

-2,-1,-0,+0,+1,+2 — word position identifiers (the features are used to determine the relation between the words at -0 and +0)

$\alpha_x, \beta_x, \rho_x, \xi_x, \epsilon_x$  — x's lowercase text, word cluster, part-of-speech, rightmost child, and leftmost child, respectively.

$\varrho_x$  — if  $\rho_x$  ends with \$:  $\rho_x$  else: first two letters of  $\rho_x$

$\kappa_x, \omega_x$  — x's original index in the sentence, first three characters (lowercase), respectively

$\tau_x$  — if  $\rho \in \{\text{IN, WRB, WP}\}$ :  $\alpha_x + \rho_x$  else  $\rho_x$

$\delta_x$  — dependency links from x; left/right direction is included; *prep*, and *punct* links include word form; *nsubj* includes form if it is 'it'; *advmod* links include word cluster; additional indicator if x has a relative pronoun *nsubj*, *nsubj*→*prep*→*pobj*, *dobj*, *iobj*, or *prep*→*pobj* or as a *poss* of these positions.

$\gamma_x$  — target of dependency arc x

$\lambda_x$  —  $\kappa_{\xi_x} - \kappa_{\epsilon_x}$

$\eta_{x,y}$  — token y places to the right of token x in original token order

$\sigma_x, \Delta_x$  — determiner (or possessor) of x, indicator if x is either possessive or a definite determiner

for  $\mathbf{x} \in \{-2, -1, -0, +0, +1, +2\}$

$\tau_x, \varrho_x, \alpha_x, \beta_x, \xi_x \neq \text{null} \ \& \ \epsilon_x \neq \text{null}, \tau_x \wedge \lambda_x, \rho_x \in \{\text{VBD, VBN}\}, \rho_x \in \{\text{NNS, VBZ}\}, \rho_x \in \{\text{NN, VB}\}, \alpha_x \in \{\text{the, a, an}\}$

for  $\mathbf{x}$  where  $\text{MAX}(\kappa_{-0}\kappa_{+0} - 10) < \kappa_x < \kappa_{+0}$ :  $\tau_x$

for  $\mathbf{x}$  where  $\text{MAX}(\kappa_{-0}, \kappa_{+0} - 10) < \kappa_x < \kappa_{+0} \ \& \ \alpha_x \in \{“; : - ( )\}$ :  $\alpha_x$

for  $\mathbf{x} \in \{-1, -0, +0, +1\}$ : if  $\rho_x == \text{CC}$ :  $\xi_x \neq \text{null} \ \& \ \epsilon_x \neq \text{null}, \alpha_x \wedge \rho_{\epsilon_x} \wedge \rho_{\xi_x}$

for  $\mathbf{x} \in \{-2, -1, -0, +0, +1, +2\}$ : for  $\mathbf{y} \in \delta_x$ :  $\mathbf{y}$

if  $\rho_{-1} == \text{CC}$

$\tau_{-2} \wedge \tau_{-0} \wedge \tau_{+0}, \tau_{-2} \wedge \tau_{-0} \wedge \tau_{+0} \wedge \rho_{\xi_{-2}} \wedge \rho_{\epsilon_{-2}}$

for  $\mathbf{x} \in \{-0, +0\}$ :  $(\rho_x == \rho_{-2}) \wedge \tau_{+0}, (\varrho_x == \varrho_{-2}) \wedge \tau_{+0}, (\alpha_x == \alpha_{-2}) \wedge \tau_{+0}$

if  $\rho_{-0} == \text{CC}$

$\rho_{-2} \wedge \rho_{-1} \wedge \rho_{+0}, \varrho_{-2} \wedge \varrho_{+0}$

for  $\mathbf{x} \in \{+0, +1, +2\}$ : for  $\mathbf{y} \in \{-2, -1\}$ :  $\alpha_x == \alpha_y, \rho_x == \rho_y, \varrho_x == \varrho_y$

if  $\rho_{+0} == \text{CC}$

$\rho_{\xi_{+0}} == \rho_{-0}, \varrho_{\xi_{+0}} == \varrho_{-0}, \alpha_{-0} == \alpha_{\xi_{+0}}, \alpha_{-1} == \alpha_{\xi_{+0}},$

$\sigma_{-0} == \sigma_{\xi_{+0}}, \sigma_{-2} == \sigma_{\xi_{+0}}, \sigma_{-2} == \sigma_{\xi_{+0}} \ \& \ \sigma_{-0} \neq \sigma_{\xi_{+0}}, \Delta_{\sigma_{-2}} \wedge \Delta_{\sigma_{-0}} \wedge \Delta_{\sigma_{\xi_{+0}}}$

if  $\rho_{+1} == \text{CC}$

$\alpha_{-0} == \alpha_{+2}, \alpha_{-0} == \alpha_{\xi_{+1}}, \alpha_{\xi_{+1}} == \alpha_{-1}, \rho_{\xi_{+1}} == \rho_{-0}, \varrho_{\xi_{+1}} == \varrho_{-0}, \alpha_{+0} == \alpha_{+2}, \alpha_{+0} == \alpha_{\xi_{+1}},$

$\tau_{-0} \wedge \tau_{+0} \wedge \rho_{\xi_{+1}} \wedge \tau_{+2}, \rho_{-0} \wedge \rho_{+0} \wedge \rho_{+2}$

if  $\rho_{+2} == \text{CC}$

if  $\xi_{+2} \neq \text{null}$ : for  $\mathbf{x} \in \{-0, +0, +1\}$ :  $\alpha_{\xi_{+2}} = \alpha_x, \rho_{\xi_{+2}} = \rho_x, \varrho_{\xi_{+2}} = \varrho_x$

else  $\xi_{+2} == \text{null}$

for  $\{\mathbf{x}, \mathbf{k}\} \in \{\{+0, \{-0, -1, -2\}\}, \{+1, \{+0, -0, -1\}\}, \{+2, \{+0, -0\}\}\}$

if  $(\rho_x == \text{CC})$ : for  $\mathbf{x} \in \delta_{\xi_x}$ : for  $\mathbf{y} \in \mathbf{k}$ : for  $\mathbf{z} \in \delta_y$ :  $\mathbf{x} \wedge \mathbf{z}$

for  $\{\mathbf{x}, \mathbf{y}\} \in \{\{+0, -0\}, \{+1, +0\}\}$

for  $\mathbf{z} \in \delta_x$ : if  $\mathbf{z} == \text{whadvmod}$ :  $\alpha_{\gamma_z} \wedge \alpha_y, \alpha_{\gamma_z} \wedge \tau_y, \alpha_{\gamma_z} \wedge \beta_y, \alpha_y, \tau_y, \beta_y$

for  $\{\mathbf{x}, \mathbf{y}\} \in \{\{-1, -0\}, \{-0, +0\}, \{+0, +1\}\}$

for  $\mathbf{a} \in \delta_x$ : for  $\mathbf{b} \in \delta_y$ :  $\mathbf{a} \wedge \mathbf{b}$

for  $\mathbf{x} \in \delta_{-0}$ :  $\tau_{+0} \wedge \mathbf{x}, \beta_{+0} \wedge \mathbf{x}$

for  $\mathbf{x} \in \delta_{+0}$ :  $\tau_{-0} \wedge \mathbf{x}, \beta_{-0} \wedge \mathbf{x}$

for  $\mathbf{x} \in \{-0, +0\}$ : for  $\mathbf{y} \in \delta_x$

$\tau_x \wedge \mathbf{y}, \beta_x \wedge \mathbf{y}$

for  $\mathbf{z} \in \delta_x$ : if  $\mathbf{y} \neq \mathbf{z}$ :  $\mathbf{y} \wedge \mathbf{z}$

$\tau_{-1} \wedge \tau_{+1}, \tau_{-0} \wedge \tau_{+0} \wedge \beta_{\xi_{+0}}, \tau_{-0} \wedge \tau_{+0} \wedge \beta_{\xi_{-0}}, \eta_{+0}, \eta_{+0}, -1 \wedge \eta_{+0}, -0$

for  $\{\mathbf{x}, \mathbf{y}\} \in \{\{-1, +0\}, \{-1, -0\}, \{-0, +0\}, \{+0, +1\}, \{-0, +1\}\}$

$\tau_x \wedge \tau_y, \alpha_x \wedge \alpha_y, \alpha_x \wedge \tau_y, \tau_x \wedge \alpha_y, \beta_x \wedge \beta_y, \tau_x \wedge \tau_y \wedge \rho_{\epsilon_x} \wedge \rho_{\epsilon_y}, \tau_x \wedge \tau_y \wedge \rho_{\epsilon_x} \wedge \rho_{\xi_y}, \tau_x \wedge \tau_y \wedge \rho_{\xi_x} \wedge \rho_{\epsilon_y}, \tau_x \wedge \tau_y \wedge \rho_{\xi_x} \wedge \rho_{\xi_y}$

if  $\varrho_{-2} \neq \text{VB} \ \& \ \varrho_{-1} \neq \text{VB} \ \& \ \varrho_{+1} \neq \text{VB} \ \& \ \varrho_{+2} \neq \text{VB}$ :  $\varrho_{-0} \neq \text{VB} \ \& \ \varrho_{+0} == \text{VB}, \varrho_{+0} \neq \text{VB} \ \& \ \varrho_{-0} == \text{VB}$

if  $(-1 == \text{null} \ \& \ +1 == \text{null})$ :  $\tau_2 \wedge \tau_{+0}, \beta_2 \wedge \beta_{+0}$

if  $\rho_{-0} == \text{IN}$ :  $\alpha_{-1} \wedge \alpha_{-0} \wedge \rho_{\xi_{-0}}, \tau_{-1} \wedge \alpha_{-0} \wedge \alpha_{\xi_{-0}}, \beta_{-1} \wedge \beta_{-0} \wedge \beta_{\xi_{-0}}, \tau_{-1} \wedge \beta_{-0} \wedge \beta_{\xi_{-0}}$

if  $\rho_{+0} == \text{IN}$

for  $\mathbf{x} \in \{-2, -1\}$ :  $\omega_x \wedge \alpha_{+0} \wedge \rho_{\xi_{+0}}, \tau_x \wedge \alpha_{+0} \wedge \alpha_{\xi_{+0}}$

$\omega_{-0} \wedge \alpha_{+0} \wedge \rho_{\xi_{+0}}, \tau_{-0} \wedge \alpha_{+0} \wedge \alpha_{\xi_{+0}}, \beta_{-0} \wedge \beta_{+0} \wedge \beta_{\xi_{+0}}, \tau_{-0} \wedge \beta_{+0} \wedge \beta_{\xi_{+0}}$

if  $\xi_{+0} \neq \text{null}$ :  $\omega_{-1} \wedge \varrho_{-0} \wedge \alpha_{+0}, \omega_{-1} \wedge \tau_{-0} \wedge \alpha_{+0}, \omega_{-0} \wedge \alpha_{+0} \wedge \tau_{+1}, \xi_{+0} \neq \text{null}, \omega_{-2} \wedge \alpha_{+0}, \omega_{-1} \wedge \alpha_{+0}, \omega_{-0} \wedge \alpha_{+0},$

$\tau_{-2} \wedge \alpha_{+0}, \tau_{-1} \wedge \alpha_{+0}, \tau_{-0} \wedge \alpha_{+0}$

if  $\rho_{+1} == \text{IN}$ :  $\omega_{+0} \wedge \alpha_{+1} \wedge \rho_{\xi_{+1}}, \omega_{-0} \wedge \alpha_{+1} \wedge \rho_{\xi_{+1}}, \tau_{+0} \wedge \alpha_{+1} \wedge \alpha_{\xi_{+1}}, \tau_{-0} \wedge \alpha_{+1} \wedge \alpha_{\xi_{+1}}, \beta_{-0} \wedge \beta_{+1}, \tau_{-0} \wedge \beta_{+1},$

$\beta_{-0} \wedge \beta_{+1} \wedge \beta_{\xi_{+1}}, \tau_{-0} \wedge \beta_{+1} \wedge \beta_{\xi_{+1}}$

for  $\{\mathbf{x}, \mathbf{y}\} \in \{\{-2, -1\}, \{-1, -0\}, \{-0, +0\}, \{+0, +1\}, \{+1, +2\}\}$ :  $|\kappa_x - \kappa_y| > 1, (|\kappa_x - \kappa_y| > 1) \wedge \tau_x \wedge \tau_y$

$(\tau_{-0} \in \{\text{JJ, NN, NNS, NNP, NNPS}\}) \wedge (\exists \mathbf{x} \in \delta_{-0}: \alpha_{\gamma_x} == \text{'too'} \vee \alpha_{\gamma_x} == \text{'enough'})$

for  $\mathbf{x} \in \{-1, -2\}, \{+1, +2\}$ : for  $\mathbf{y} \in \mathbf{x}$ :  $\beta_y, \varrho_y$

Figure 1: Feature templates. If the variable is in bold, then the value of the variable is considered to be part of the identifier for any features produced by the templates utilizing that variable (only variables holding word indices may appear in bold).