Blackwell Optimal Policies

⊛ If (g^*, h^*) is the common gain-bias pair that is common to all BOP, then they satisfy these nested optimizations:

$$g^*(i) = \min_{a \in A(i)} \sum_{j}^{|S|} p_{ij}(a)\, g^*(j), \quad i \in S.$$

&

$$g^*(i) + h^*(i) = \min_{a \in \bar{A}(i)} \left\{ r(i,a) + \sum_j p_{ij}(a)\, h(j) \right\}, \quad i \in S$$

where $\bar{A}$ is those actions that attain minima in first optimization problem.

⊛ The above are interpreted in the class via this inequality (following Laurent Series expansion & looking @ dominating terms one after another)

$$0 \le r_\pi - r_{\pi^*} + \alpha(P_\pi - P_{\pi^*})\left(\frac{1}{(1-\alpha)} g^* + h^* + O(|1-\alpha|)\right)$$

⊛ Converse also holds.