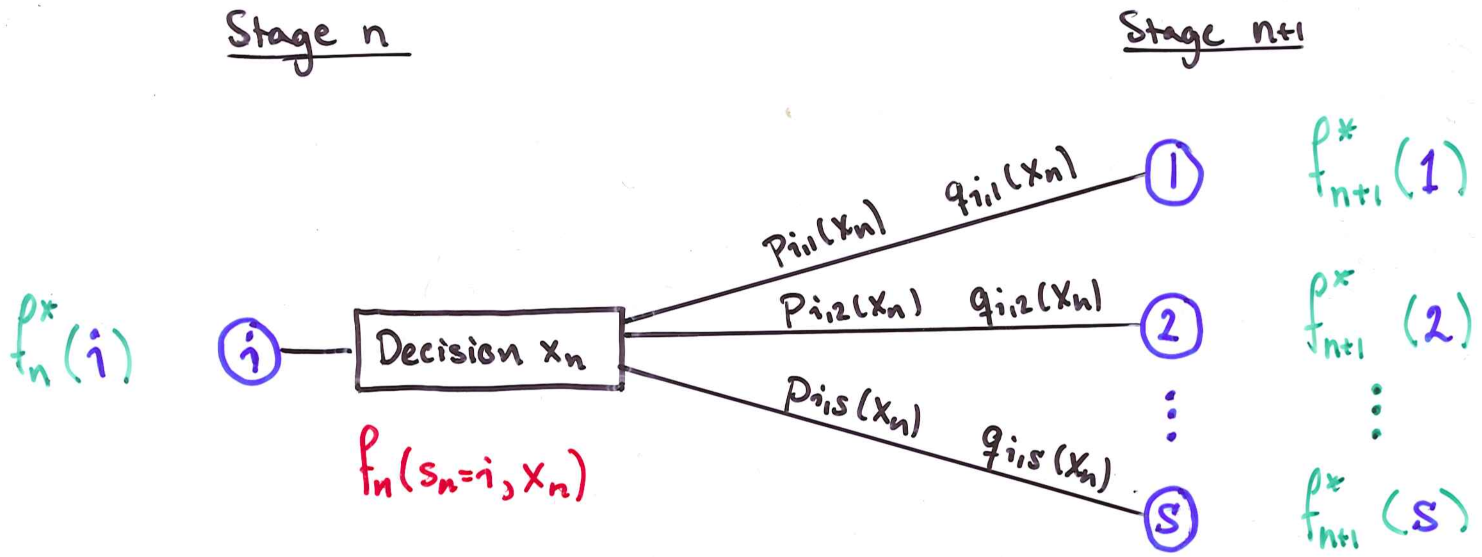
Probabilistic Dynamic programming



where

$p_{ij}(x_n)$ = probability of transition $i \to j$ given decision x_n

$q_{ij}(x_n)$ = expected cost given transition $i \to j$ and decision x_n

$f_n^*(s_n)$ = expected cost for stages $n \dots N$ starting at state s_n and using optimal decision policy.

$f_n(s_n, x_n)$ = expected cost for stages $n \dots N$ starting at state s_n making decision x_n and after that optimal policy.

$$\Rightarrow f_n(s_n=i, x_n) = E[\text{cost} \mid x_n, s_n=i] = \sum_{j=1}^{s} E[\text{cost} \mid x_n, s_n=i, s_{n+1}=j]\, P(s_{n+1}=j \mid s_n=i, x_n)$$

$$= \sum_{j=1}^{s} \left[q_{ij}(x_n) + f_{n+1}^*(s_{n+1}=j) \right] p_{ij}(x_n)$$

$$= C_{i,x_n} + \sum_{j=1}^{s} p_{ij}(x_n)\, f_{n+1}^*(s_{n+1}=j)$$

where $C_{i,x_n} = \sum_{j=1}^{s} p_{ij}(x_n)\, q_{ij}(x_n)$ = expected cost when decision x_n is taken in state i

then $$f_n^*(s_n=i) = \min_{x_n \in F_n(s_n)} f_n(s_n=i, x_n) = \min_{x_n \in F_n(s_n)} \left\{ C_{i,x_n} + \sum_{j=1}^{s} p_{ij}(x_n)\, f_{n+1}^*(s_{n+1}=j) \right\}$$