Bigram language model

So that's what we get for n = 2:

$$p(\mathbf{w}) = \cancel{p(w_1)} p(w_2|w_1) \ldots p(w_k|w_{k-1})$$

$$p(w_1|start)$$

Toy corpus:

This is the malt
That lay in the house that Jack built.

$$\underset{}{p(\text{this is the house})} = \underset{1/2}{p(\text{this})}\ \underset{1}{p(\text{is}|\ \text{this})}\ \underset{1}{p(\text{the}|\ \text{is})}\ \underset{1/2}{p(\text{house}|\ \text{the})}$$