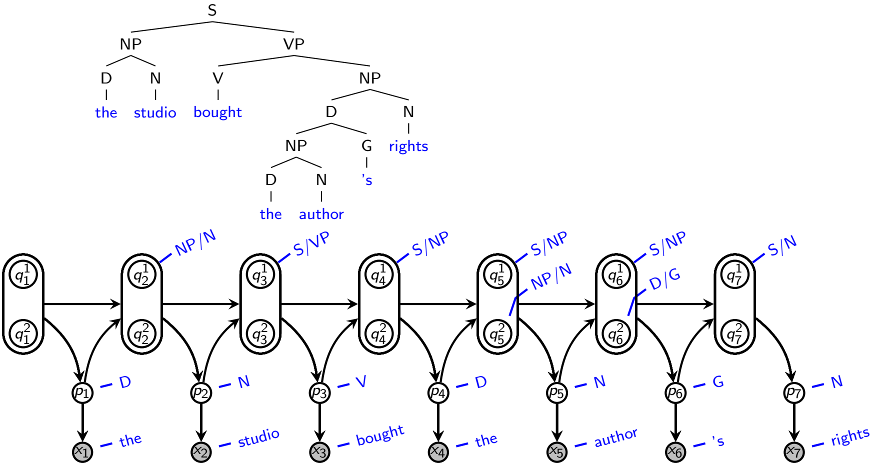
Practice Parse #1364



Practice Parse #1364

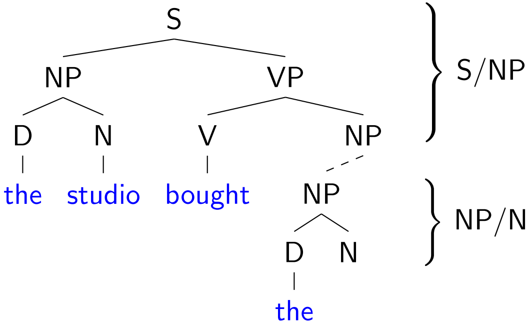


Practice Parse #1364



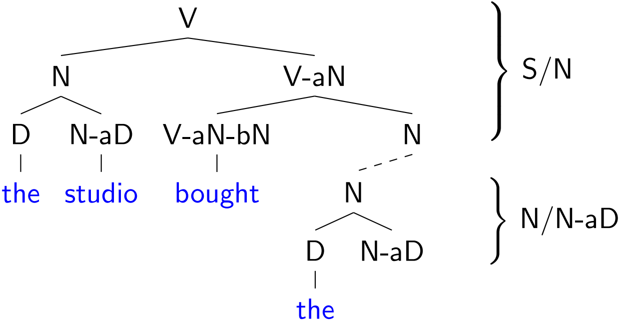
Interpretation

Connected Components

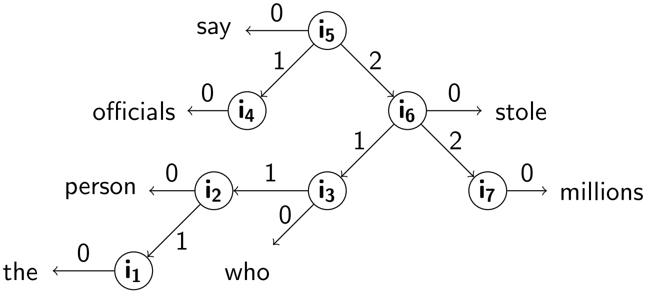


Interpretation

Reannotated Connected Components



Interpretation: Referent States



Interpretation: Fa/La

First or Last element of a CC

$$\frac{\exists_{i^1 j^1 .. \, i^\ell j^\ell} \; \ldots \; \wedge \; (g^\ell{:}c/d\,\{j^\ell\}\, i^\ell) \quad x_t}{\exists_{i^1 j^1 .. \, i^\ell} \; \ldots \; \wedge \; ((g^\ell f){:}c\, i^\ell)} \quad x_t \mapsto_M f{:}d \qquad (\text{–Fa})$$

$$\frac{\exists_{i^1 j^1 .. \, i^\ell j^\ell} \; \ldots \; \wedge \; (g^\ell{:}c/d\,\{j^\ell\}\, i^\ell) \quad x_t}{\exists_{i^1 j^1 .. \, i^\ell j^\ell i^{\ell+1}} \; \ldots \; \wedge \; (g^\ell{:}c/d\,\{j^\ell\}\, i^\ell) \; \wedge \; (f{:}e\, i^{\ell+1})} \quad x_t \mapsto_M f{:}e \qquad (+\text{Fa})$$

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1} i^\ell} \; \ldots \; \wedge \; (g^\ell{:}d\, i^\ell)}{\exists_{i^1 j^1 .. \, i^\ell j^\ell} \; \ldots \; \wedge \; ((f g^\ell){:}c/e\,\{j^\ell\}\, i^\ell)} \begin{cases} g{:}d \;\; h{:}e \Rightarrow (f\; g\; h){:}c \\ g{:}d \;\; h{:}e \Rightarrow \lambda_k(f\; (g\; k)\; h){:}c \\ g{:}d \;\; h{:}e \Rightarrow \lambda_k(f\; g\; (h\; k)){:}c \\ g{:}d \;\; h{:}e \Rightarrow \lambda_k(f\; (g\; k)\; (h\; k)){:}c \end{cases} \qquad (\text{–La})$$

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1} i^\ell} \; \ldots \; \wedge \; (g^{\ell-1}{:}a/c\,\{j^{\ell-1}\}\, i^{\ell-1}) \; \wedge \; (g^\ell{:}d\, i^\ell)}{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1}} \; \ldots \; \wedge \; (g^{\ell-1} \circ (f\; g^\ell){:}a/e\,\{j^{\ell-1}\}\, i^{\ell-1})} \begin{cases} g{:}d \;\; h{:}e \Rightarrow (f\; g\; h){:}c \\ g{:}d \;\; h{:}e \Rightarrow \lambda_k(f\; (g\; k)\; h){:}c \\ g{:}d \;\; h{:}e \Rightarrow \lambda_k(f\; g\; (h\; k)){:}c \\ g{:}d \;\; h{:}e \Rightarrow \lambda_k(f\; (g\; k)\; (h\; k)){:}c \end{cases} \qquad (+\text{La})$$

Interpretation

$$\frac{\exists_{i_1}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \qquad \text{the}}{\exists_{i_1\,i_3}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})\ \wedge\ (..:\mathbf{N/N\text{-}aD}\,\{\mathbf{i_3}\}\,\mathbf{i_3}) \qquad \text{person}}\ {\scriptstyle +Fa,-La,-N}$$

$$\frac{}{\exists_{i_1\,i_3}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})\ \wedge\ (..:\mathbf{N/V\text{-}rN}\,\{\mathbf{i_3}\}\,\mathbf{i_3}) \qquad \text{who}}\ {\scriptstyle -Fa,-La,-N}$$

$$\frac{}{\exists_{i_1\,i_3\,i_6}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})\ \wedge\ (..:\mathbf{N/V\text{-}gN}\,\{\mathbf{i_6}\}\,\mathbf{i_3}) \qquad \text{officials}}\ {\scriptstyle +Fa,+Lc,-N}$$

$$\frac{}{\exists_{i_1\,i_3\,i_6\,i_9}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})\ \wedge\ (..:\mathbf{N/V\text{-}gN}\,\{\mathbf{i_6}\}\,\mathbf{i_3})\ \wedge\ (..:\mathbf{V\text{-}gN/V\text{-}aN\text{-}gN}\,\{\mathbf{i_9}\}\,\mathbf{i_9}) \qquad \text{say}}\ {\scriptstyle +Fa,-La,-N}$$

$$\frac{}{\exists_{i_1\,i_3\,i_{11}}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})\ \wedge\ (..:\mathbf{N/V\text{-}aN}\,\{\mathbf{i_{11}}\}\,\mathbf{i_3}) \qquad \text{stole}}\ {\scriptstyle +Fb,+La,+N}$$

$$\frac{}{\exists_{i_1\,i_3\,i_{13}}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})\ \wedge\ (..:\mathbf{N/N}\,\{\mathbf{i_{13}}\}\,\mathbf{i_3}) \qquad \text{millions}}\ {\scriptstyle +Fa,+La,-N}$$

$$\frac{}{\exists_{i_1}\ (..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})}\ {\scriptstyle -Fa,+La,-N}$$

Interpretation: Fb/Lb

$$\psi \in \{\text{-r,-i}\} \times \text{C}$$

$$\frac{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^\ell j^\ell} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge (g^\ell{:}c/d\,\{j^\ell\}\,i^\ell) \quad x_t}{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^\ell} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge ((g^\ell(f'\{j^n\}f)){:}c\,i^\ell)}$$

$$x_t \mapsto_M \lambda_k(f'\{k\}\,f){:}d \quad (\text{–Fb})$$

$$\frac{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^\ell j^\ell} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge (g^\ell{:}c/d\,\{j^\ell\}\,i^\ell) \quad x_t}{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^\ell j^\ell i^{\ell+1}} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge (g^\ell{:}c/d\,\{j^\ell\}\,i^\ell) \wedge ((f'\{j^n\}f){:}e\,i^{\ell+1})}$$

$$x_t \mapsto_M \lambda_k(f'\{k\}\,f){:}e \quad (+\text{Fb})$$

$$\frac{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^{\ell-1} j^{\ell-1} i^\ell} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge (g^\ell{:}d\,i^\ell)}{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^\ell j^\ell} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge ((fg^\ell) \circ (f'\{j^n\}){:}c\psi/e\,\{j^\ell\}\,i^\ell)}$$

$$g{:}d\ \ h{:}e \Rightarrow \lambda_k(fg\,(f'\{k\}\,h)){:}c\psi \quad (\text{–Lb})$$

$$\frac{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^{\ell-1} j^{\ell-1} i^\ell} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge (g^{\ell-1}{:}a/c\psi\,\{j^{\ell-1}\}\,i^{\ell-1}) \wedge (g^\ell{:}d\,i^\ell)}{\exists_{i^1 j^1 ..\ i^n j^n ..\ i^{\ell-1} j^{\ell-1}} \ldots \wedge (g^n{:}y/z\psi\,\{j^n\}\,i^n) \wedge \ldots \wedge (g^{\ell-1} \circ (fg^\ell) \circ (f'\{j^n\}){:}a/e\,\{j^{\ell-1}\}\,i^{\ell-1})}$$

$$g{:}d\ \ h{:}e \Rightarrow \lambda_k(fg\,(f'\{k\}\,h)){:}c\psi \quad (+\text{Lb})$$

Interpretation: Lc/N

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1} i^{\ell}} \; \dots \; \wedge \; (g^{\ell}{:}d \; i^{\ell})}{\exists_{i^1 j^1 .. \, i^{\ell} j^{\ell}} \; \dots \; \wedge \; ((fg^{\ell}) \circ (\lambda_{h\,k\,i} \, (h\,k)){:}a/e\psi \, \{j^{\ell}\} \, i^{\ell})} \quad g{:}d \;\; h{:}e\psi \Rightarrow (fg \; h){:}c \tag{$-$Lc}$$

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1} i^{\ell}} \; \dots \; \wedge \; (g^{\ell-1}{:}a/c \, \{j^{\ell-1}\} \, i^{\ell-1}) \; \wedge \; (g^{\ell}{:}d \; i^{\ell})}{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1}} \; \dots \; \wedge \; (g^{\ell-1} \circ (fg^{\ell}) \circ (\lambda_{h\,k\,i} \, (h\,k)){:}a/e\psi \, \{j^{\ell-1}\} \, i^{\ell-1})} \quad g{:}d \;\; h{:}e\psi \Rightarrow (fg \; h){:}c \tag{+Lc}$$

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell} j^{\ell}} \; \dots \; \wedge \; (g^{\ell-1}{:}c/d\psi \, \{j^{\ell-1}\} \, i^{\ell-1}) \; \wedge \; (g^{\ell}{:}d\psi/e \, \{j^{\ell}\} \, i^{\ell})}{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1}} \; \dots \; \wedge \; (g^{\ell-1} \circ (\lambda_{h\,i} \exists_j (h\,j)) \circ g^{\ell}{:}c/e \, \{j^{\ell-1}\} \, i^{\ell-1})} \tag{+N}$$

All of these rules may be made probabilistic

Interpretation

$$\exists_{i_1}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \quad \text{the}$$

$$+\text{Fa},-\text{La},-\text{N}$$

$$\exists_{i_1\,i_3}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \;\wedge\; (..:\mathbf{N/N\text{-}aD}\,\{\mathbf{i_3}\}\,\mathbf{i_3}) \quad \text{person}$$

$$-\text{Fa},-\text{La},-\text{N}$$

$$\exists_{i_1\,i_3}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \;\wedge\; (..:\mathbf{N/V\text{-}rN}\,\{\mathbf{i_3}\}\,\mathbf{i_3}) \quad \text{who}$$

$$+\text{Fa},+\text{Lc},-\text{N}$$

$$\exists_{i_1\,i_3\,i_6}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \;\wedge\; (..:\mathbf{N/V\text{-}gN}\,\{\mathbf{i_6}\}\,\mathbf{i_3}) \quad \text{officials}$$

$$+\text{Fa},-\text{La},-\text{N}$$

$$\exists_{i_1\,i_3\,i_6\,i_9}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \;\wedge\; (..:\mathbf{N/V\text{-}gN}\,\{\mathbf{i_6}\}\,\mathbf{i_3}) \;\wedge\; (..:\mathbf{V\text{-}gN/V\text{-}aN\text{-}gN}\,\{\mathbf{i_9}\}\,\mathbf{i_9}) \quad \text{say}$$

$$+\text{Fb},+\text{La},+\text{N}$$

$$\exists_{i_1\,i_3\,i_{11}}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \;\wedge\; (..:\mathbf{N/V\text{-}aN}\,\{\mathbf{i_{11}}\}\,\mathbf{i_3}) \quad \text{stole}$$

$$+\text{Fa},+\text{La},-\text{N}$$

$$\exists_{i_1\,i_3\,i_{13}}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1}) \;\wedge\; (..:\mathbf{N/N}\,\{\mathbf{i_{13}}\}\,\mathbf{i_3}) \quad \text{millions}$$

$$-\text{Fa},+\text{La},-\text{N}$$

$$\exists_{i_1}\,(..:\mathbf{T/T}\,\{\mathbf{i_1}\}\,\mathbf{i_1})$$

Evaluation: Syntactic vs Semantic

Syntactic Parser [van Schijndel et al., ress]

- Only Fa/La
- Trained on WSJ 02-21
- Split-merged $\times 5$ [Petrov et al., 2006]

Semantic Parser

- Trained on Reannotated WSJ 02-21
- Split-merged $\times 3$

Evaluation: Syntactic vs Semantic

Test Corpus: Dundee

- Log-transformed go-past durations
- Omit:
  - first and last of each line (wrap-up)
  - <5 times in WSJ (accuracy) [Fossum and Levy, 2012]
  - saccade length > 4 (track loss) [Demberg and Keller, 2008]

Evaluation: Syntactic vs Semantic

Linear Mixed Effects Baseline

- Number of characters
- Previous (next) word fixated?
- Unigram and Bigram probs
- Sentence position
- Total Surprisal [Hale, 2001]
- Number of intervening words
- Cum. Total Surprisal
- Cum. Entropy Reduction [Hale, 2003]
- Joint interactions
- Spillover Predictors

Factors are residualized off next simpler model
Subject and Item random intercepts were included

Evaluation: Syntactic vs Semantic

AIC and p-values

AIC can be used to test null hypothesis [Burnham and Anderson, 2002]

$\exp((- |\mathrm{AIC}_A - \mathrm{AIC}_B|)/2)$

Model	log-likelihood	AIC
syntactic	-74312	148874
semantic	-74314	148877
syntactic+(F–L+)	-74281	148816
semantic+(F–L+)	**-74277**	**148809**

Goodness-of-fits

Upper rows: Baseline comparison ($p = .14$)

Lower rows: With integration (F–L+) cost ($p = .03$)

Not shown: With encoding (F+L–) cost ($p = .14$)

Evaluation: Semantic Factor Correlations

Factor	t-score	p-value
F+L– (encoding)	4.17	$3.05 \cdot 10^{-05}$
F–L+ (integration)	-8.16	$3.38 \cdot 10^{-16}$

Significance of residualized factors on reading time.

Positive t-score: inhibition

Negative t-score: facilitation

Conclusion

Results

- Described incremental semantic dependency parser
- General metrics are not hurt by semantic calculation
- Semantic metrics predict reading times better than syntactic
- Replicated negative integration cost without FG confound

Fin

Thanks to Elliot Schumacher (and viewers like you)!
Questions?

Extras 1: Probabilistic Formulae

$$\mathsf{P}_{\phi_\ell}(\text{`}-\text{'}\ r^F \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \mathsf{E}_{\gamma^*_\ell}(c \xrightarrow{0} d\ \ldots) \cdot \textstyle\sum_x \mathsf{P}_\gamma(d \rightarrow x) \cdot [\![ r^F = \langle i, \text{`id'}, j\rangle]\!] \tag{1a}$$

$$\mathsf{P}_{\phi_\ell}(\text{`}+\text{'}\ r^F \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \mathsf{E}_{\gamma^*_\ell}(c \xrightarrow{+} d\ \ldots) \cdot \textstyle\sum_x \mathsf{P}_\gamma(d \rightarrow x) \cdot [\![ r^F = \langle \text{`}-\text{'}, \text{`}-\text{'}, \text{`}-\text{'}\rangle]\!] \tag{1b}$$

$$\mathsf{P}_{\lambda_\ell}(\text{`}+\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \textstyle\sum_{c',e} \mathsf{E}_{\gamma^*_\ell}(c \xrightarrow{0} c'\ \ldots) \cdot \mathsf{P}_{\gamma_{B,\ell}}(c' \rightarrow d\ e) \tag{2a}$$

$$\mathsf{P}_{\lambda_\ell}(\text{`}-\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \textstyle\sum_{c',e} \mathsf{E}_{\gamma^*_\ell}(c \xrightarrow{+} c'\ \ldots) \cdot \mathsf{P}_{\gamma_{A,\ell}}(c' \rightarrow d\ e) \tag{2b}$$

$$\mathsf{P}_{\nu_\ell}(\text{`}+\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle\ \langle j',d'\rangle\ \langle k,e\rangle) \overset{\text{def}}{=} [\![c,d,d' \in C \times \{\textbf{-g}\} \times C \ \wedge\ e \notin C \times \{\textbf{-g}\} \times C]\!] \tag{3a}$$

$$\mathsf{P}_{\nu_\ell}(\text{`}-\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle\ \langle j',d'\rangle\ \langle k,e\rangle) \overset{\text{def}}{=} [\![c,d,d' \notin C \times \{\textbf{-g}\} \times C \ \vee\ e \in C \times \{\textbf{-g}\} \times C]\!] \tag{3b}$$

Extras 1: Probabilistic Formulae

$$\mathsf{P}_{\alpha_\ell}(\langle i',c'\rangle\ r^A \mid l\,\langle i,c\rangle\,\langle j,d\rangle) \overset{\text{def}}{\propto} \begin{cases}\text{if } l = \text{'+'} : \sum_e \mathsf{E}_{\gamma^*_\ell}(c \xrightarrow{+} c'\ \ldots)\cdot \mathsf{P}_{\gamma_{A,\ell}}(c' \to d\ e) \\ \text{if } l = \text{'–'}\ : [\![c' = d]\!]\end{cases}$$
$$\cdot \begin{cases}\text{if } l = \text{'+'} \vee [d \ldots \Rightarrow c'] \in \text{Ae–h,Me–h} : [\![i' = j]\!] \\ \text{if } l = \text{'–'} \wedge [d \ldots \Rightarrow c'] \in \text{Aa–d,Ma–d}\ : [\![i' = \mathbf{i}_{\mathbf{Z}+1}]\!]\end{cases}$$
$$\cdot \begin{cases}\text{if } l = \text{'+'} : [\![r^A = \langle i', \text{'id'}, j\rangle]\!] \\ \text{if } l = \text{'–'}\ : [\![r^A = \langle \text{'–'}, \text{'–'}, \text{'–'}\rangle]\!]\end{cases} \tag{1}$$

$$\mathsf{P}_{\beta_{s,\ell}}(\langle k,e\rangle\ r^B \mid \langle i,c\rangle\,\langle j,d\rangle) \overset{\text{def}}{\propto} \mathsf{P}_{\gamma_{s,\ell}}(c \to d\ e)\cdot \begin{cases}\text{if } [d\,e \Rightarrow c] \in \text{Aa–d,Ma–d} : [\![k = i]\!] \\ \text{if } [d\,e \Rightarrow c] \in \text{Ae–h,Me–h}\ : [\![k = \mathbf{i}_{\mathbf{Z}+1}]\!]\end{cases}$$
$$\cdot \begin{cases}\text{if } [d\,e \Rightarrow c] \in \text{Aa–d,Me–h} : [\![r^B = \langle k, V(e), j\rangle]\!] \\ \text{if } [d\,e \Rightarrow c] \in \text{Ae–h,Ma–d} : [\![r^B = \langle j, V(d), k\rangle]\!] \\ \text{if } [d\,e \Rightarrow c] \in \text{Fa–c} \quad : [\![r^B = \langle \text{'–'}, \text{'–'}, \text{'–'}\rangle]\!]\end{cases} \tag{2}$$

$$\mathsf{P}_{\kappa_\ell}(r^K \mid \langle i,c\rangle\,\langle i',c'\rangle\,\langle j,d\rangle\,\langle k,e\rangle) \overset{\text{def}}{=}$$
$$\begin{cases}\text{if } c \in C\times\{\textbf{-g}\}\times C \ \wedge\ \exists_{d'}\ d'\ e \Rightarrow c' \ \wedge\ [d \Rightarrow d'] \in \text{Ga–b} : [\![r^K = \langle i', V(d), i\rangle]\!] \\ \text{if } c \in C\times\{\textbf{-g}\}\times C \ \wedge\ \exists_{e'}\ d\ e' \Rightarrow c' \ \wedge\ [e \Rightarrow e'] \in \text{Ga–b} : [\![r^K = \langle i', V(e), i\rangle]\!] \\ \text{if } c \in C\times\{\textbf{-g}\}\times C \ \wedge\ \exists_{d'}\ d'\ e \Rightarrow c' \ \wedge\ [d \Rightarrow d'] \in \text{Gc} : [\![r^K = \langle i, 1, i'\rangle]\!] \\ \text{if } c \in C\times\{\textbf{-g}\}\times C \ \wedge\ \exists_{e'}\ d\ e' \Rightarrow c' \ \wedge\ [e \Rightarrow e'] \in \text{Gc} : [\![r^K = \langle i, 1, i'\rangle]\!] \\ \text{otherwise} : [\![r^K = \langle \text{'–'}, \text{'–'}, \text{'–'}\rangle]\!]\end{cases} \tag{3}$$

Extras 1: Probabilistic Formulae

$$\begin{aligned}
\mathsf{P}_{\sigma}(q_t^{1..N} \mid q_{t-1}^{1..N} x_{t-1}) &\stackrel{\text{def}}{=} \mathsf{P}_{\phi_\ell}(\text{'--'} \mid b_{t-1}^{\ell}\; x_{t-1}) \cdot \mathsf{P}_{\sigma'_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N}\; a_{t-1}^{\ell}) \\
&\quad + \mathsf{P}_{\phi_\ell}(\text{'+'} \mid b_{t-1}^{\ell}\; x_{t-1}) \cdot \mathsf{P}_{\sigma'_{\ell+1}}(q_t^{1..N} \mid q_{t-1}^{1..N}\; x_{t-1}); \quad \ell \stackrel{\text{def}}{=} \max\{\ell' \mid q_{t-1}^{\ell'} \neq \text{'--'}\} \qquad (4) \\
\mathsf{P}_{\sigma'_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a') &\stackrel{\text{def}}{=} \mathsf{P}_{\lambda_\ell}(\text{'+'} \mid b_{t-1}^{\ell-1}\; a') \cdot [\![a = a_{t-1}^{\ell-1}]\!] \cdot \mathsf{P}_{\beta_{\mathrm{B},\ell-1}}(b \mid b_{t-1}^{\ell-1}\; a') \cdot \mathsf{P}_{\sigma''_{\ell-1}}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b\, a') \\
&\quad + \mathsf{P}_{\lambda_\ell}(\text{'--'} \mid b_{t-1}^{\ell-1}\; a') \cdot \mathsf{P}_{\alpha_\ell}(a \mid b_{t-1}^{\ell-1}\; a') \cdot \mathsf{P}_{\beta_{\mathrm{A},\ell}}(b \mid a_t^{\ell}\; a') \cdot \mathsf{P}_{\sigma''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b\, a') \qquad (5) \\
\mathsf{P}_{\sigma''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b\, a') &\stackrel{\text{def}}{=} \mathsf{P}_{\nu_{\ell-1}}(\text{'+'} \mid a_{t-1}^{\ell-1}\; b_{t-1}^{\ell-1}\; a_{t-1}^{\ell}\; b) \cdot \mathsf{P}_{\kappa_{\ell-1}}(r^K \mid b_{t-1}^{n}\; b_{t-1}^{\ell}\; a'\; b) \cdot \mathsf{P}_{\sigma'''_{\ell-1}}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b) \\
&\quad + \mathsf{P}_{\nu_{\ell-1}}(\text{'--'} \mid a_{t-1}^{\ell-1}\; b_{t-1}^{\ell-1}\; a_{t-1}^{\ell}\; b) \cdot \mathsf{P}_{\kappa_{\ell-1}}(r^K \mid b_{t-1}^{n}\; b_{t-1}^{\ell}\; a'\; b) \cdot \mathsf{P}_{\sigma'''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b) \qquad (6) \\
\mathsf{P}_{\sigma'''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b) &\stackrel{\text{def}}{=} [\![q_t^{1..\ell-1} = q_{t-1}^{1..\ell-1}]\!] \cdot [\![a_t^{\ell} = a]\!] \cdot [\![b_t^{\ell} = b]\!] \cdot [\![q_t^{\ell+1..N} = \text{'--'}]\!] \qquad (7)
\end{aligned}$$

Bibliography I

Burnham, K. P. and Anderson, D. R. (2002).
Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.
Springer-Verlag, 2nd edition.

Chen, E., Gibson, E., and Wolf, F. (2005).
Online syntactic storage costs in sentence comprehension.
Journal of Memory and Language, 52(1):144–169.

Demberg, V. and Keller, F. (2008).
Data from eye-tracking corpora as evidence for theories of syntactic processing complexity.
Cognition, 109(2):193–210.

Bibliography II

Fossum, V. and Levy, R. (2012).
Sequential vs. hierarchical syntactic models of human incremental sentence processing.
In *Proceedings of CMCL-NAACL 2012*. Association for Computational Linguistics.

Gibson, E. (2000).
The dependency locality theory: A distance-based theory of linguistic complexity.
In *Image, language, brain: Papers from the first mind articulation project symposium*, pages 95–126.

Hale, J. (2001).
A probabilistic earley parser as a psycholinguistic model.
In *Proceedings of the second meeting of the North American chapter of the Association for Computational Linguistics*, pages 159–166, Pittsburgh, PA.

Bibliography III

Hale, J. (2003).
Grammar, Uncertainty and Sentence Processing.
PhD thesis, Cognitive Science, The Johns Hopkins University.

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)*, Mumbai, India.

Petrov, S., Barrett, L., Thibaux, R., and Klein, D. (2006).
Learning accurate, compact, and interpretable tree annotation.
In *Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics (COLING/ACL'06)*.

Bibliography IV

van Schijndel, M., Exley, A., and Schuler, W. (in press).
A model of language processing as hierarchic sequential prediction.
Topics in Cognitive Science.