$s' =$ Policy

$V(s) = \sum_{s',r} p(s',r \mid s, \pi(s))[r + \gamma V(s')]$

Initially, $V(s') = 0$

$\cancel{V(s)} = (0.8)(40) + (0.2)(30)$

$V(\text{prod}) = 32 + 6 = \underline{38}$

$V(\text{exh}) = \cancel{2}(1)(20) = \underline{20}$

$V(\text{fit}) = \underline{0}$

$\pi(s) = \arg\max_a \sum$

$\pi(\text{prod}) = \cancel{\max}\Big[\overbrace{(0.8)(40 + (0.86)(20))}^{\text{Coding}},$
$(0.2)(30 + 0.86)(38))\Big]$

$= \cancel{\max}[45.76, + 12.536] = 58.296$

$\pi(\text{prod})$ exercise $= [\cancel{(100)}(1)(-10 + 0.86(0)]$
$= -8.6$

$\cancel{\pi(\text{prod}})$

$\therefore \pi(\text{prod}) = \max(\cancel{45}\ 58.29, -8.6)$
$= 58.29 = \underline{\text{coding}}$

$\pi(\text{exhausted})$ coding $= (1)\cancel{(100)}(20 + (0.86)(20))$
$= \underline{37.2}$

$\pi(\text{exhausted})$ exercise $= [(0.5)(-10 + (0.86)(38)) + (0.5)(-10 + (0.86)(20)]$
$= 11.34 + 3.6 = \underline{\underline{14.94}}$

$$\pi(\text{exhausted})_{\text{rest}} = \cancel{20}[(0.2)(0+(0.86)(20)) + (0.8)(0+(38)(0\ldots$$

$$= \cancel{0.344+} \; 3.44 + 26.144 = \underline{29.584}$$

$\therefore \pi(\text{exhausted}) = \max(37.2, 14, 29.584)$

$= 37.2 = \text{coding}$

$$\pi(\text{fit})_{\text{coding}} = (1)[(100) + (0.86)(20)]$$

$$= 117.2$$

$\therefore \pi(\text{fit}) = \text{coding}$

$\therefore$ After 1^{st} iteration,

$V(\text{fit}) = 0$ $\quad V(\text{exhausted}) = 38$

$V(\text{prod}) = 38$

$\pi(\text{fit}) = \text{coding}$, $\pi(\text{prod}) = \text{coding}$

$\pi(\text{exhausted}) = \text{coding}$