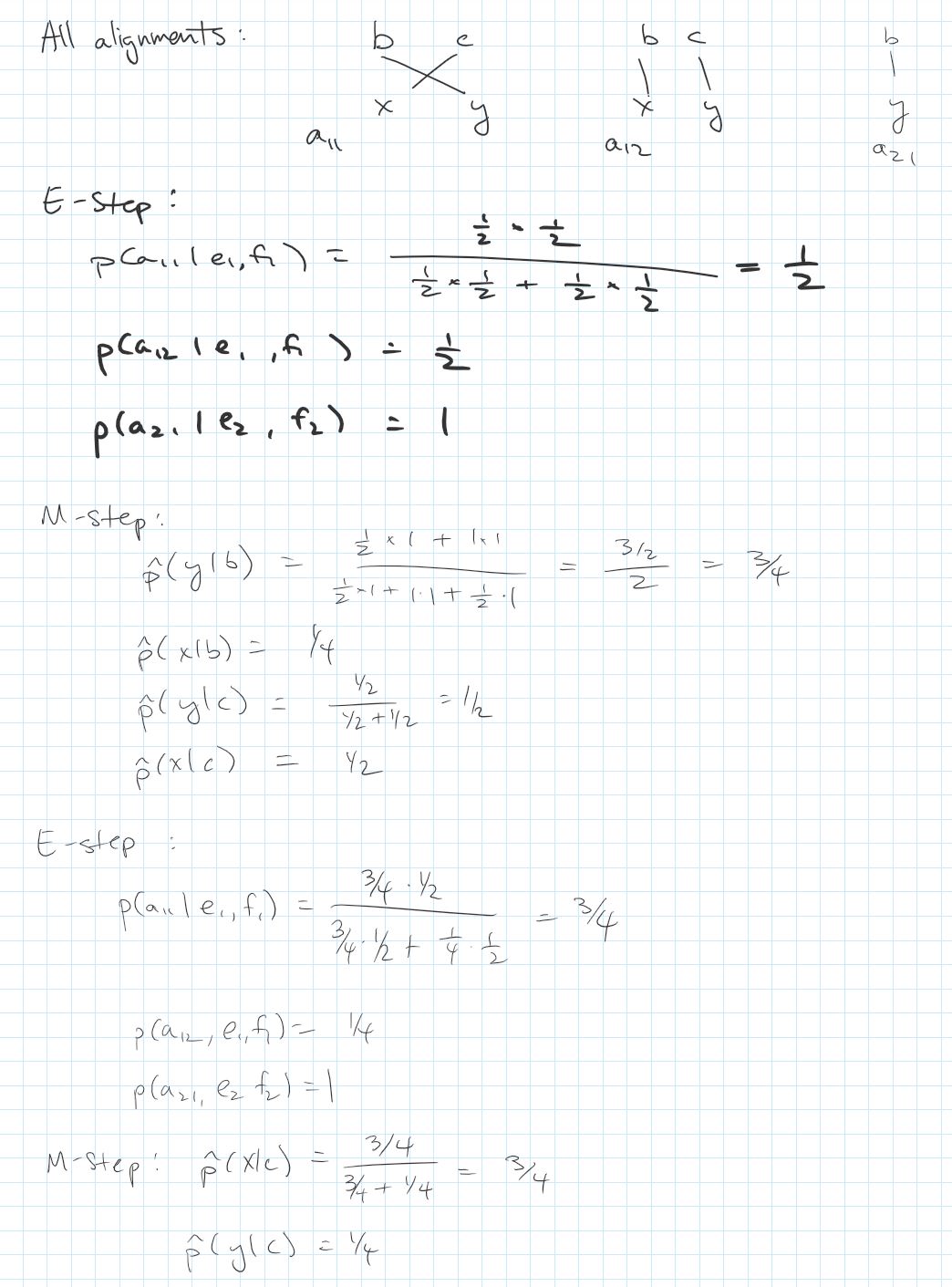
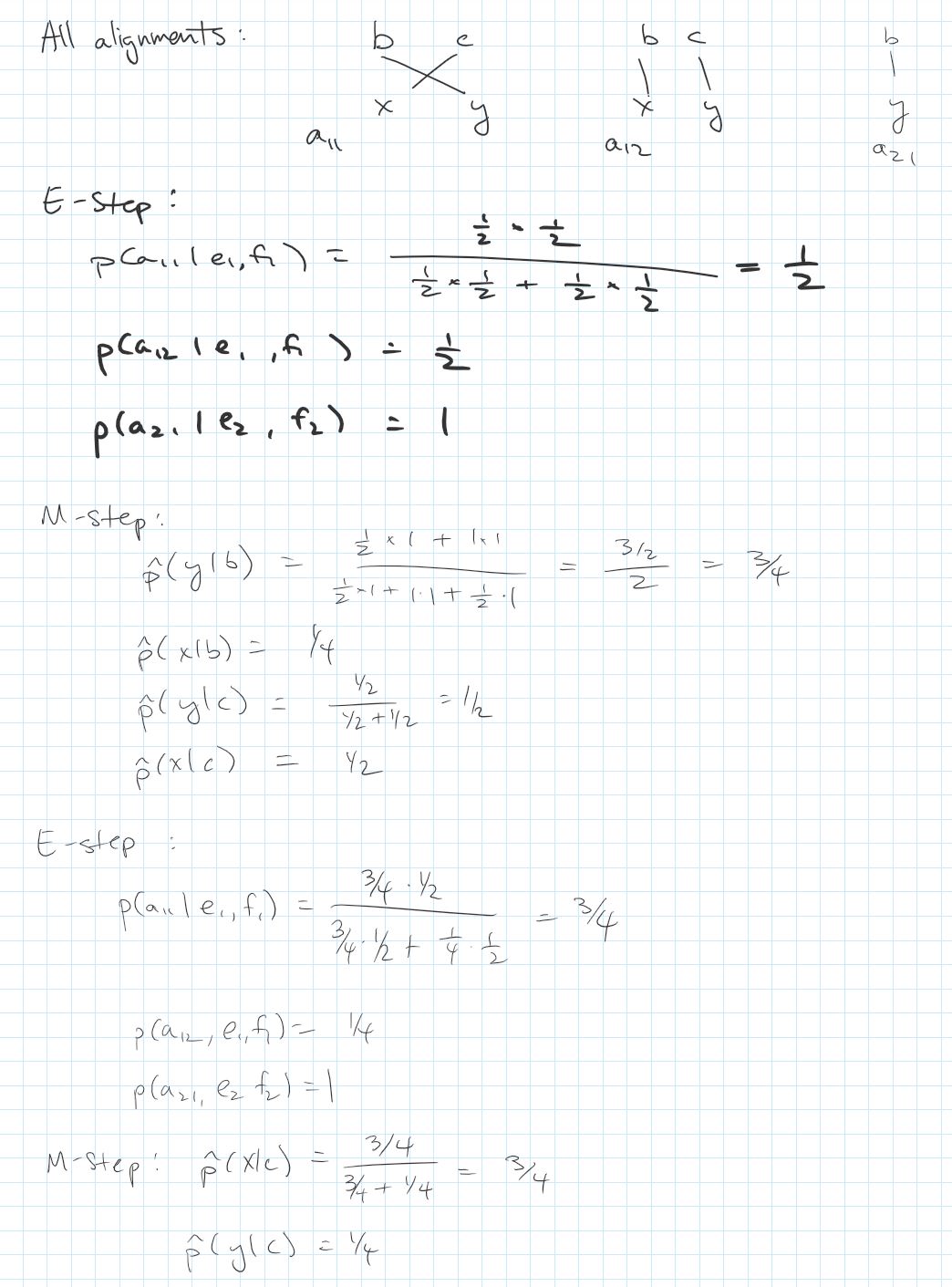
**Lecture 5**

* Continuing from last lecture (IBM Model 1)
  + Hidden variable (alignment, a one-to-one function)
    - Assume uniform
  + Example
    - : “b c”, “b”
    - : “x y”, “y”
    - Initialize





* Neural Models (We will be studying embeddings)
  + Review of neural models – refer to 6.036 lecture notes on Stellar
  + Huge impact on computer vision…not so much in NLP. Why?
    - Neural models do not really take into count the structure of language
    - Compared to previous models, neural models also need much more data
* Neural Language Models
  + Start with bigram model

