Learning: Hidden-Variable Perceptron

passes over dataset

for each entity pair *i*

most likely sentence labels and inferred facts (ignoring DB facts)

most likely sentence labels given DB facts

initialize parameter vector $\Theta \leftarrow \mathbf{0}$

for $t = 1...T$ **do**

    for $i = 1...n$ **do**

        $(\mathbf{y}', \mathbf{z}') \leftarrow \arg\max_{\mathbf{y},\mathbf{z}} p(\mathbf{y}, \mathbf{z}|\mathbf{x_i}; \theta)$

        if $\mathbf{y}' \neq \mathbf{y_i}$ **then**

        $\mathbf{z}^* \leftarrow \arg\max_{\mathbf{z}} p(\mathbf{z}|\mathbf{x_i}, \mathbf{y_i}; \theta)$

        $\Theta \leftarrow \Theta + \phi(\mathbf{x_i}, \mathbf{z}^*) - \phi(\mathbf{x_i}, \mathbf{z}')$

        end if

    end for

end for

Return Θ