2. (a) Consider a multi-label problem with 2 variables y_1 and y_2 each taking 3 states $\{1, 2, 3\}$. Let the unary terms be given by $\theta^l_{y_1} = \{[l = 1] \rightarrow 0.5, [l = 2] \rightarrow 1.5, [l = 3] \rightarrow 1.0\}$, and $\theta^l_{y_2} = \{[l = 1] \rightarrow 1.5, [l = 2] \rightarrow 1.5, [l = 3] \rightarrow 0.0\}$. We use the pairwise model and it is given by $\theta^{lm}_{y_1 y_2} = |l - m|$, where $|l - m|$ denotes the absolute value of the difference between the two labels. Show that the pairwise potentials satisfy the metric condition. Show the iterations in alpha-expansion till it converges. Note that you will construct a Boolean function in every iteration. You will manually compute the solution for every maxflow/mincut step. [20 points]

(b) Is the following function submodular or not?

$$f(x_1, x_2, x_3) = x_1 x_2 x_3 - x_1 - x_2 \tag{1}$$

Prove your result using the set definition for submodularity: $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$. [10 points]

$\theta^{lm}_{y_1 y_2} = |l-m|$

	1	2	3
y_1	0.5	1.5	1.0
y_2	1.5	2	0.0

$y_1 \backslash y_2$	1	2	3
1	0	1	2
2	1	0	1
3	2	1	0

(a) Matrix condition:

$$\begin{cases} \theta^{l_a l_a}_{y_1 y_2} = 0 & (1) \\ \theta^{l_a l_b}_{y_1 y_2} = \theta^{l_b l_a}_{y_1 y_2} \geq 0 & (2) \\ \theta^{l_a l_b}_{y_1 y_2} + \theta^{l_b l_c}_{y_1 y_2} \geq \theta^{l_a l_c}_{y_1 y_2} & (3) \end{cases}$$

Proof:

(1) $\theta^{l_a l_a}_{y_1 y_2} = |l_a - l_a| = 0$

(2) $\theta^{l_a l_b}_{y_1 y_2} = |l_a - l_b| = \theta^{l_b l_a}_{y_1 y_2}$

(3). $\theta^{l_a l_b}_{y_1 y_2} + \theta^{l_b l_c}_{y_1 y_2} = |l_a - l_b| + |l_b - l_c|$

$\theta^{l_a l_c}_{y_1 y_2} = |l_a - l_c|$

As $|x| + |y| \geq |x + y|$, for $\forall x, y \in R$,

$|l_a - l_b| + |l_b - l_c| \geq |l_a - l_c|$

So $\theta^{l_a l_b}_{y_1 y_2} + \theta^{l_b l_c}_{y_1 y_2} \geq \theta^{l_a l_c}_{y_1 y_2}$.

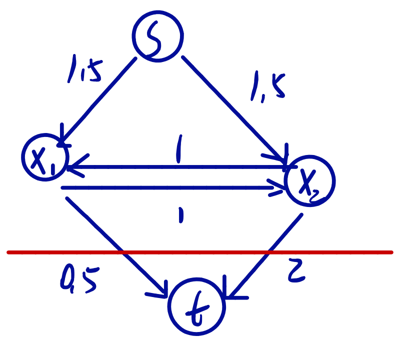
(b) $y_1 = 1,\ y_2 = 1$

① $\alpha = 1$, no change.

② $\alpha = 2$.

$$\begin{cases} y_1 = 1 \\ y_1 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ x_1 = 1 \end{cases} \qquad \begin{cases} y_2 = 1 \\ y_2 = 2 \end{cases} \Rightarrow \begin{cases} x_2 = 0 \\ x_2 = 1 \end{cases}$$

$$E = 0.5\bar{x}_1 + 1.5x_1 + 1.5\bar{x}_2 + 2x_2 + \bar{x}_1 x_2 + \bar{x}_2 x_1$$



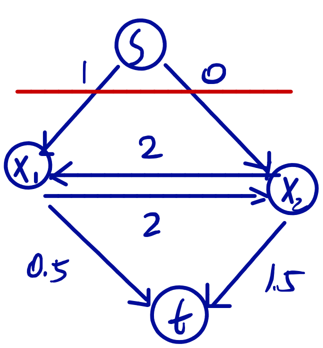
so $x_1 = x_2 = 0$

$y_1 = y_2 = 1$.

③ $\alpha = 3$.

$$\begin{cases} y_1 = 1 \\ y_1 = 3 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ x_1 = 1 \end{cases} \qquad \begin{cases} y_2 = 1 \\ y_2 = 3 \end{cases} \Rightarrow \begin{cases} x_2 = 0 \\ x_2 = 1 \end{cases}$$

$$E = 0.5\bar{x}_1 + x_1 + 1.5\bar{x}_2 + 2\bar{x}_1 x_2 + 2x_2\bar{x}_1$$



$x_1 = x_2 = 1$

$y_1 = y_2 = 3$

	1	2	3
y_1	0.5	1.5	1.0
y_2	1.5	2	0.0

④ $\alpha = 1$

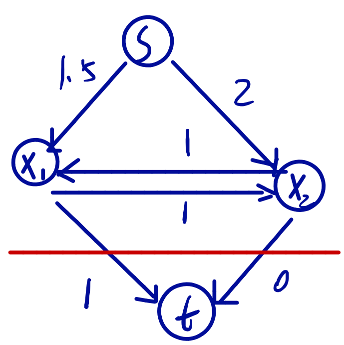
$$\begin{cases} y_1 = 1 \\ y_2 = 3 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ x_1 = 1 \end{cases} \qquad \begin{cases} y_1 = 1 \\ y_2 = 3 \end{cases} \Rightarrow \begin{cases} x_2 = 0 \\ x_2 = 1 \end{cases}.$$

The energy function is the same as ③, so $y_1 = y_2 = 3$.

⑤. $\alpha = 2$.

$$\begin{cases} y_1 = 3 \\ y_2 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 1 \end{cases} \qquad \begin{cases} y_2 = 3 \\ y_2 = 2 \end{cases} \Rightarrow \begin{cases} x_2 = 0 \\ x_2 = 1 \end{cases}$$

$$E = \bar{x}_1 + 1.5 x_1 + 2 x_2 + \bar{x}_1 x_2 + x_1 \bar{x}_2$$



so $x_1 = x_2 = 0$

$y_1 = y_2 = 3$.

⑥ $\alpha = 3$. no change.

As there is no change in a whole iteration $\alpha = 1, 2, 3$, it converges.

$$f(x_1,x_2,x_3) = x_1x_2x_3 - x_1 - x_2$$

(b) Let $A = \{1,2\} \rightarrow x_1=1, x_2=1, x_3=0$

$B = \{3\} \rightarrow x_1=0, x_2=0, x_3=1$

$$f(A) = (1)(1)(0) - (1) - (1) = -2$$
$$f(B) = (0)(0)(0) - (0) - (0) = 0$$
$$f(A\cup B) = (1)(1)(1) - (1) - (1) = -1$$
$$f(A\cap B) = (0)(0)(0) - (0) - (0) = 0$$

$$\Longrightarrow f(A) + f(B) = -2$$
$$f(A\cup B) + f(A\cap B) = -1$$

$$\Longrightarrow f(A) + f(B) < f(A\cup B) + f(A\cap B)$$

The function is not Submodular

One counter example is sufficient to show that the function is not Submodular.

———x———

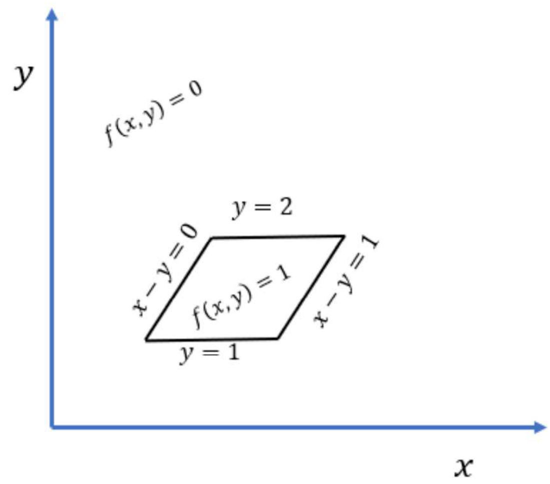


Figure 1:

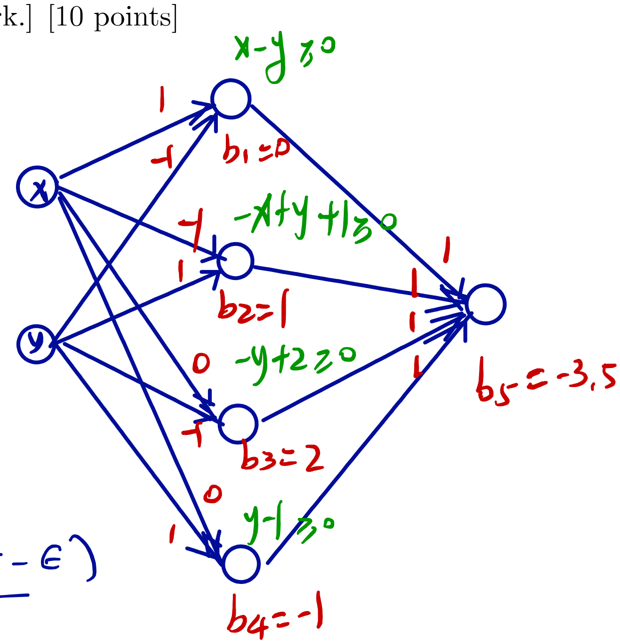
4. Consider a linear threshold unit T. Let $x_1, x_2, , x_n$ be the real inputs to the linear threshold unit and $w_1, w_2, , w_n$ be the learned real weights and let b be the bias term. Then the output from T will be 1 if $\sum_{i=1}^{n}(w_i x_i + b) \geq 0$ and 0 otherwise. Consider a function $f(x, y)$ that takes two real inputs (x, y) and gives Boolean output 1 or 0 as shown in Fig 1. In particular, the function $f(x, y)$ is 1 inside the quadrilateral, and 0 outside.

   (a) Build the function using linear threshold units. You are free to use as many linear threshold units as you need. Manually come up with weights and biases for each linear threshold unit. [20 points]

   (b) Build the function using RELU to implement the same function $f(x, y)$. Note that $RELU(x) = max(0, x)$. You are free to use as many RELU units as you need. Manually come up with weights and biases in the network. [Hint: Try to model the threshold function T using one or more RELU activation functions, and replace all threshold units using RELU units in the previously designed network.] [10 points]

(a)

$$\begin{cases} x-y \geq 0 \\ x-y \leq 1 \\ y \leq 2 \\ y \geq 1 \end{cases} \quad \begin{cases} x-y \geq 0 \\ -x+y+1 \geq 0 \\ -y+2 \geq 0 \\ y-1 \geq 0 \end{cases} \Rightarrow$$



(b) $$T(x) = \frac{RELU(x+\epsilon) - RELU(x-\epsilon)}{2\epsilon}$$

$\epsilon = 0.00001$, Other correct solutions will also get points.