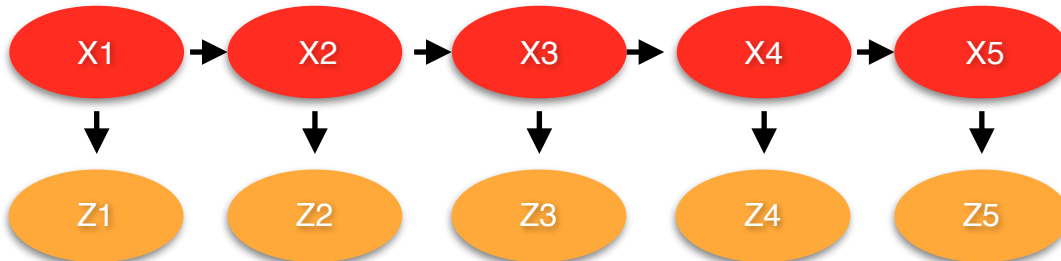


# Machine Discovery

## hw1-1 report

- Graphical model



- Assumption

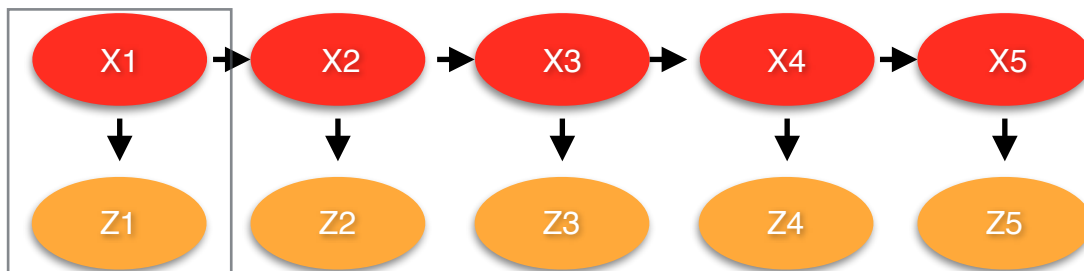
the probability of  $P(X_i | X_{i+1})$  is given from bigram.txt

the probability of  $P(Z_i | X_i)$  is given from encode.txt

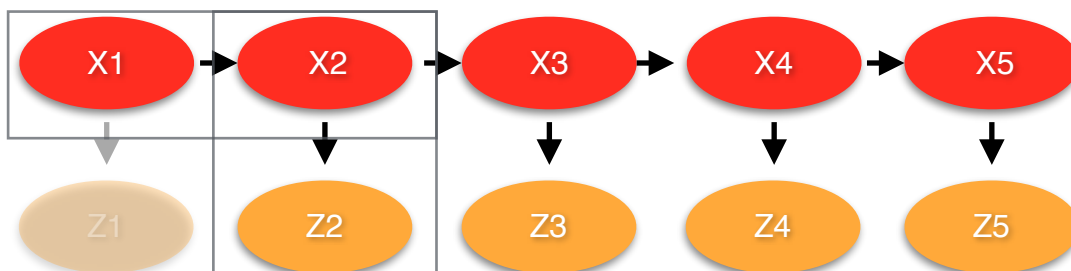
since  $P(Z_i | X_i) = P(X_i | Z_i) * P(Z_i) / P(X_i)$ , and regard  $P(Z_i)$ ,  $P(X_i)$  as same value for all  $Z_i \in Z$ ,  $X_i \in X$

so we use  $P(Z_i | X_i)$  to represent the value of given  $Z_i$  prob of  $P(X_i)$

$$M(X_1) = \max P(X_1 | Z_1) P(X_1 | Z_1)$$

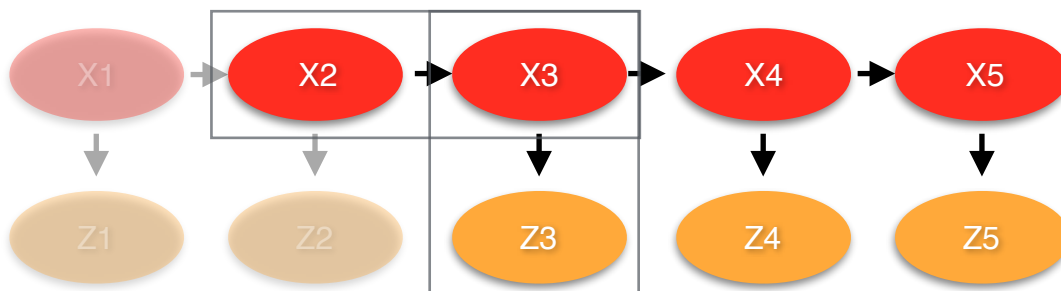


store maximum path from  $X_1$  to  $X_2$



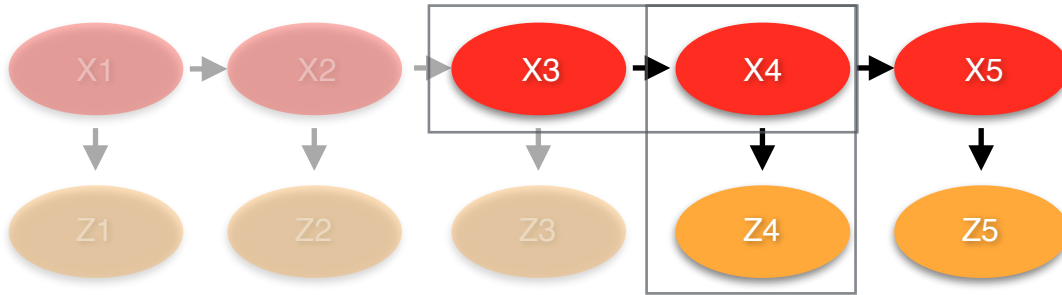
$$M(X_3) = \max P(X_3 | X_2) P(X_3 | Z_3) M(X_2)$$

store maximum path from  $X_2$  to  $X_3$



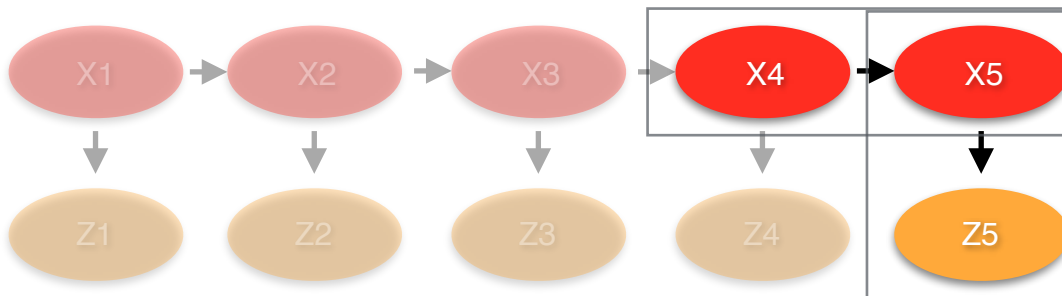
$$M(X4) = \max P(X4 | X3)P(X4 | Z4)M(X3)$$

store maximum path from X3 to X4

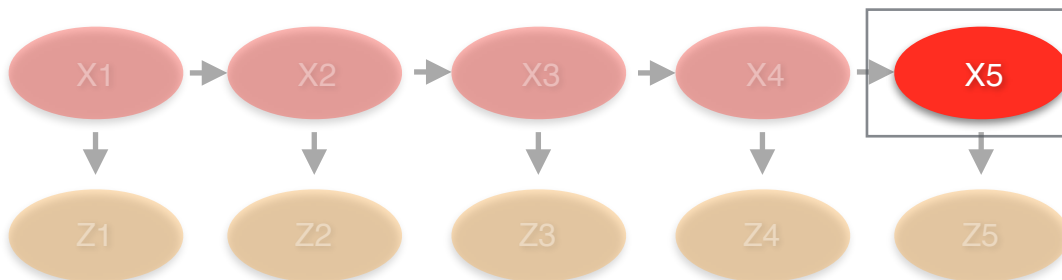


$$M(X5) = \max P(X5 | X4)P(X5 | Z5)M(X4)$$

store maximum path from X4 to X5



$$M(\text{end}) = \max P(' | X5)M(X5)$$



Calculate the maximum probability path for each encoded words in test file. Since space can't be encoded to other characters, we can split each word by space directly. Finally, output the final result to output file.

### • Reference

- [https://en.wikipedia.org/wiki/Viterbi\\_algorithm](https://en.wikipedia.org/wiki/Viterbi_algorithm)