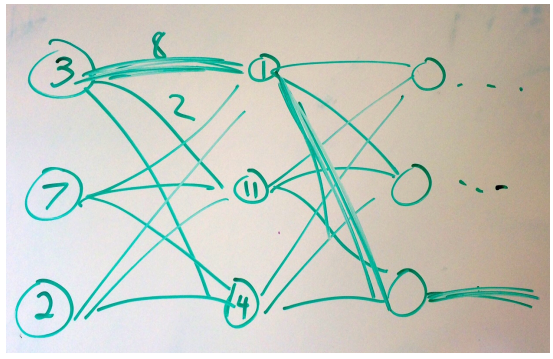
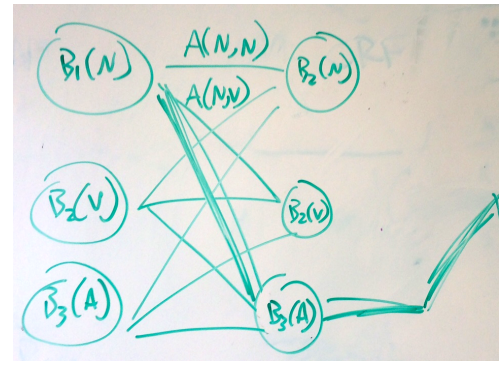
CS585 10/8/15 stuff

Office hours:
alternate explanation of Viterbi and calc_factor_scores.

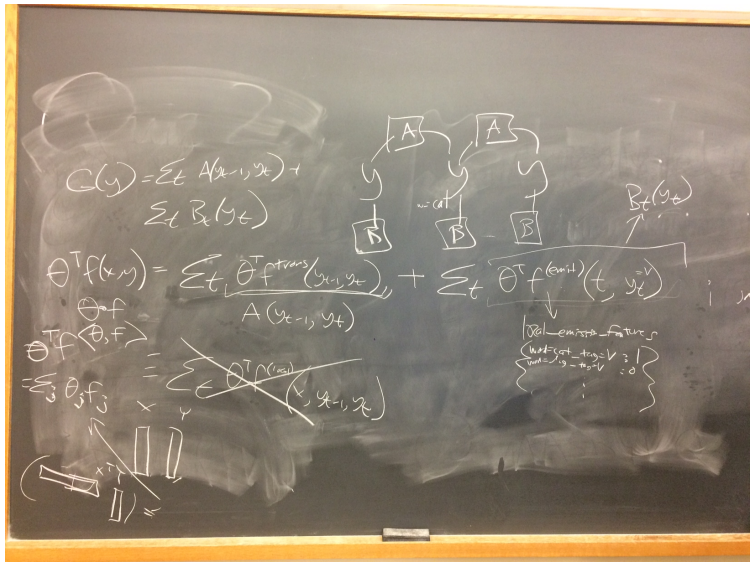
LEFT: Think of Viterbi as an algorithm that takes in as input a graph where there are weights on each node and edge. The graph has a lattice structure. It returns the highest-scoring path through that lattice. (where path score = sum of all node and edge weights on path.)

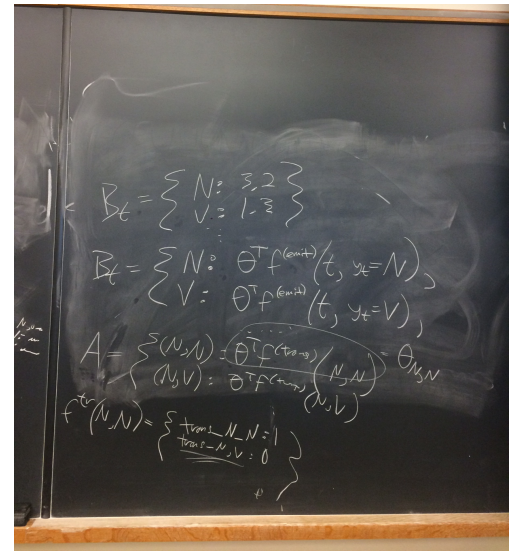
RIGHT: calc_factor_scores builds this graph, represented as (Ascores,Bscores). The B_t scores are node weights and A scores are edge weights (same pattern of edge weights at each timestep).

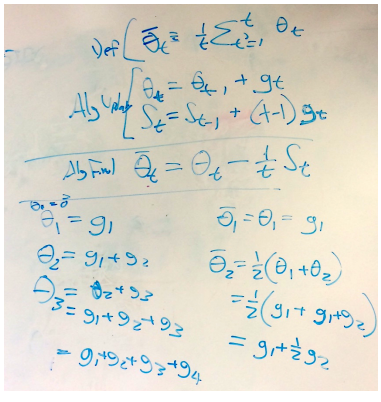




Blackboard at the end of 10/8/15 lecture: defining local factor scores from local CRF features







Office hours:
starting Problem 3, averaged perc. weightsums trick.

Some intuition:
bar{theta} overcounts the gradients near the start.
they get summed in for all the later thetas.
so theta undercounts early gradients relative to bar{theta}.
thus we make S, which highly weights gradients near the end.
subtracting S out of theta reduces weights of gradients near the end.