

$$u(k) = \underbrace{m_t(k)}_{\text{expected reward}} + \beta \underbrace{v_t(k)}_{\text{uncertainty}} + \lambda \underbrace{I(\{k, r\}; r^*)}_{\text{information gain about max}}$$