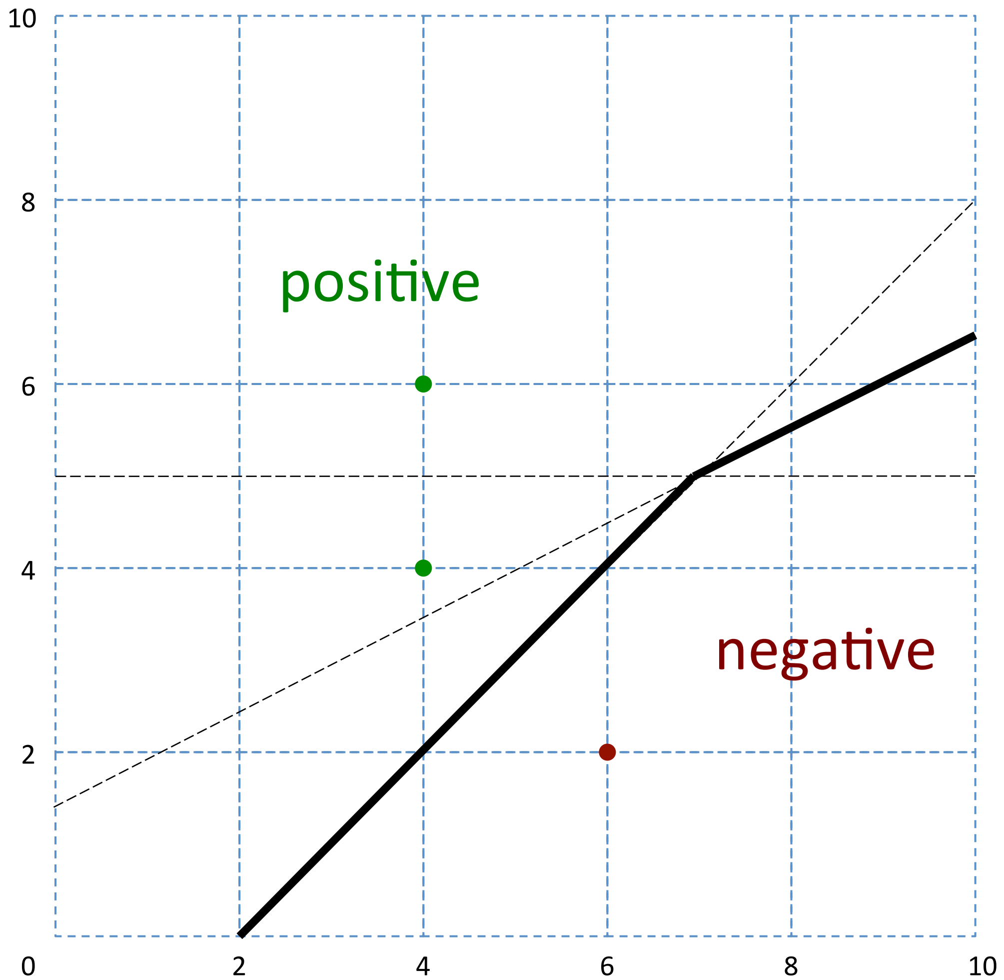
1.



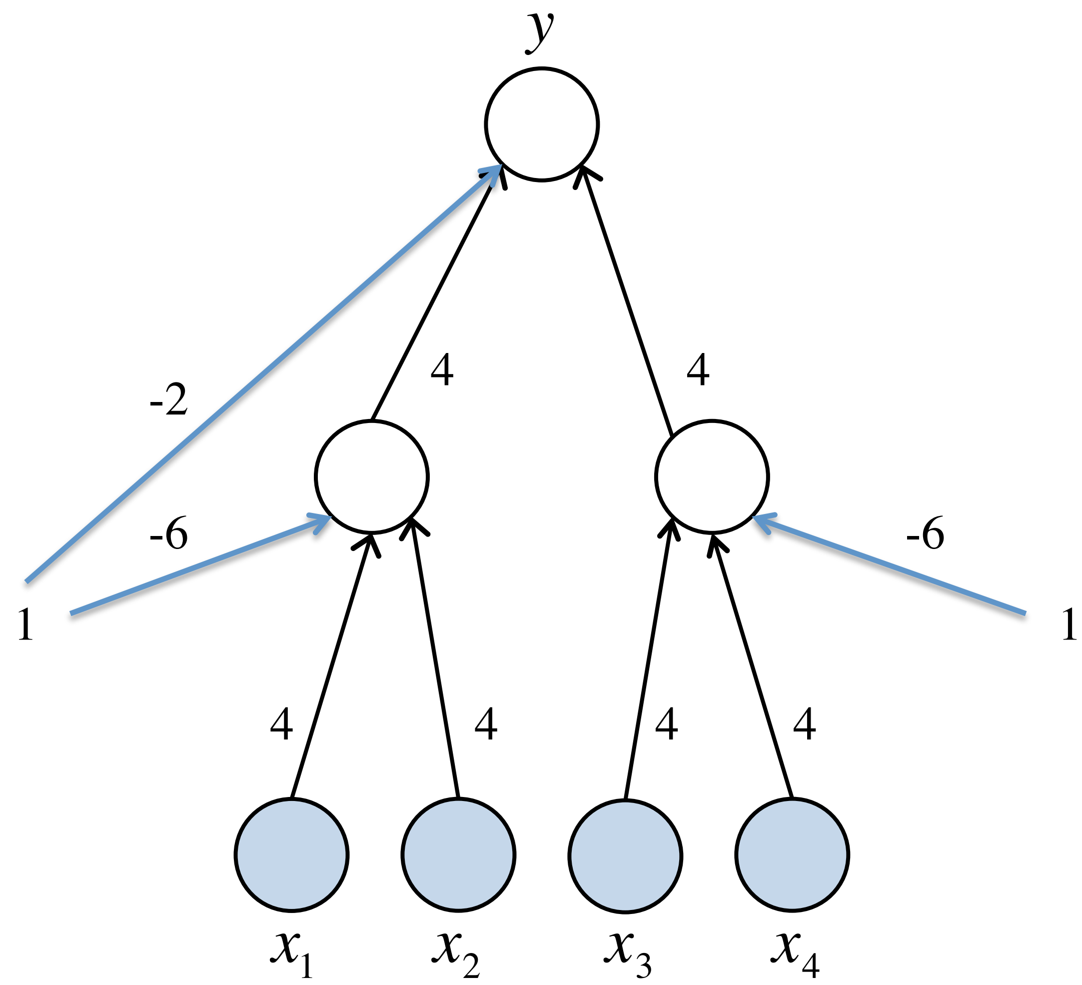
2.

distances are approximate since it’s hard to tell exactly what the coordinates of some of the instances are

node	distance	best distance	best node	priority queue
		∞		(f, 0)
f	7.07	7.07	f	(h, 0) (c, 1)
h	7.07	7.07	f	(i, 0) (c, 1) (g, 5)
i	3.0	3.0	i	(c, 1) (j, 3) (g, 5)
c	2.0	2.0	c	(b, 0) (e, 0) (j, 3) (g, 5)
b	4.47	2.0	c	(e, 0) (j 3) (a, 4) (g, 5)
e	7.81	2.0	c	(d, 0) (j, 3) (a, 4) (g, 5)
d	5.39	2.0	c	(j, 3) (a, 4) (g, 5)

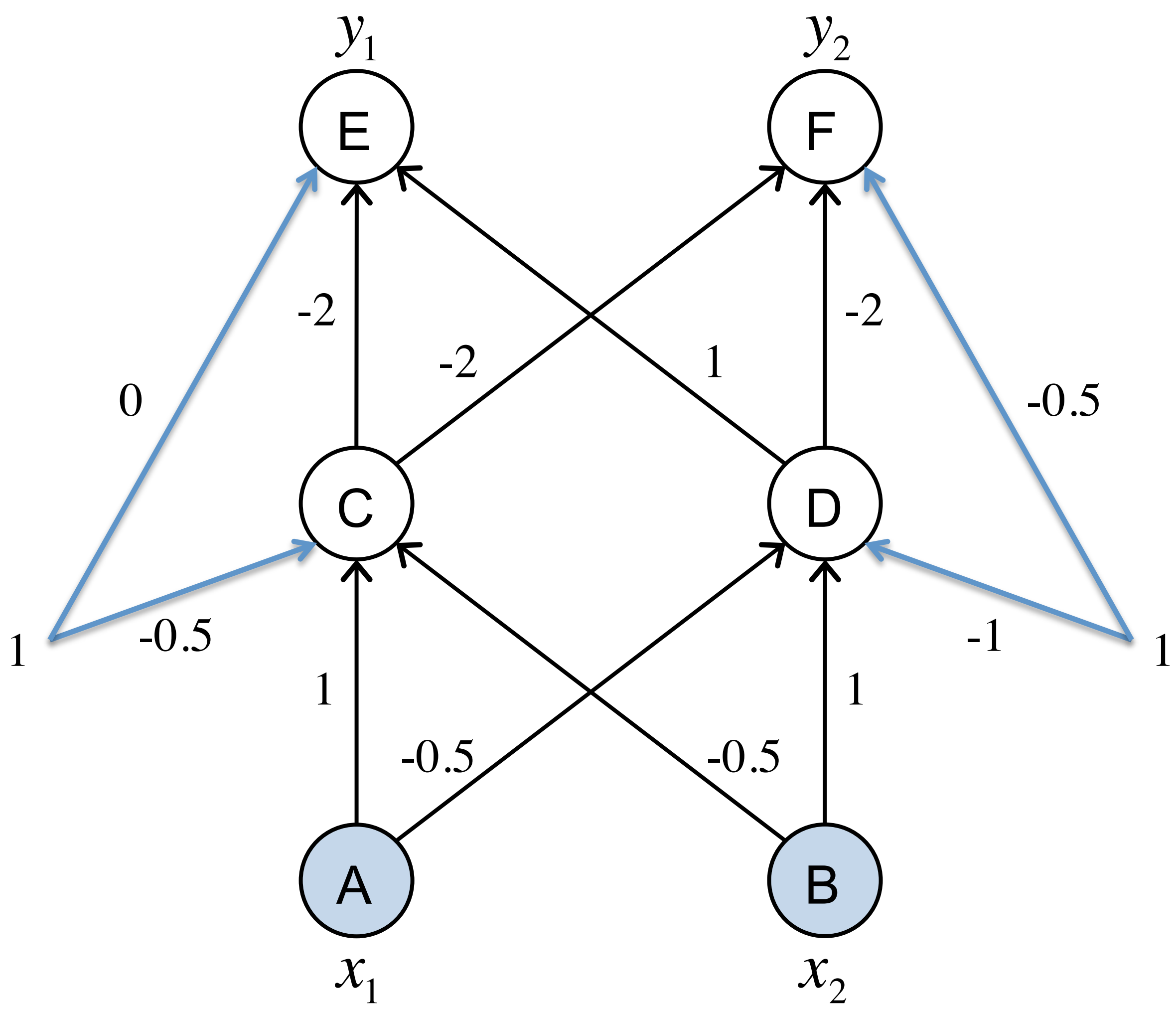
node c would be returned as the nearest neighbor

3. the weights don't have to be these exact numbers, but they should be set such that the hidden units are approximating AND's and the output unit is approximating an OR



$$y = (x_1 \wedge x_2) \vee (x_3 \wedge x_4)$$

$\boxed{4.}$ I'll show how to do the backprop calculations in general



training instance: $\boldsymbol{x} = [0, 1]$ $\boldsymbol{y} = [1, 0]$

first determine output of each unit

$$o_C = \frac{1}{1+e^{-(-1)}} = 0.2689$$

$$o_D = \frac{1}{1+e^{-(0)}} = 0.5$$

$$o_E = \frac{1}{1+e^{-(-0.5378\ +\ 0.5)}} = 0.4905$$

$$o_F = \frac{1}{1+e^{-(-0.5378\ -\ 1\ -\ 0.5)}} = 0.1153$$

then calculate δ's

$$\delta_E = o_E(1-o_E)(1-o_E) = .1273$$

$$\delta_F = o_F(1-o_F)(0-o_F) = -0.01176$$

$$\begin{aligned}\delta_C &= o_C(1-o_C)\left(\delta_E w_{EC} + \delta_F w_{FC}\right)\\ &= 0.2689(1-0.2689)\left(0.1273(-2)+(-0.01176)(-2)\right)\\ &= -0.0454\end{aligned}$$

$$\begin{aligned}\delta_D &= o_D(1-o_D)\left(\delta_E w_{ED} + \delta_F w_{FD}\right)\\ &= 0.5(1-0.5)\left(0.1273(1)+(-0.01176)(-2)\right)\\ &= 0.03771\end{aligned}$$

now determine the weight updates

$$\Delta w_{EC} = \eta\ \delta_E\ o_C = 0.1 \times 0.1273 \times 0.2689 = 0.003423$$

$$\Delta w_{ED} = \eta\ \delta_E\ o_D = 0.1 \times 0.1273 \times 0.5 = 0.006365$$

$$\Delta w_{Eb} = \eta\ \delta_E\ (1) = 0.1 \times 0.1273 \times 1 = 0.01273$$

don't forget the bias parameters
(I'll use *b* to denote the bias "unit")

$$\Delta w_{FC} = \eta\ \delta_F\ o_C = 0.1 \times (-0.01176) \times 0.2689 = -0.0003162$$

$$\Delta w_{FD} = \eta\ \delta_F\ o_D = 0.1 \times (-0.01176) \times 0.5 = -0.000588$$

$$\Delta w_{Fb} = \eta\ \delta_F\ o_b = 0.1 \times (-0.01176) \times 1 = -0.001176$$

$$\Delta w_{CA} = \eta\ \delta_C\ o_A = 0.1 \times (-0.0454) \times 0 = 0$$

$$\Delta w_{CB} = \eta\ \delta_C\ o_B = 0.1 \times (-0.0454) \times 1 = -0.00454$$

$$\Delta w_{Cb} = \eta\ \delta_C\ (1) = 0.1 \times (-0.0454) \times 1 = -0.00454$$

$$\Delta w_{DA} = \ldots$$