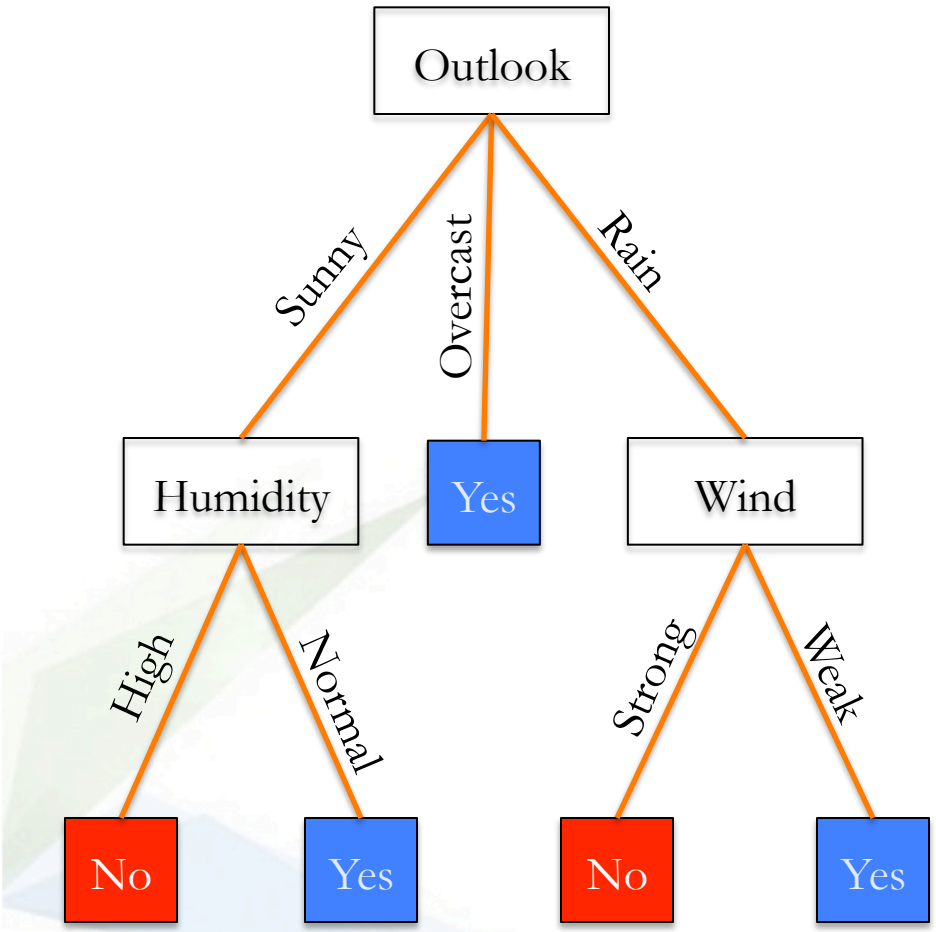
Will Nadal Play Tennis?

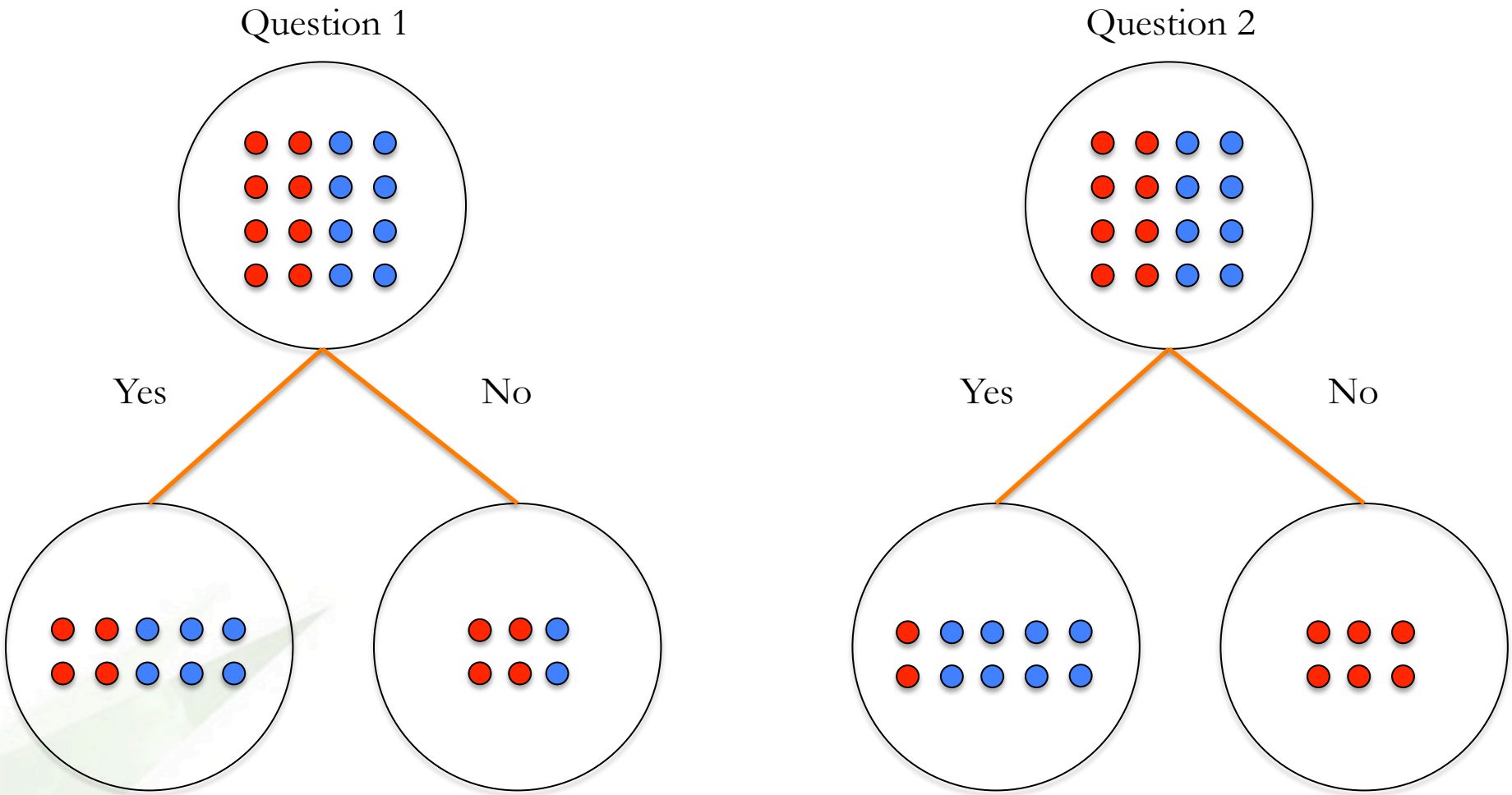
Rafael Nadal

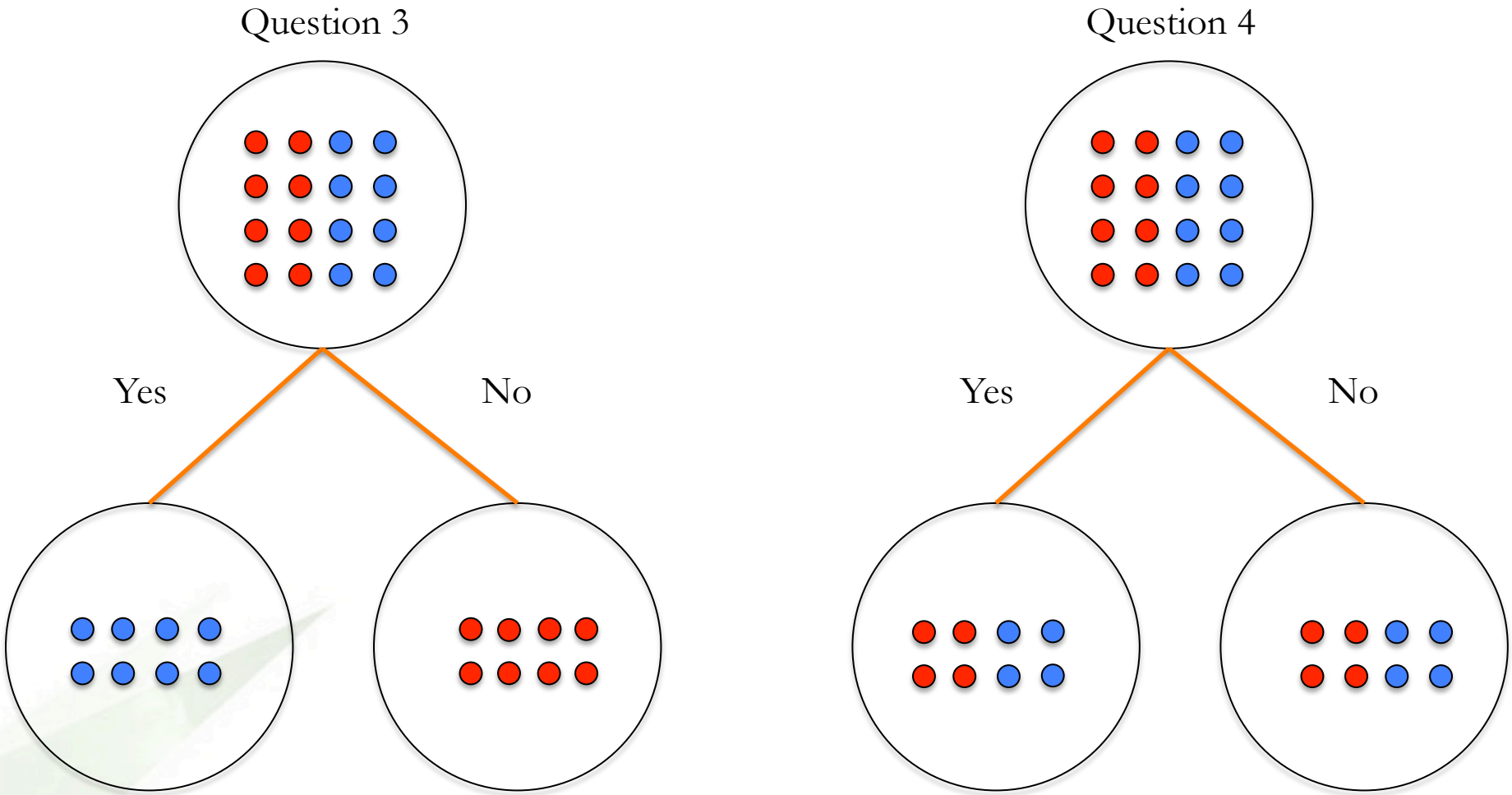
Day	Outlook	Temp	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Weak	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Strong	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Will Nadal Play Tennis?

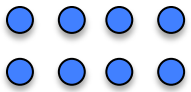


Day	Outlook	Temp	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Weak	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Strong	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

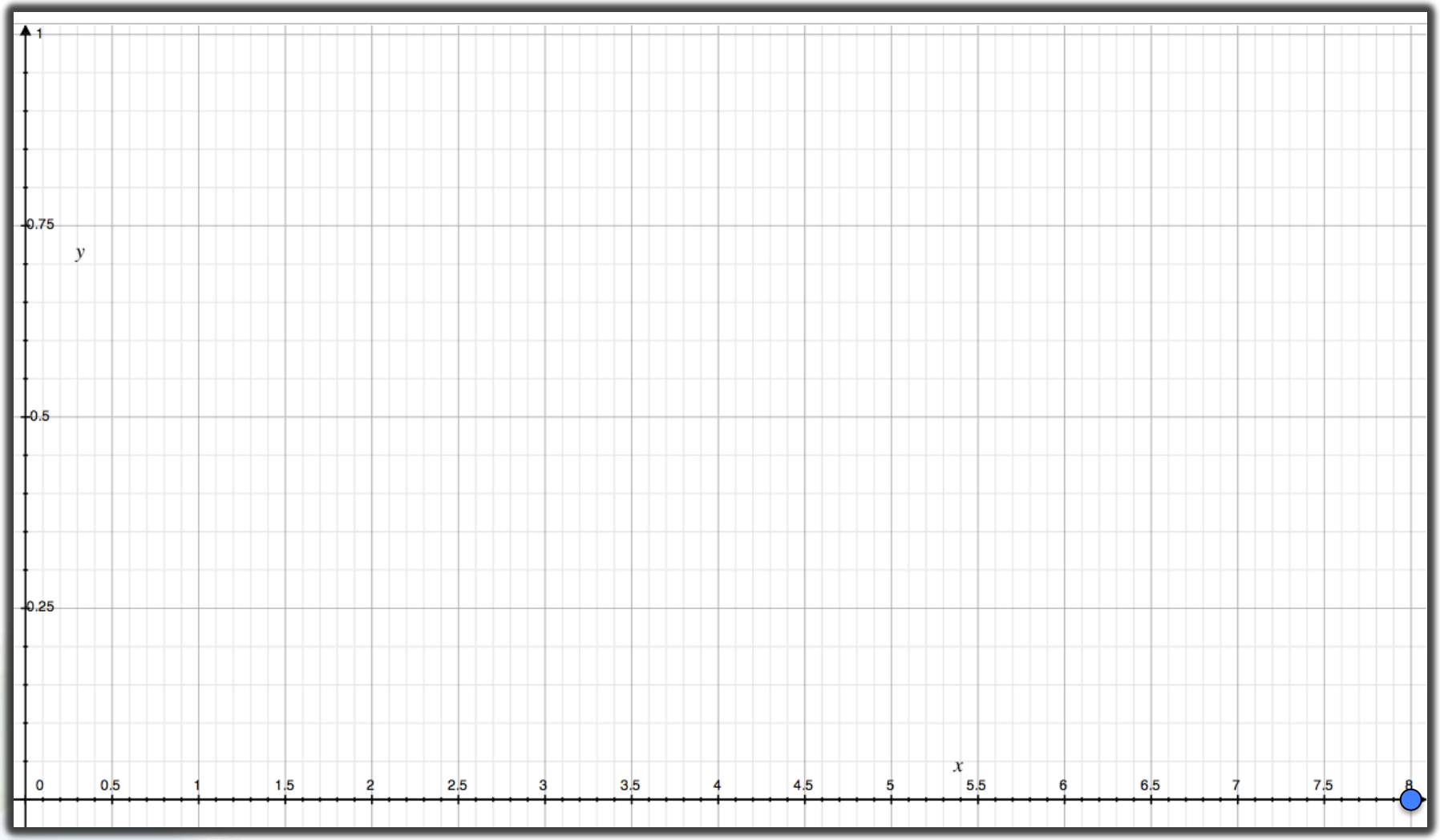

Question 1
Yes
No
Question 2
Yes
No

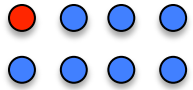


$$E(S) = -p_+ log(p_+) - p_- log(p_-)$$

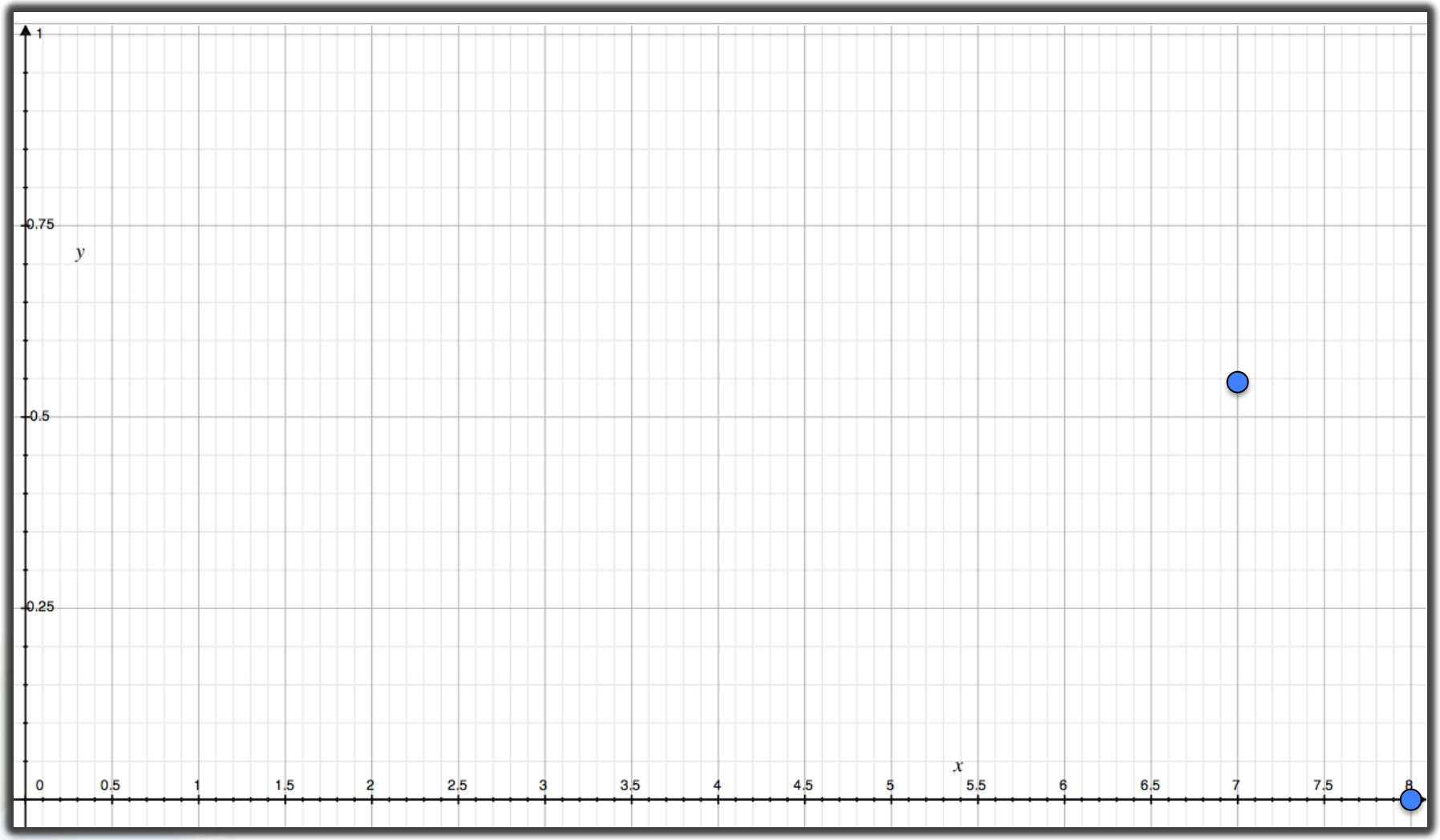


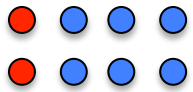
$$-\left(\frac{8}{8}\right) log_2 \left(\frac{8}{8}\right) - \left(\frac{0}{8}\right) log_2 \left(\frac{0}{8}\right) = 0$$



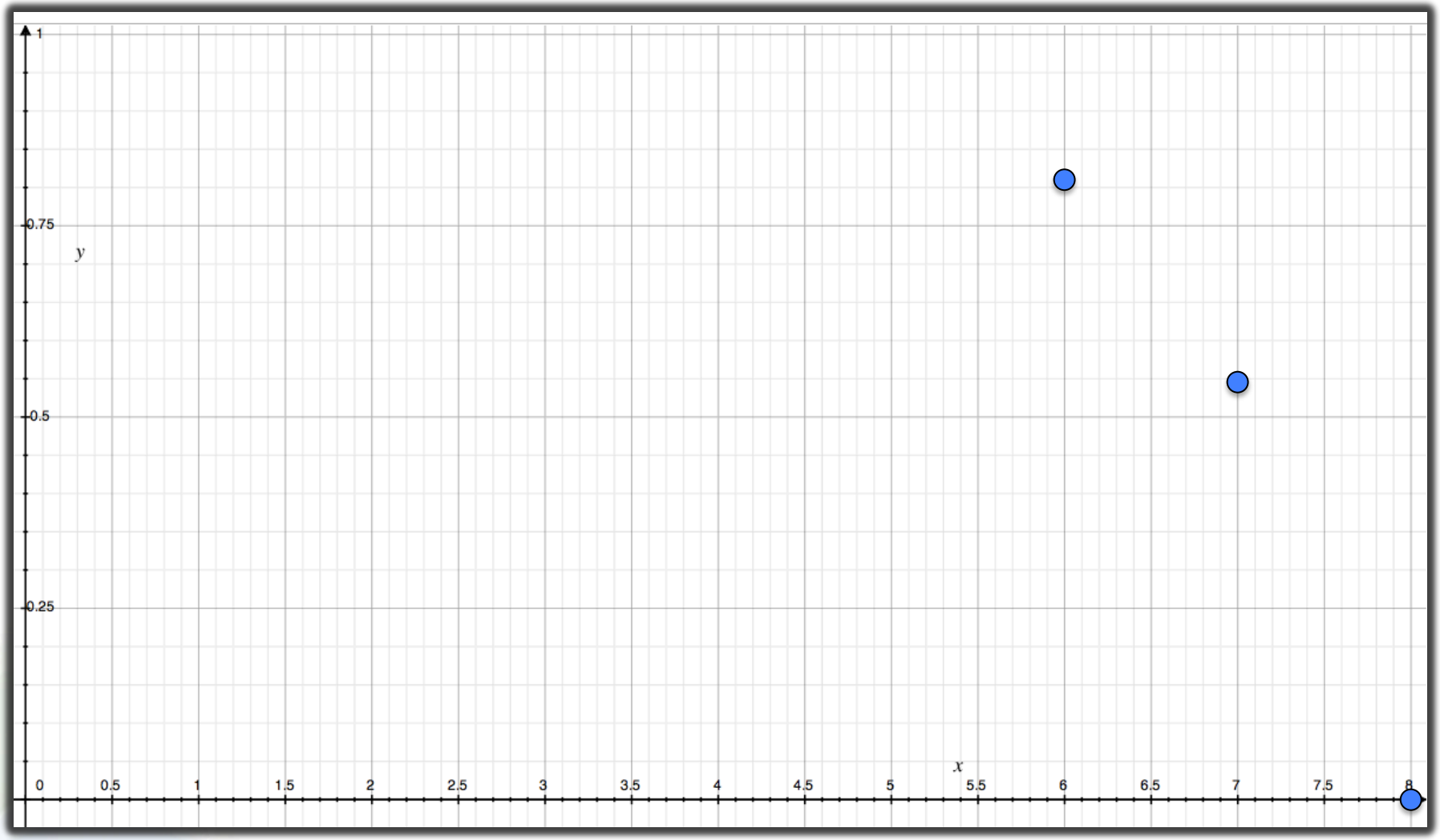


$$-\left(\frac{7}{8}\right) log_2\left(\frac{7}{8}\right) - \left(\frac{1}{8}\right) log_2\left(\frac{1}{8}\right) = 0.54$$

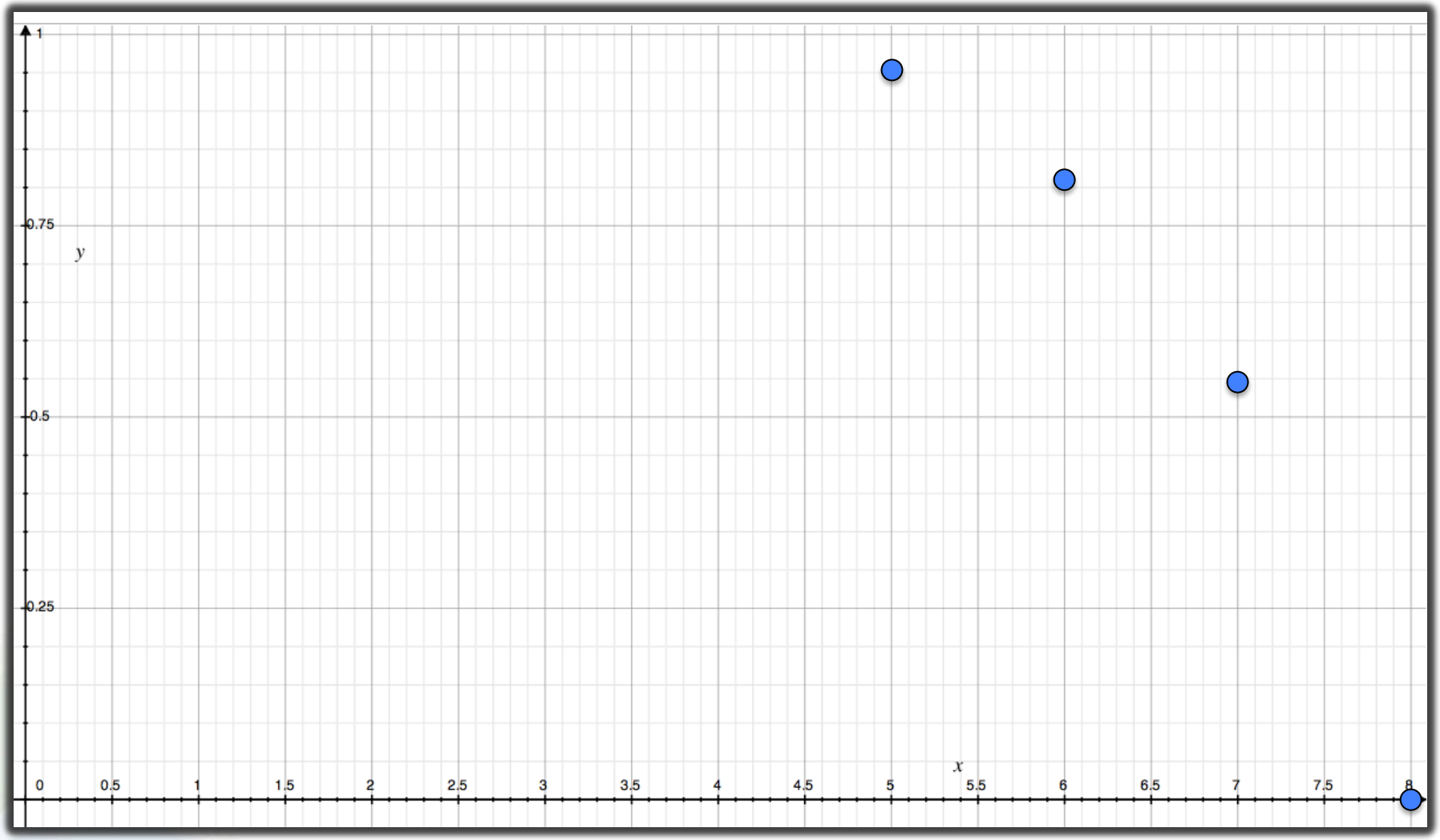




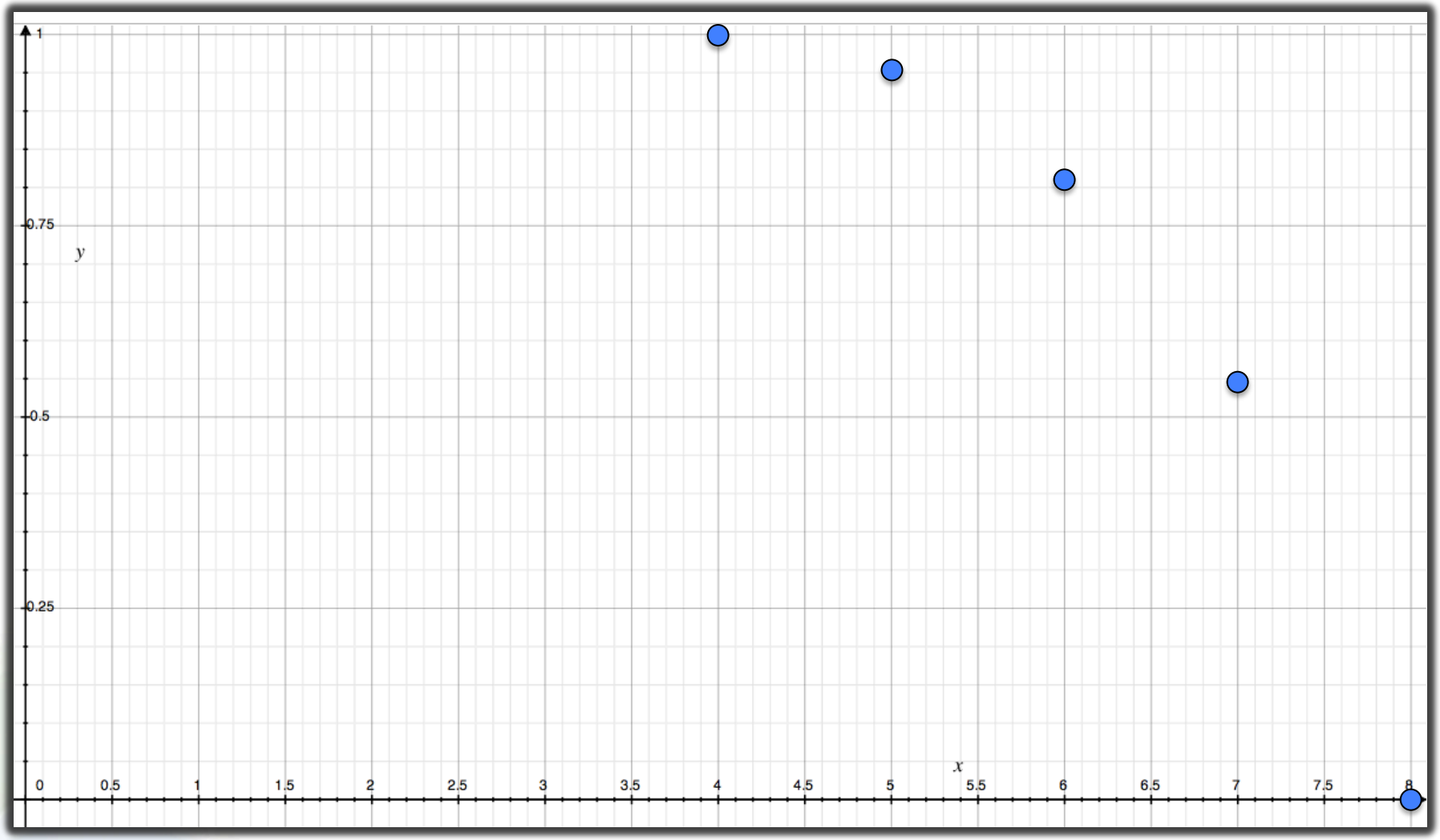
$$-\left(\frac{6}{8}\right) log_2 \left(\frac{6}{8}\right) - \left(\frac{2}{8}\right) log_2 \left(\frac{2}{8}\right) = 0.81$$



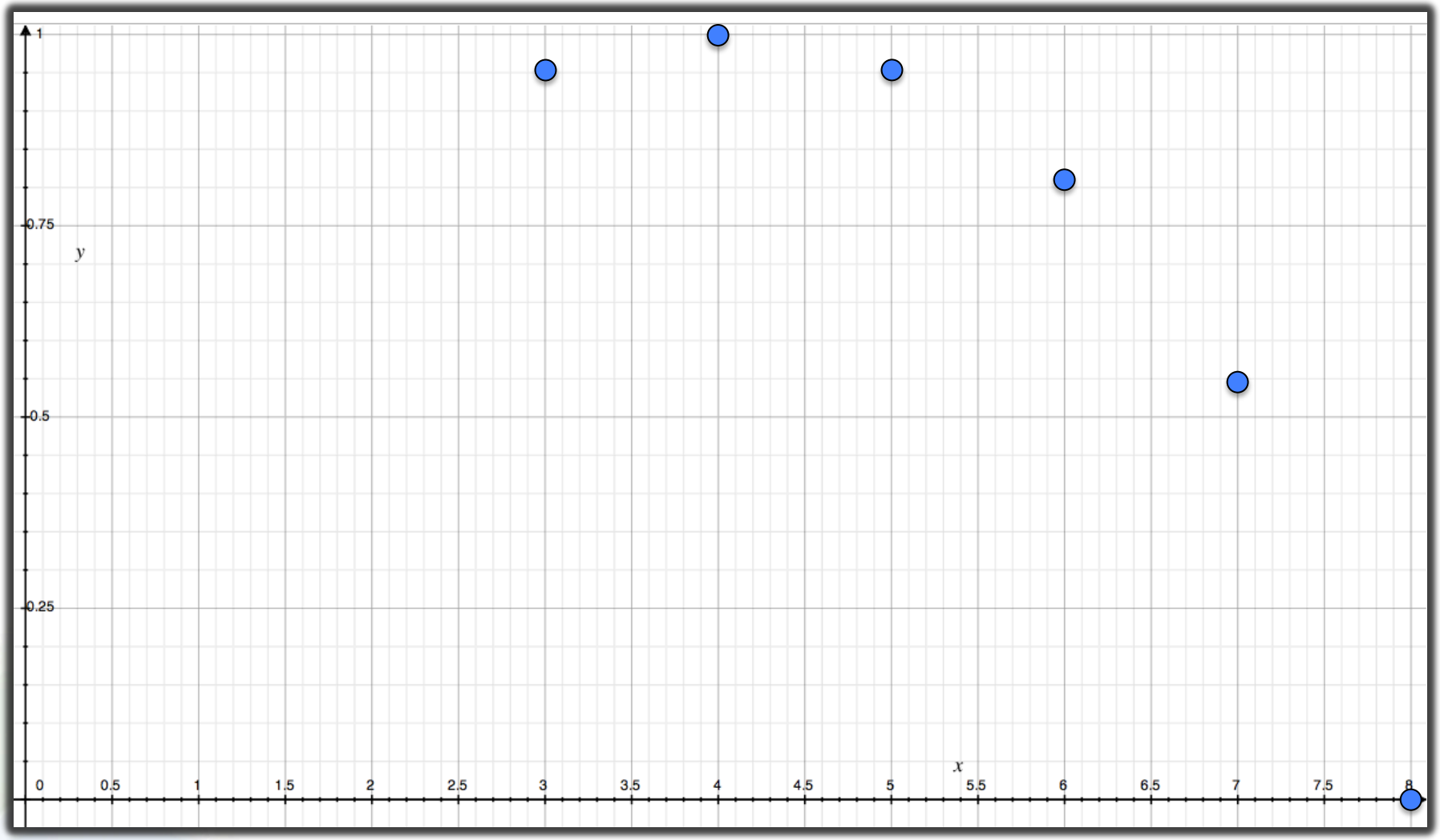
$$-\left(\frac{5}{8}\right) log_2 \left(\frac{5}{8}\right) - \left(\frac{3}{8}\right) log_2 \left(\frac{3}{8}\right) = 0.95$$



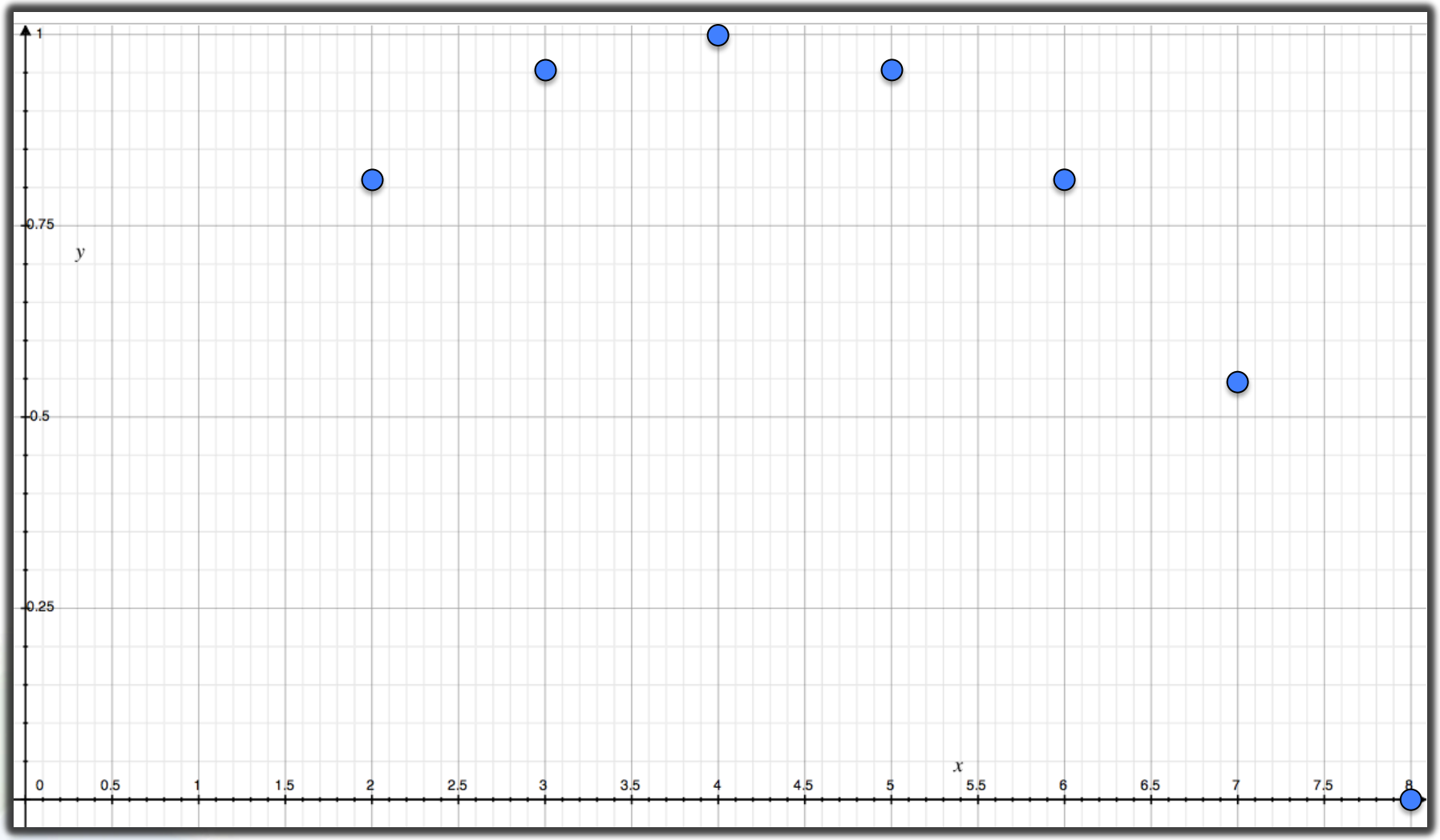
$$-\left(\frac{4}{8}\right) log_2 \left(\frac{4}{8}\right) - \left(\frac{4}{8}\right) log_2 \left(\frac{4}{8}\right) = 1$$



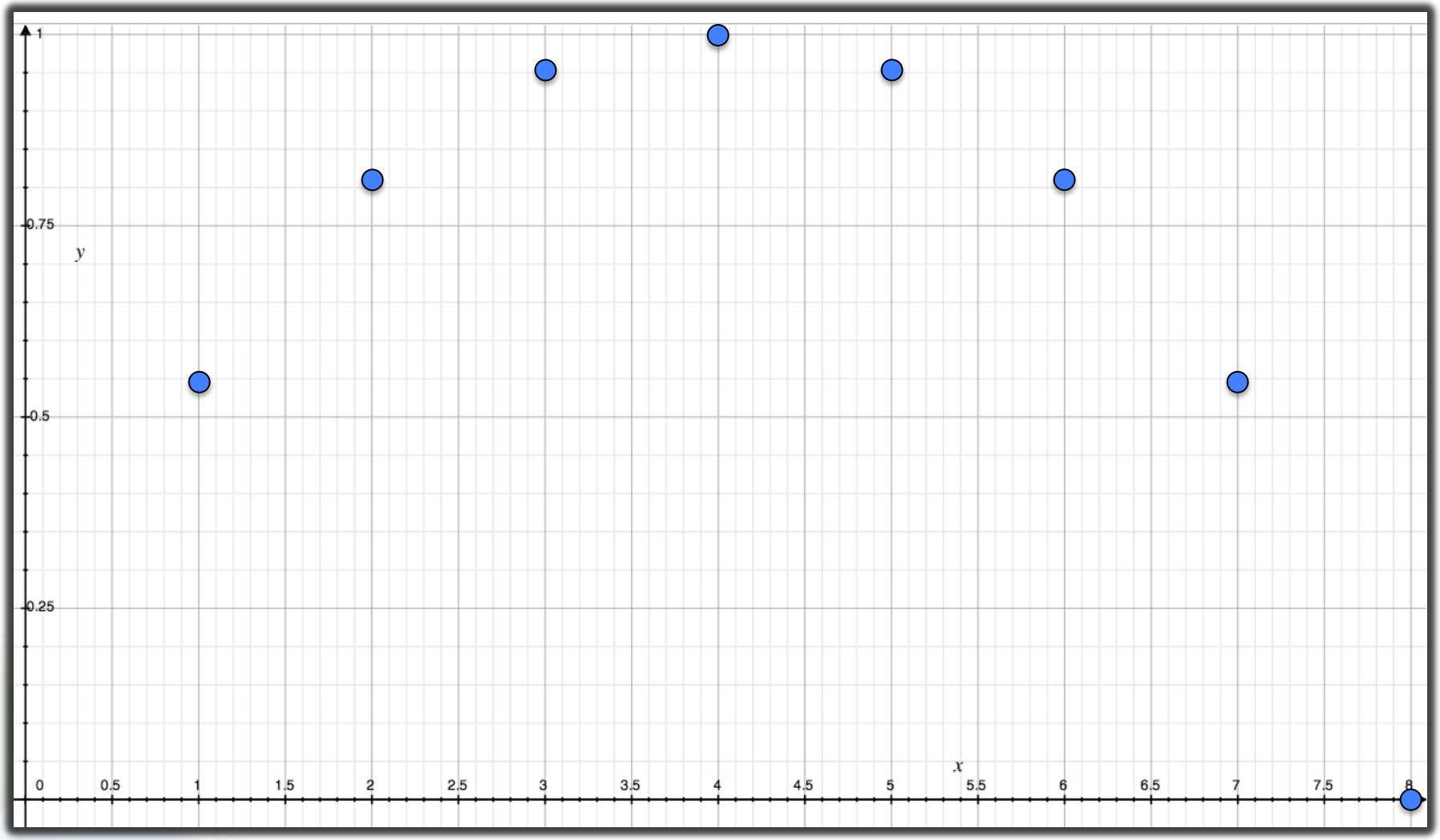
$$-\left(\frac{3}{8}\right) log_2 \left(\frac{3}{8}\right) - \left(\frac{5}{8}\right) log_2 \left(\frac{5}{8}\right) = 0.95$$



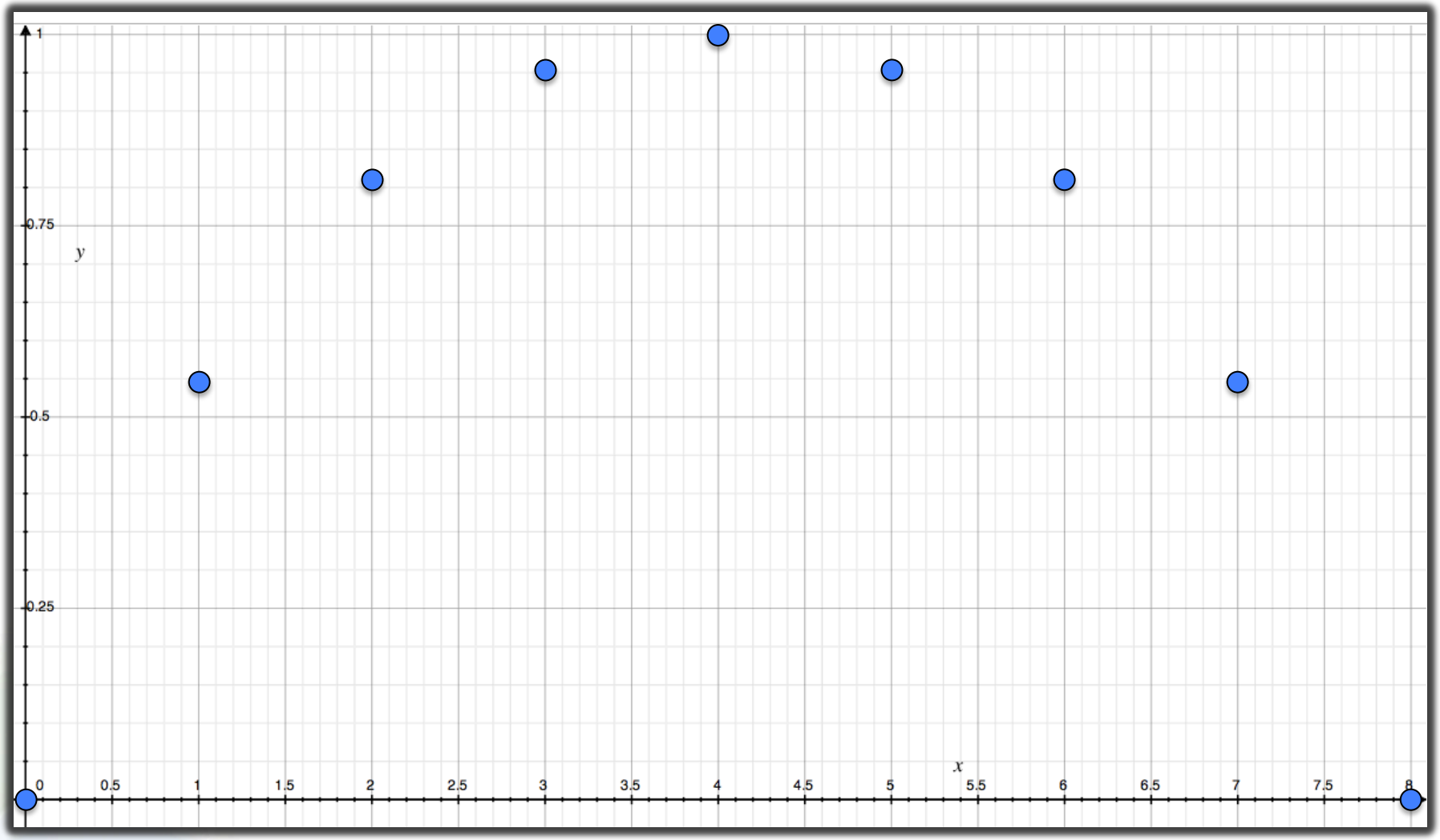
$$-\left(\frac{2}{8}\right) log_2 \left(\frac{2}{8}\right) - \left(\frac{6}{8}\right) log_2 \left(\frac{6}{8}\right) = 0.81$$

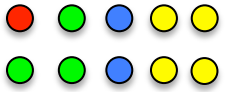


$$-\left(\frac{1}{8}\right) log_2 \left(\frac{1}{8}\right) - \left(\frac{7}{8}\right) log_2 \left(\frac{7}{8}\right) = 0.54$$



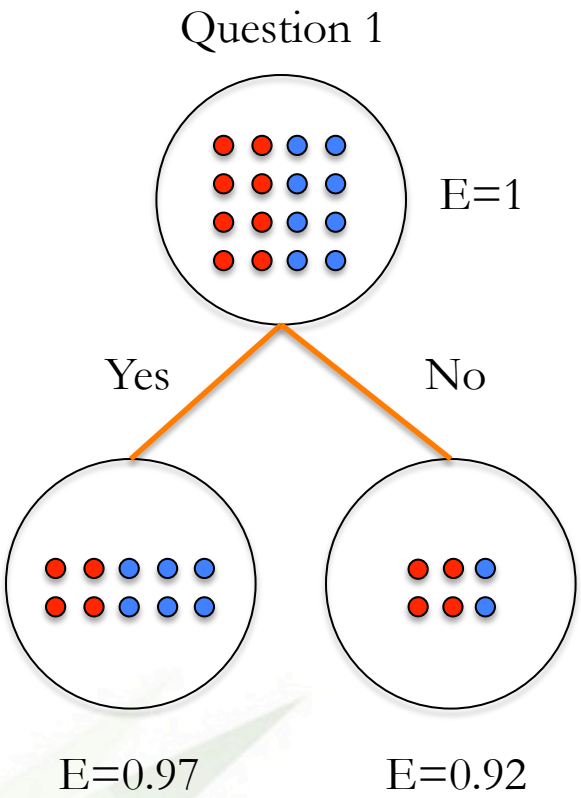
$$-\left(\frac{0}{8}\right) log_2 \left(\frac{0}{8}\right) - \left(\frac{8}{8}\right) log_2 \left(\frac{8}{8}\right) = 0$$

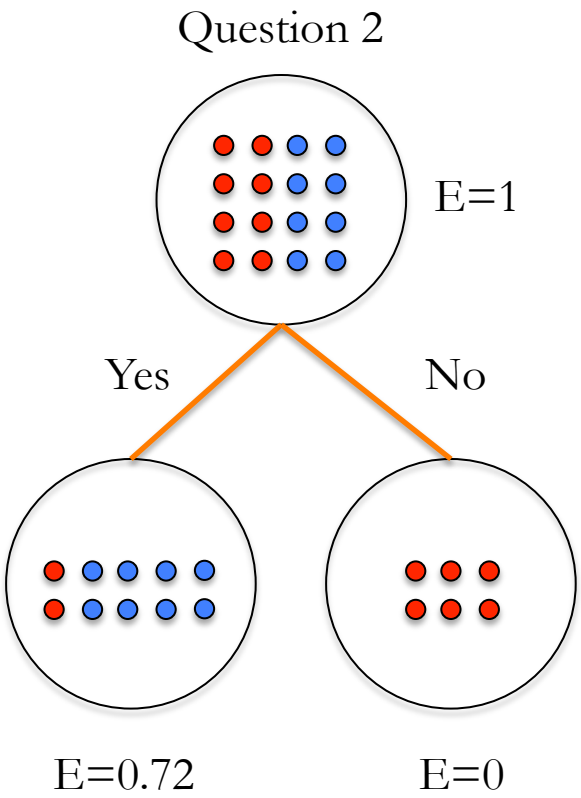




$$y = -\sum_{i=1}^{k} p_i log_k(p_i)$$

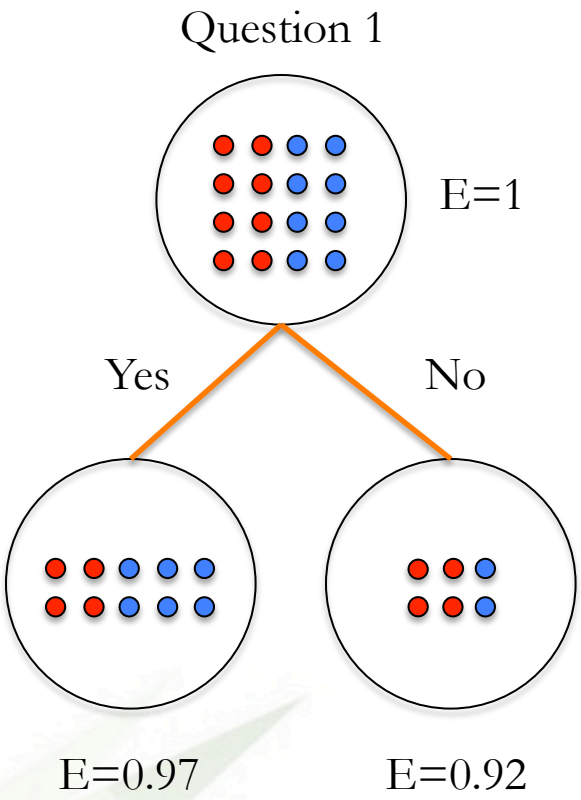
$$y = -\underbrace{\left[\left(\frac{1}{10}\right) log_4 \left(\frac{1}{10}\right)\right]}_{\text{Red}} - \underbrace{\left[\left(\frac{3}{10}\right) log_4 \left(\frac{3}{10}\right)\right]}_{\text{Green}} - \underbrace{\left[\left(\frac{2}{10}\right) log_4 \left(\frac{2}{10}\right)\right]}_{\text{Blue}} - \underbrace{\left[\left(\frac{4}{10}\right) log_4 \left(\frac{4}{10}\right)\right]}_{\text{Yellow}}$$

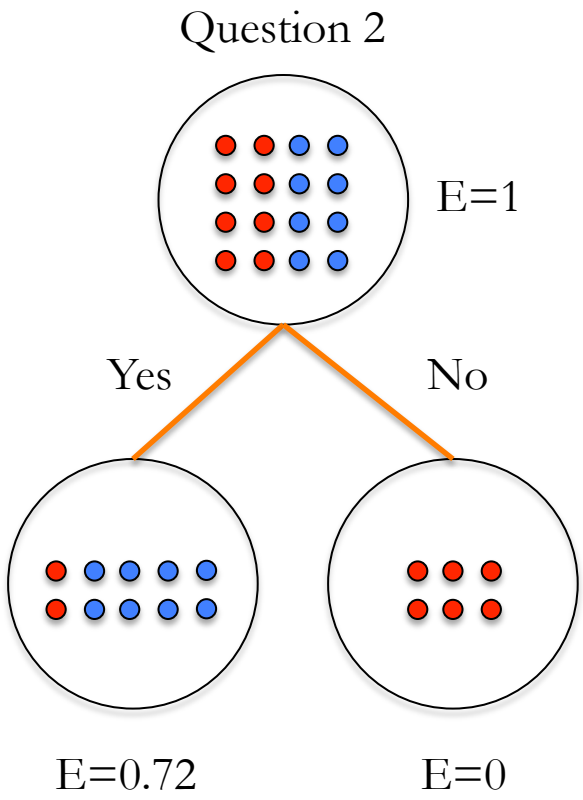




Information Gain

$$G(S,Q) = E(S) - \sum_{i=1}^{k} p_i E(S,Q_i)$$





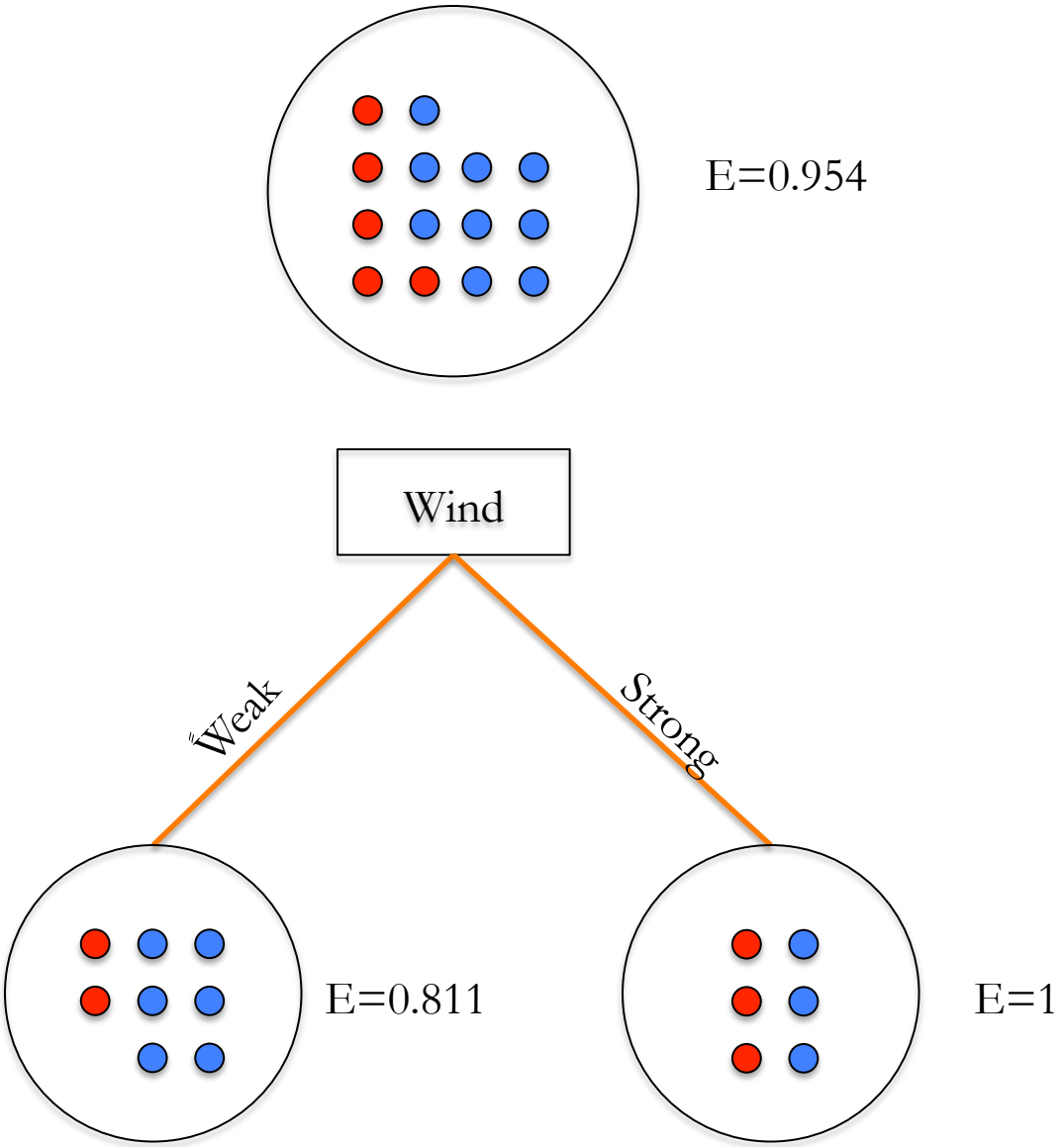
$$G(S, Q_1) = 1 - \left(\frac{10}{16}\right) \times 0.97 - \left(\frac{6}{16}\right) \times 0.92$$

$$G(S, Q_1) = 0.049$$

$$G(S, Q_2) = 1 - \left(\frac{10}{16}\right) \times 0.72 - \left(\frac{6}{16}\right) \times 0$$

$$G(S, Q_2) = 0.55$$

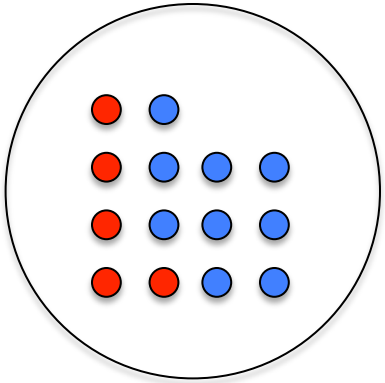
Day	Outlook	Temp	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Weak	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Strong	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No



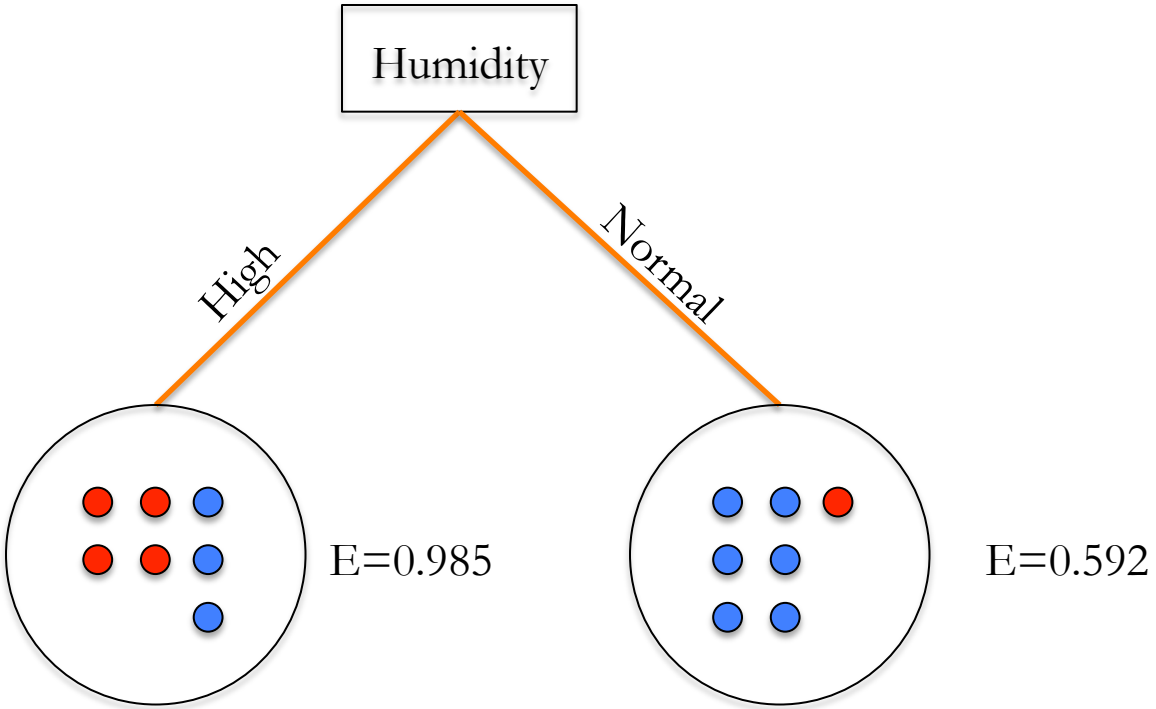
$$G(S, Wind) = 0.954 - \frac{8}{14}0.811 - \frac{6}{14}1$$

$$G(S, Wind) = 0.048$$

$G(S, Wind) = 0.048$



E=0.954

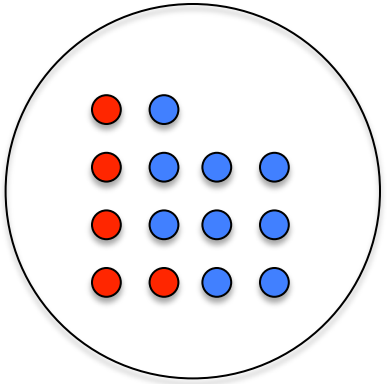


$$G(S, Humidity) = 0.954 - \frac{7}{14}0.985 - \frac{7}{14}0.592$$

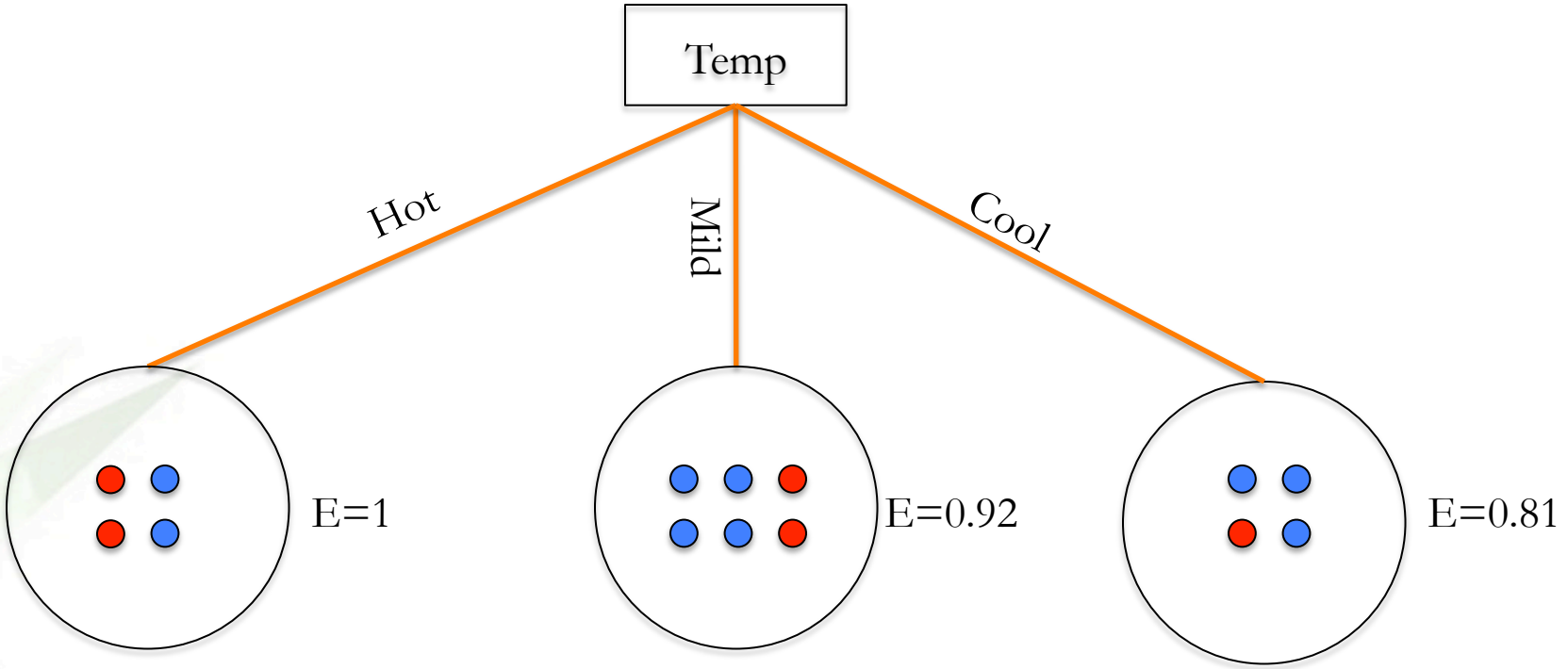
$$G(S, Humidity) = 0.151$$

$G(S, Wind) = 0.048$

$G(S, Humidity) = 0.151$



E=0.954



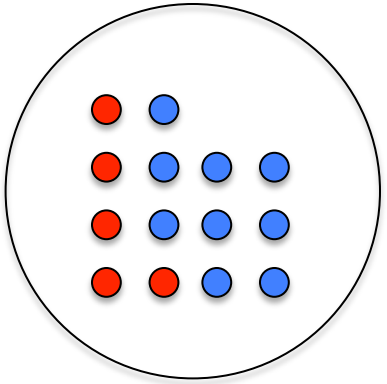
$$G(S, Temp) = 0.954 - \frac{4}{14}1 - \frac{6}{14}0.92 - \frac{4}{14}0.81$$

$$G(S, Temp) = 0.042$$

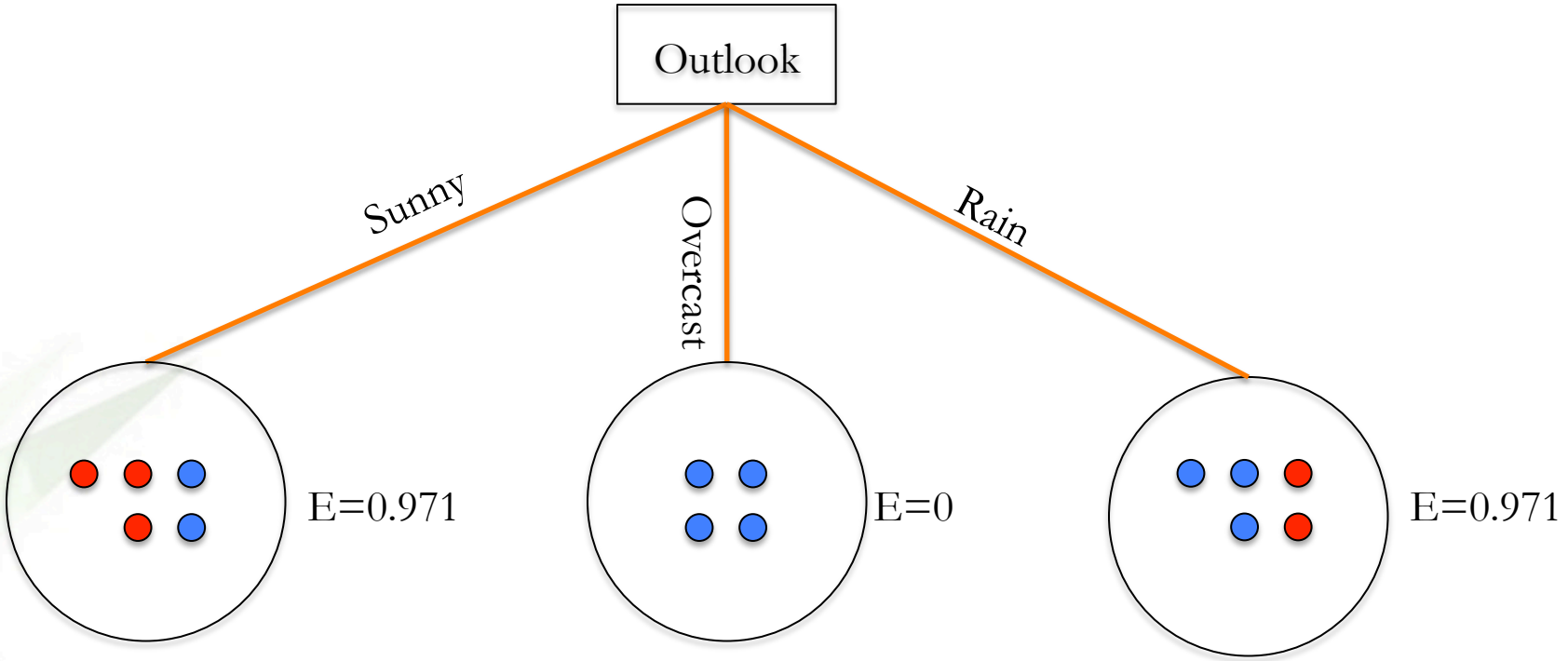
$G(S, Wind) = 0.048$

$G(S, Humidity) = 0.151$

$G(S, Temp) = 0.042$

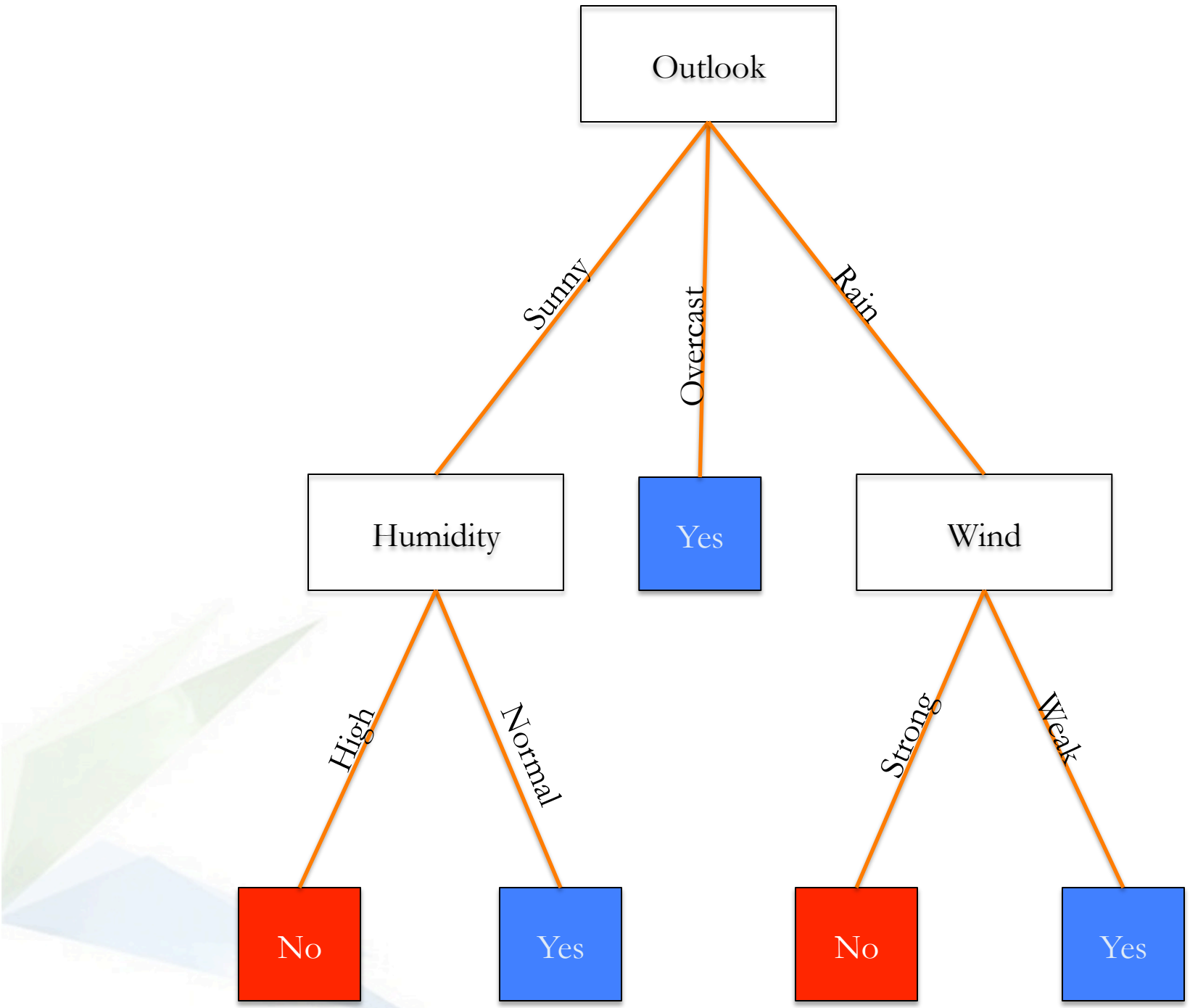


E=0.954



$$G(S, Outlook) = 0.954 - \frac{5}{14}0.971 - \frac{4}{14}0 - \frac{5}{14}0.971$$

$$G(S, Outlook) = 0.247$$


Outlook
Sunny
Overcast
Rain
Humidity
Yes
Wind
High
Normal
Strong
Weak
No
Yes
No
Yes