a)

By chain rule we have:

$$\frac{\partial L}{\partial z_6} = \frac{\partial L}{\partial x_6} \cdot \frac{\partial x_6}{\partial z_6}$$

we first find $\frac{\partial L}{\partial x_6}$:

$$\frac{\partial L(x_6, t)}{\partial x_6} = \frac{\partial}{\partial x_6} \cdot \left(\frac{t}{x_6} - \ln\frac{t}{x_6} - 1\right)$$

$$= -\frac{t}{x_6^2} - \left(\frac{x_6}{t}\right)\left(-\frac{t}{x_6^2}\right)$$

$$= \frac{-t}{x_6^2} + \frac{1}{x_6}$$

$$= \frac{x_6 - t}{x_6^2}$$

Now we find $\frac{\partial x_6}{\partial z_6}$:

We know $x_6 = \rho(z_6)$

$$= z_6^2$$

$$\frac{\partial x_6}{z_6} = \frac{\partial}{\partial z_6} \cdot z_6^2$$

$$= 2z_6$$

Thus $\frac{\partial L}{\partial z_6} = \left(\frac{x_6 - t}{x_6^2}\right)(2z_6)$

$$= 2z_6 \cdot \frac{x_6 - t}{x_6^2}$$

b) To determine update rule for w_{46} we need to find $\frac{\partial L}{\partial w_{46}}$.

By chain rule $\frac{\partial L}{\partial w_{46}} = \frac{\partial L}{\partial z_6} \frac{\partial z_6}{\partial w_{46}}$

we compute $\frac{\partial z_6}{\partial w_{46}}$:

we have $z_6 = w_{36}x_3 + w_{46}x_4 + w_{56}x_5 + b_6$

$$\frac{\partial z_6}{\partial w_{46}} = \frac{\partial}{\partial w_{46}}\left(w_{36}x_3 + w_{46}x_4 + w_{56}x_5 + b_6\right)$$

$$= x_4$$

Thus reusing our a. we have:

$$\frac{\partial L}{\partial w_{46}} = 2z_6 \cdot \frac{x_6 - t}{x_6^2} \cdot x_4$$

Thus by gradient descent our update rule is:

$$w_{46}^{(n+1)} = w_{46}^{(n)} - k\frac{\partial L}{\partial w_{46}}$$

$$= w_{46}^{(n)} - k\left(2z_6 \cdot \frac{x_6 - t}{x_6^2} \cdot x_4\right)$$

where k is step multiplier

$w_{46}^{(n+1)}$ is updated connection weight

$w_{46}^{(n)}$ is old connection weight

c) We need to find $\frac{\partial L}{\partial b_6}$:

$$\frac{\partial L}{\partial b_6} = \frac{\partial L}{\partial z_6} \cdot \frac{\partial z_6}{\partial b_6}$$

Similar to b.

$$\frac{\partial z_6}{\partial b_6} = \frac{\partial}{\partial b_6}\left(w_{36}x_3 + w_{46}x_4 + w_{56}x_5 + b_6\right)$$

$$= 1$$

Thus $\frac{\partial L}{\partial b_6} = 2z_6 \cdot \frac{x_6 - t}{x_6^2}$

Thus by gradient descent our update rule is:

$$b_6^{(n+1)} = b_6^{(n)} - k\frac{\partial L}{\partial b_6}$$

$$= b_6^{(n)} - k\left(2z_6 \cdot \frac{x_6 - t}{x_6^2}\right)$$

where k is step multiplier

$b_6^{(n+1)}$ is updated bias

$b_c^{(n)}$ is old bias

d) By chain rule:

$$\frac{\partial L}{\partial z_3} = \frac{\partial L}{\partial x_3} \cdot \frac{\partial x_3}{\partial z_3}$$

$$= \frac{\partial L}{\partial z_6} \cdot \frac{\partial z_6}{\partial x_3} \cdot \frac{\partial x_3}{\partial z_3}$$

Then $\frac{\partial z_6}{\partial x_3} = \frac{\partial}{\partial x_3} \left(w_{36} x_3 + w_{46} x_4 + w_{56} x_5 + b_6\right)$

$$= \frac{\partial}{\partial x_3} w_{36} x_3 + \frac{\partial}{\partial x_3} w_{46} x_4 + \frac{\partial}{\partial x_3} w_{56} x_5$$

Clearly $\frac{\partial}{\partial x_3} w_{36} x_3 = w_{36}$

For $\frac{\partial}{\partial x_3} w_{46} x_4 = \frac{\partial}{\partial x_3} w_{46}\, \sigma(z_4)$

$$= \frac{\partial}{\partial x_3} w_{46}\, \sigma\left(w_{34} x_3 + w_{14} x_1 + w_{24} x_2 + b_4\right)$$

$$= w_{46}\, \sigma'(z_4)\, w_{34}$$

where $\sigma'(\cdot)$ is derivative of $\sigma(\cdot)$

Similarily

$$\frac{\partial}{\partial x_3} w_{56} x_5 = w_{56}\, \sigma'(z_5)\, w_{35}$$

so $\frac{\partial z_6}{\partial x_3} = w_{36} + w_{46}\, \sigma'(z_4) w_{34} + w_{56}\, \sigma'(z_5) w_{35}$

And $\frac{\partial x_3}{\partial z_3} = \frac{\partial}{\partial z_3} \sigma(z_3)$

$$= \sigma'(z_3)$$

Thus substituting $\frac{\partial L}{\partial z_6}, \frac{\partial z_6}{\partial x_3}, \frac{\partial x_3}{\partial z_3}$

$$\frac{\partial L}{\partial z_3} = \left(2z_6 \cdot \frac{x_6 - t}{x_6^2}\right)\left(w_{36} + w_{46}\,\sigma'(z_4)w_{34} + w_{56}\,\sigma'(z_5)w_{35}\right)\left(\sigma'(z_3)\right)$$