Neural Coreference Resolution Models

Natalie Parde
UIC CS 421

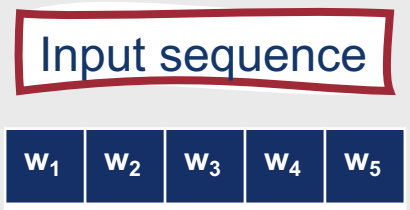
Neural Classification Models

- Generally end-to-end systems
- May not have a separate mention detection step
  - Instead, consider every possible text span of length < *k* as a possible mention
- Same overall goal as usual:
  - Assign to each span *i* an antecedent y_i ranging over the values $Y(i) = \{1, \ldots, i-1, \varepsilon\}$

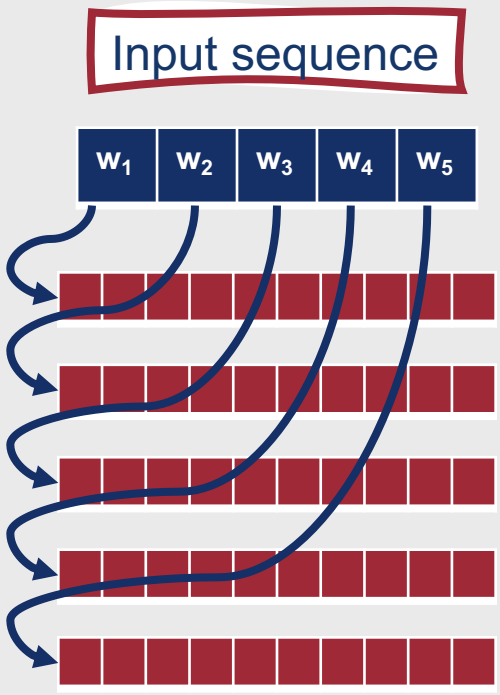
What goes on behind the scenes?

- For each pair of spans *i* and *j*, the system assigns a score $s(i,j)$ for the coreference link between the two
  - $s(i,j) = m(i) + m(j) + c(i,j)$
    - $m(i)$: Whether span *i* is a mention
    - $m(j)$: Whether span *j* is a mention
    - $c(i,j)$: Whether *j* is the antecedent of *i*
- The functions $m(\cdot)$ and $c(\cdot,\cdot)$ are computed using neural models:
  - $m(i) = w_m \cdot FFNN_m(g_i)$
  - $c(i,j) = w_c \cdot FFNN_c([g_i, g_j, g_i \circ g_j, \phi(i,j)])$
    - Where g_i is a vector representation of span *i*, and $\phi(i,j)$ encodes manually-defined characteristics of the relationship between *i* and *j*
    - Note that the exact definition of $c(i,j)$ may differ across models!

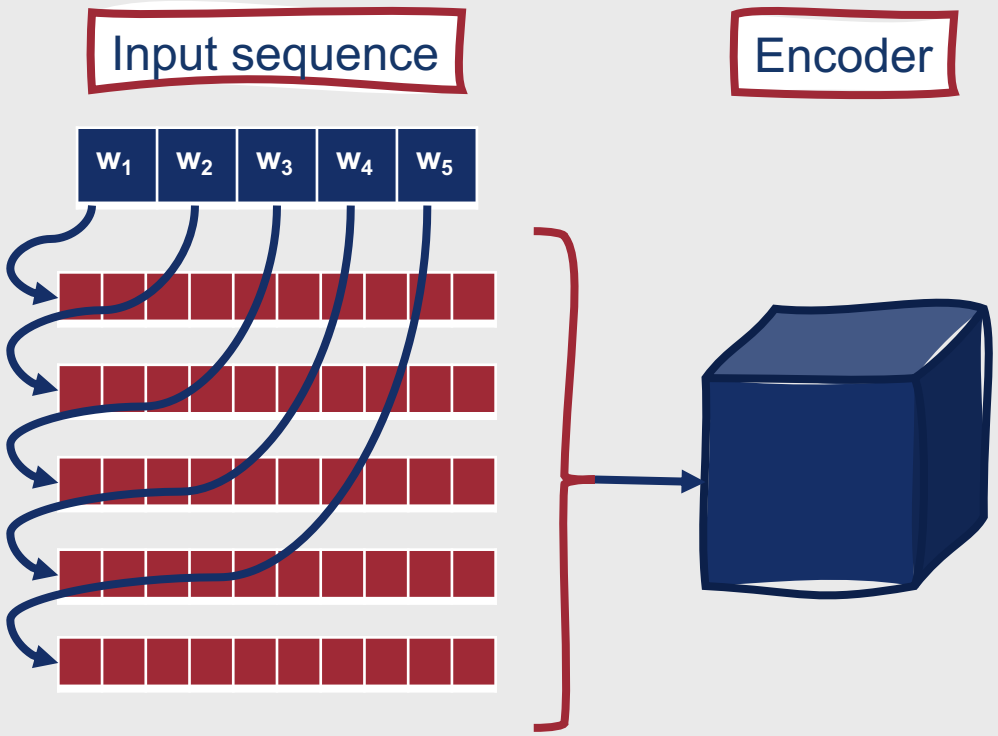
Altogether, a neural coreference resolution model might look like the following….



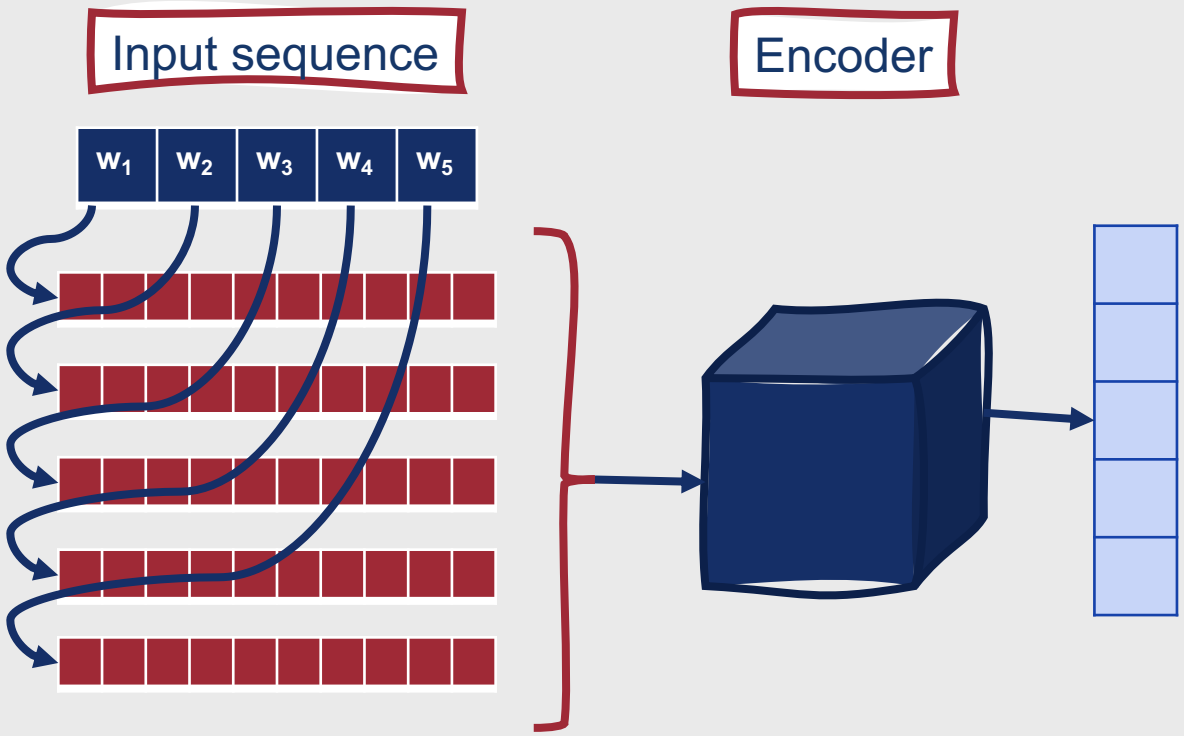
Altogether, a neural coreference resolution model might look like the following….



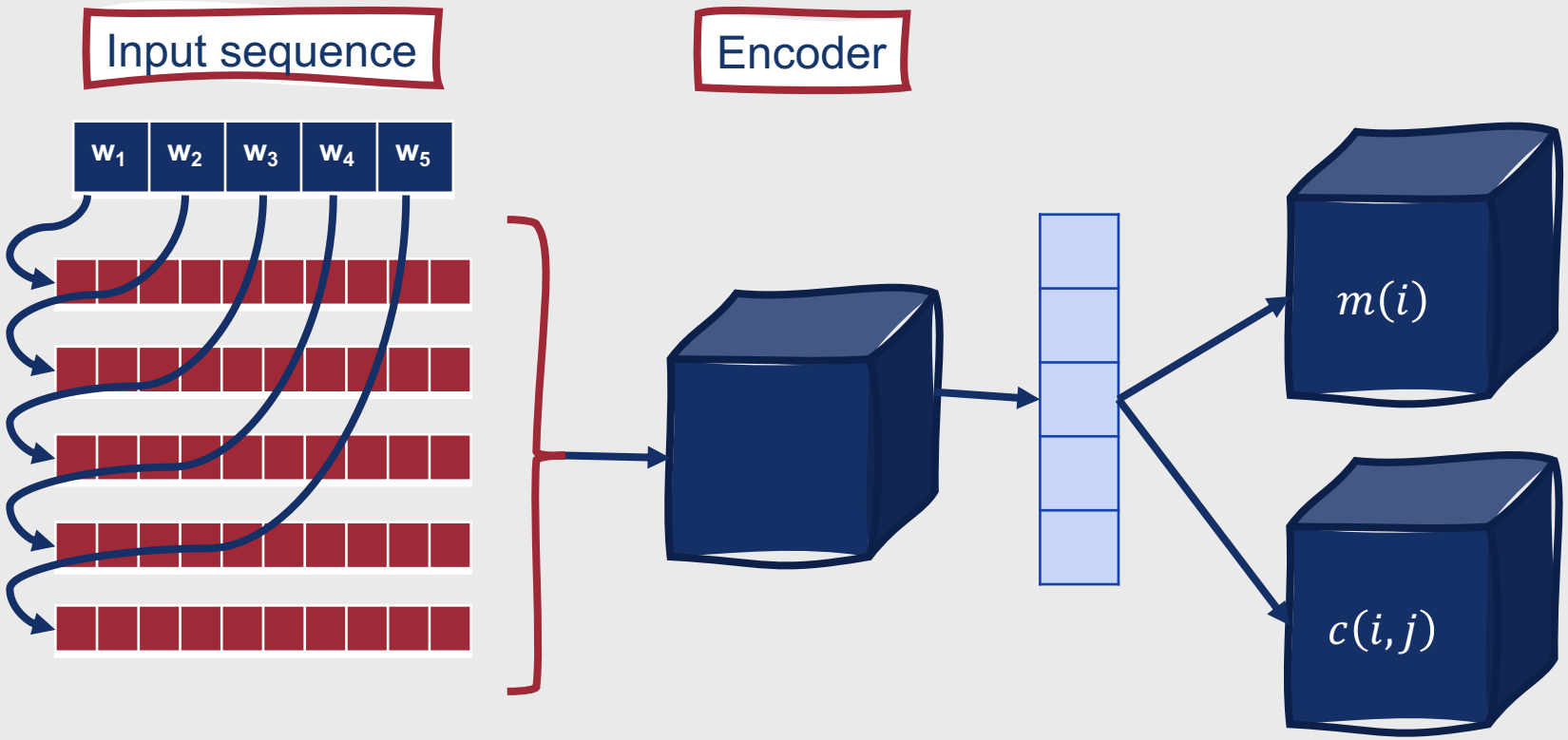
Altogether, a neural coreference resolution model might look like the following….



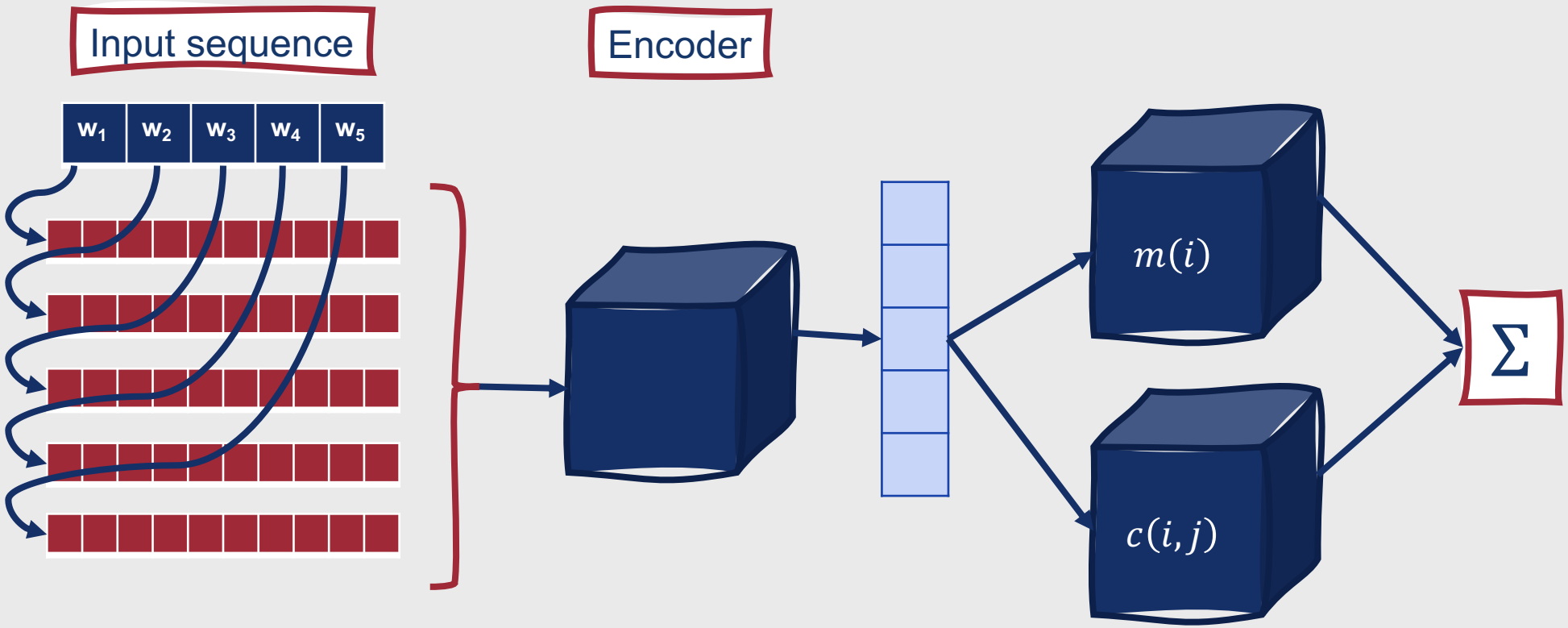
Altogether, a neural coreference resolution model might look like the following….



Altogether, a neural coreference resolution model might look like the following….



Altogether, a neural coreference resolution model might look like the following….



Altogether, a neural coreference resolution model might look like the following….