

问题

1. 如何理解HMM和朴素贝叶斯、CRF和逻辑斯蒂回归的关系，以及HMM和CRF的关系？（相同点、不同点）
2. 为什么一般认为CRF比HMM的效果好？

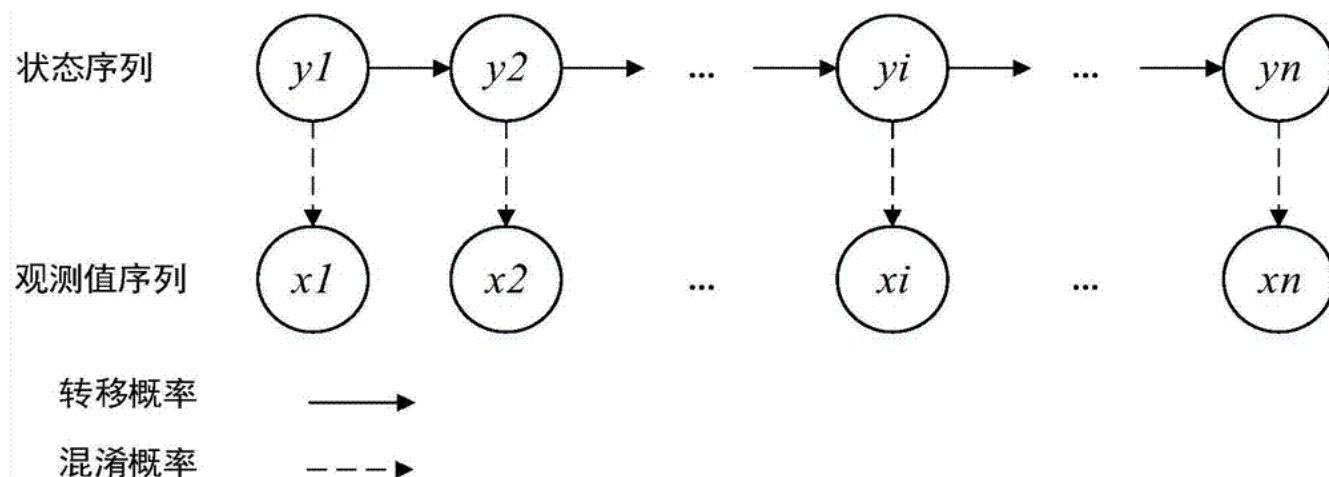
疑问

1. P229最后一段，“在算法S中需要使常数S取足够大，这样一来，每步迭代的增量向量会变大，算法收敛会变慢。”，然而S是出现在 δ 的分母上的，是不是写反了？“每步迭代的增量向量会变小，算法收敛会变慢”。

HMM & CRF

	图的类型	模型类型	特点（共同点）	概率计算	预测算法	学习算法
HMM	有向图	生成模型	状态之间的马尔科夫性	前向后向	维特比算法	EM算法
CRF	无向图	判别模型	状态之间的马尔科夫性	前向后向	维特比算法	IIS、拟牛顿法

超全的博客：[HMM](#)



CRF

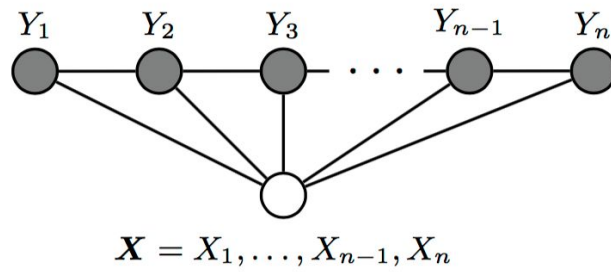
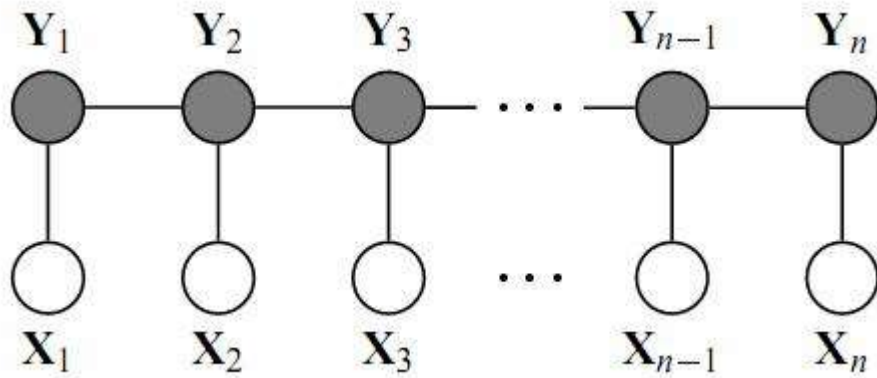


Figure 1: Graphical structure of a chain-structured CRFs for sequences. The variables corresponding to unshaded nodes are *not* generated by the model.



HMM计算的向量化形式

前向

$$\begin{aligned}\alpha_{t+1} &= (A^T \alpha_t) \odot b_{o_t} \\ &= (A \odot b_{o_t})^T \alpha_t\end{aligned}$$

后向

$$\begin{aligned}\beta_t &= (A \beta_{t+1}) \odot b_{o_{t+1}} \\ &= (A \odot b_{o_{t+1}}) \beta_{t+1}\end{aligned}$$

其中, \odot 为 element-wise product.

朴素贝叶斯、逻辑斯蒂回归、HMM、CRF的关系

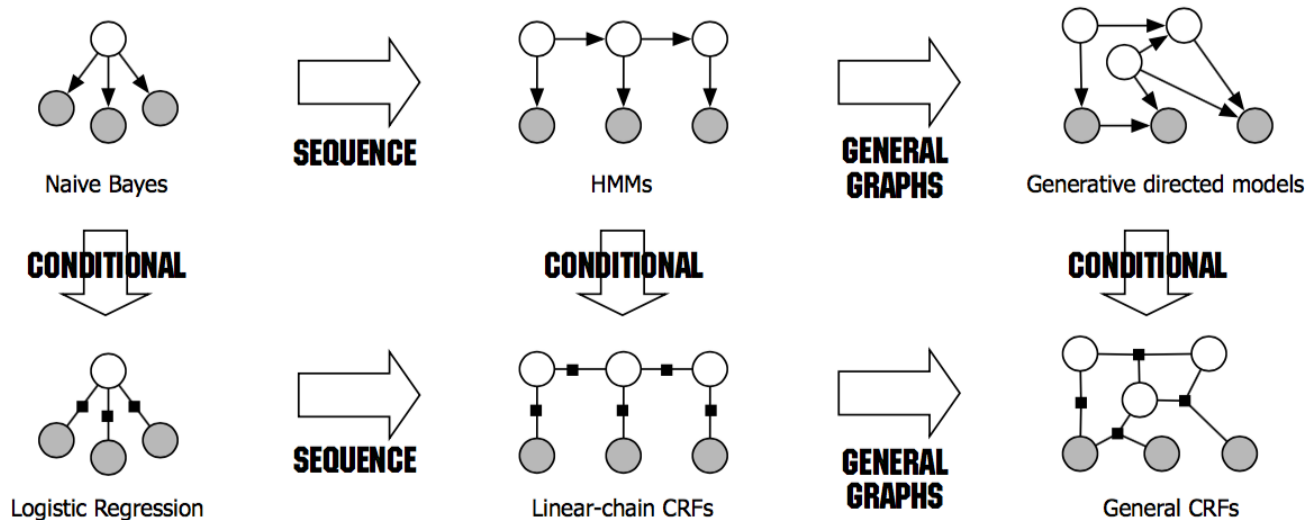


Figure 1.2 Diagram of the relationship between naive Bayes, logistic regression, HMMs, linear-chain CRFs, generative models, and general CRFs.