$$p(a_t = k) = \frac{\exp[(\propto m_t(k) + \beta v_t(k) + \lambda I(\{k, r\}; r^*))/\tau]}{\sum_{k' \in A} \exp[(m_t(k') + \beta v_t(k') + \lambda I(\{k', r\}; r^*))/\tau]}$$