CSE 473/573 Fall 2007 December 13 2007

Final Exam Solution Set

1. In an image we have found four objects {a',b',c',d'}. There are five possible labels {a,b,c,d,e} for these objects. Here are the unary properties for the label set:

Pa={capacitor}
Pb=Pe={capacitor, resistor}
Pc=Pd={transistor}

The relations for the label set are:

a is smaller than b
c is smaller than d
c is smaller than a
e is smaller than a
e is smaller than d
b is smaller than d

There are no m-ary relations for m>2. Here is what we determine about these four objects by preprocessing the image:

b' is a transistor
c' is a capacitor
a' is a capacitor or a resistor
d' is a capacitor
a' is smaller than c'
b' is smaller than c'
c' is smaller than d'

Apply the discrete relaxation constraint propogationalgorithm to find all distinct consistent labellings for these objects. Specify each consistent labelling. Show your work.

The columns below show the steps in the constaint propogation discrete relaxation algorithm. U is the unary pass, followed by two relation-checking passes.

	U	R1	R2
a'	*abe*	*ae*	*e*
b'	*cd*	*c*	*c*
c'	*abe*	*a*	*a*
d'	*abe*	*b*	*b*

The algorithm is stationary after R2. There is just one consistent labelling: *a'* is and *e*, *b'* is a *c*, *c*' an *a* and *d'* is a *b*.

2. Let f(i,j) be a binary image with foreground pixels f(i,j)=1 and background pixels f(i,j)=0. Denote the ordinary moments of f by m_{pq} and the central moments of f by μ_{pq}.

(a) How can the number of foreground pixels in f be determined from its ordinary moments? From its central moments?

$m_{00} = \sum\sum i^0 j^0 f(i,j)$, so m_{00} equals the number of foreground pixels. The same is true for μ_{00}.

(b) What does μ_{10} tell us about f?

For any image,

$\mu_{10} = \sum\sum (i - x_c) f(i,j) = m_{10} - x_c m_{00} = 0$. So μ_{10} tells us nothing about *f*.

(c) Show that by expanding the formula for μ_{11}, you can express that central moment in terms of several ordinary moments m_{pq}.

$$\begin{aligned}
\mu_{11} &= \sum\sum (i - x_c)(j - y_c) f(i,j). \text{ Expanding the product,} \\
&= \sum\sum ijf(i,j). - \sum\sum iy_c f(i,j). - \sum\sum x_c jf(i,j). + \sum\sum x_c y_c f(i,j). \\
&= m_{11} - y_c m_{10} - x_c m_{01} + x_c y_c m_{00} \\
&= m_{11} - \frac{m_{01} m_{10}}{m_{00}}
\end{aligned}$$

3. Consider a two-class {ω1,ω2} minimum-error classification problem using a 2D feature vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. The class-conditional probabilities are Gaussian with means and covariance matrices given by $\mu_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mu_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \varphi_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \varphi_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$. The prior probabilities are $P(\omega_1)$ and $P(\omega_2)=1- P(\omega_1)$.

(a) For what range of values of the prior $P(\omega_1)$ will $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ be classified as ω_1?

The discrimination functions are

$$g_1'''(x) = -\frac{1}{2}(\begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix}) + \ln(2\pi) - \frac{1}{2}\ln(1) + \ln P_1$$

$$g_2'''(x) = -\frac{1}{2}\left(\begin{bmatrix} x_1 \\ x_2+1 \end{bmatrix}^T \begin{bmatrix} x_1 \\ 2(x_2+1) \end{bmatrix}\right) + \ln(2\pi) - \frac{1}{2}\ln(\frac{1}{2}) + \ln(1-P_1)$$

Plugging in $x_1 = x_2 = 0$,

$$g_1'''\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = -\frac{1}{2}(1) + \ln(2\pi) + \ln P_1$$

$$g_2'''\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = -\frac{1}{2}(2) + \ln(2\pi) - \frac{1}{2}\ln(\frac{1}{2}) + \ln(1-P_1)$$

Now we will choose ω_1 if $g_1''' > g_2'''$. If they are equal, we can pick either. So the condition to select ω_1 is

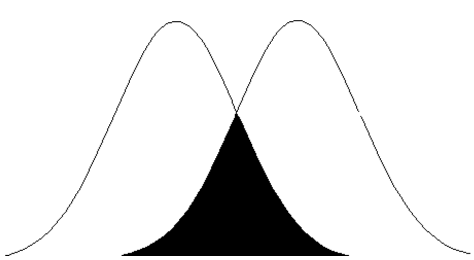
$$-\frac{1}{2} + \ln P_1 > -1 + \ln\sqrt{2} + +\ln(1-P_1)$$

Solving

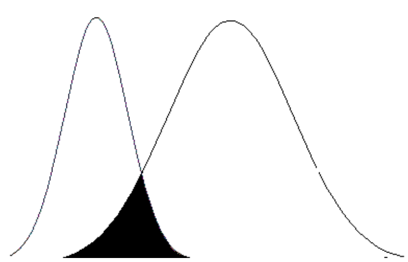
$$P_1 > \frac{\sqrt{2}}{e^{1/2} + \sqrt{2}}$$

(b) If we could only use one feature, either x_1 or x_2, but not both, which would give the better error performance? Explain your reasoning. A graph showing the pdf's might be helpful to illustrate your thinking.

The variances of x_1 for both classes is 1, while that of class 2 for x_2 is ½. This implies there is less uncertainty using x_2, which leads to less error. The diagrams below indicate the errors using x_1 and x_2. The dark areas are proportional to the expected classification error.



Using x_1



Using x_2