DSP Midterm

2012 Spring

1. (a) initial model · observations · iteration times

(b) 算法跟 2011 Autumn 3 差不多

2. 見 2011 Autumn 第5題

3. 見 2011 Autumn 第7題

4.

(a)

$$H(C_1) = -\sum_{A\sim D} P(x)\log[P(x)] = -[0.25\cdot\log_2(0.25) + 0.25\log_2(0.25) + 0.3\log_2(0.3) + 0.2\log_2(0.2)] = 1.985$$

Perplexity of $C_1 = 2^{H(C_1)} = 2^{1.985} = 3.96$

$$H(C_2) = -\sum_{A\sim D} P(x)\log[P(x)] = -[3\times 0.2\times\log_2(0.2) + 0.4\log_2(0.4)] = 1.92$$

Perplexity of $C_2 = 2^{H(C_2)} = 2^{1.92} = 3.78$

$$H(C_3) = -\sum_{A\sim D} P(x)\log[P(x)] = -[2\times 0.1\times\log_2(0.1) + 2\times 0.4\times\log_2(0.4)] = 1.72$$

Perplexity of $C_3 = 2^{H(C_3)} = 2^{1.72} = 3.29$

(b)

從 C_1 到 C_3 的 perplexity 是遞減的，由分布來看，C_1 的分佈最平均，ABCD 4種的數量差不多，C_2 次之，而 C_3 的分佈最不平均，B和D特別多。由 $C_1\sim C_3$ 的perplexity與其數量的關係，驗證了perplexity是表示"混淆度"的一個量。

(c)

① K 跟 C_1 的 cross entropy:

$$D[K\|C_1] = \sum_{A\sim D} P_K(x)\cdot\log\left[\frac{P_K(x)}{P_{C_1}(x)}\right] = 0.3\cdot\log_2\left(\frac{0.3}{0.25}\right) + 0.3\cdot\log_2\left(\frac{0.3}{0.25}\right) + 0.2\cdot\log_2\left(\frac{0.2}{0.2}\right) + 0.2\cdot\log_2\left(\frac{0.2}{0.3}\right) = 0.04083$$

② K 跟 C_2 的 cross entropy:

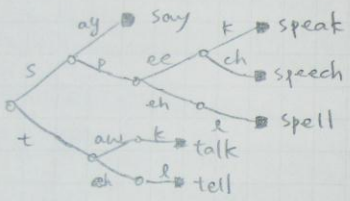
$$D[K\|C_2] = \sum_{A\sim D} P_K(x)\log\left[\frac{P_K(x)}{P_{C_2}(x)}\right] = 0.3\cdot\log_2\left(\frac{0.3}{0.2}\right) + 0.3\log_2\left(\frac{0.3}{0.2}\right) + 0.2\log_2\left(\frac{0.2}{0.2}\right) + 0.2\log_2\left(\frac{0.2}{0.4}\right) = 0.151$$

③ K 跟 C_3 的 cross entropy

$$D[K\|C_3] = 0.3\log_2\left(\frac{0.3}{0.1}\right) + 0.3\cdot\log_2\left(\frac{0.3}{0.4}\right) + 0.2\cdot\log_2\left(\frac{0.2}{0.1}\right) + 0.2\cdot\log_2\left(\frac{0.2}{0.4}\right) = 0.3509$$

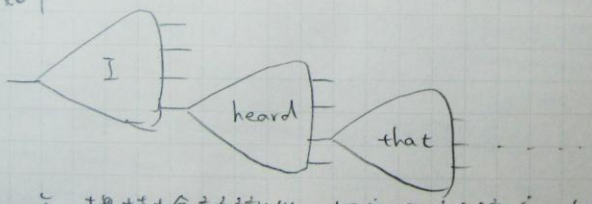
∵ K 跟 C_1 的 cross entropy 最小 ∴ K 應是屬於 C_1

5. 以下面這棵 tree lexicon 為例說明



每一條邊都是一個 HMM model (以 phone 為 unit)，每一個 leaf node 都是一個字，對於不同但是擁有某一部份相同 phonem 的字群 (ex speak 跟 speech)，在搜尋時可以共用一棵樹。

所以在做 large vocabulary continuous speech recognition 時，演算法示意圖如下



每一棵樹負責辨識出一個字，而辨識完一個字之後會使用 language model 找出下一棵樹，最後辨識出整段 speech。