



The diagram shows the equation $[\gamma Q(s_{i+1}, \pi(s_{i+1})|\theta) - Q(s_i, a_i|\theta)] - r$. A large red oval encloses the first part of the equation, $[\gamma Q(s_{i+1}, \pi(s_{i+1})|\theta) - Q(s_i, a_i|\theta)]$, with a red line pointing to the text 'Predicted immediate reward' below it. A smaller red circle encloses the variable r , with a red line pointing to the text 'Actual reward' below it.

$$[\gamma Q(s_{i+1}, \pi(s_{i+1})|\theta) - Q(s_i, a_i|\theta)] - r$$

Predicted immediate reward

Actual
reward