

Notation

The following notation is used throughout this work.

a a scalar (real, integer, or word (word's are treated as in.))

Words are treated as integers

For reasons that will become clear in ??, we will consistently notate words, as if they were scalar integer values. writing for example $w^{(1)}$ as to be the first word in a sequence.

Superscripts are never exponents

Just to reiterate, we use $x^{(t)}$ not to represent $\prod_{i=1}^{i=t} x$, but as a variable name. The only exponential that occurs in this work is the natural exponential, which we write $\exp x$.

Readers may be more familiar with subscripts being used to allow more variable names e.g. x_t . However, we use this for indexing: that would be the t th element of x .

In the case of recurrent neural networks, we often will want to reference both the i element of the t th time-step variable: $x^{(t)}$. This is thus unambiguously written written $x_i^{(t)}$.