1 Task

Action set consists of choosing one of four tasks or doing nothing {choose-1,choose-2,choose-3,choose-4,do-nothing}. We suppose that the student will always try to solve something until getting all the tasks solved. State set consists of all possible combinations of solved tasks, such as each task, each pair, each three or all. Overall we have $1 + 4 + C_4^2 + C_4^3 + 1 = 16$ states. All possible transitions with rewards depicted in the following table.

Current state	Next state	Action	Transition probability	Expected reward
$\emptyset$	$\emptyset$	choose-1	$1 - p_1 = 0.7$	0
$\emptyset$	$\{1\}$	choose-1	$p_1 = 0.3$	$r_1 = 4$
$\emptyset$	$\emptyset$	choose-2	$1 - p_2 = 0.5$	0
$\emptyset$	$\{2\}$	choose-2	$p_2 = 0.5$	$r_2 = 6$
$\emptyset$	$\emptyset$	choose-3	$1 - p_3 = 0.2$	0
$\emptyset$	$\{3\}$	choose-3	$p_3 = 0.8$	$r_3 = 3$
$\emptyset$	$\emptyset$	choose-4	$1 - p_4 = 0.8$	0
$\emptyset$	$\{4\}$	choose-4	$p_4 = 0.2$	$r_4 = 9$
$\{1\}$	$\{1\}$	choose-2	$1 - p_2 = 0.5$	0
$\{1\}$	$\{1,2\}$	choose-2	$p_2 = 0.5$	$r_2 = 6$
$\{1\}$	$\{1\}$	choose-3	$1 - p_3 = 0.2$	0
$\{1\}$	$\{1,3\}$	choose-3	$p_3 = 0.8$	$r_3 = 3$
$\{1\}$	$\{1\}$	choose-4	$1 - p_4 = 0.8$	0
$\{1\}$	$\{1,4\}$	choose-4	$p_4 = 0.2$	$r_4 = 9$
$\{2\}$	$\{2\}$	choose-1	$1 - p_1 = 0.7$	0
$\{2\}$	$\{1,2\}$	choose-1	$p_1 = 0.3$	$r_1 = 4$
$\{2\}$	$\{2\}$	choose-3	$1 - p_3 = 0.2$	0
$\{2\}$	$\{2,3\}$	choose-3	$p_3 = 0.8$	$r_3 = 3$
$\{2\}$	$\{2\}$	choose-4	$1 - p_4 = 0.8$	0
$\{2\}$	$\{2,4\}$	choose-4	$p_4 = 0.2$	$r_4 = 9$
$\{3\}$	$\{3\}$	choose-2	$1 - p_2 = 0.5$	0
$\{3\}$	$\{3,2\}$	choose-2	$p_2 = 0.5$	$r_2 = 6$
$\{3\}$	$\{3\}$	choose-1	$1 - p_1 = 0.7$	0
$\{3\}$	$\{1,3\}$	choose-1	$p_1 = 0.3$	$r_1 = 4$
$\{3\}$	$\{3\}$	choose-4	$1 - p_4 = 0.8$	0
$\{3\}$	$\{3,4\}$	choose-4	$p_4 = 0.2$	$r_4 = 9$
$\{4\}$	$\{4\}$	choose-2	$1 - p_2 = 0.5$	0
$\{4\}$	$\{4,2\}$	choose-2	$p_2 = 0.5$	$r_2 = 6$
$\{4\}$	$\{4\}$	choose-3	$1 - p_3 = 0.2$	0
$\{4\}$	$\{4,3\}$	choose-3	$p_3 = 0.8$	$r_3 = 3$
$\{4\}$	$\{4\}$	choose-1	$1 - p_1 = 0.7$	0
$\{4\}$	$\{1,4\}$	choose-1	$p_1 = 0.3$	$r_1 = 4$

Current state	Next state	Action	Transition probability	Expected reward
{1,2}	{1,2}	choose-3	$1 - p_3 = 0.2$	0
{1,2}	{1,2,3}	choose-3	$p_3 = 0.8$	$r_3 = 3$
{1,2}	{1,2}	choose-4	$1 - p_4 = 0.8$	0
{1,2}	{1,2,4}	choose-4	$p_4 = 0.2$	$r_4 = 9$
{1,3}	{1,3}	choose-2	$1 - p_2 = 0.5$	0
{1,3}	{1,2,3}	choose-2	$p_2 = 0.5$	$r_2 = 6$
{1,3}	{1,3}	choose-4	$1 - p_4 = 0.8$	0
{1,3}	{1,3,4}	choose-4	$p_4 = 0.2$	$r_4 = 9$
{1,4}	{1,4}	choose-3	$1 - p_3 = 0.2$	0
{1,4}	{1,3,4}	choose-3	$p_3 = 0.8$	$r_3 = 3$
{1,4}	{1,4}	choose-2	$1 - p_2 = 0.5$	0
{1,4}	{1,2,4}	choose-2	$p_2 = 0.5$	$r_2 = 6$
{2,3}	{2,3}	choose-1	$1 - p_1 = 0.7$	0
{2,3}	{1,2,3}	choose-1	$p_1 = 0.3$	$r_1 = 4$
{2,3}	{2,3}	choose-4	$1 - p_4 = 0.8$	0
{2,3}	{2,3,4}	choose-4	$p_4 = 0.2$	$r_4 = 9$
{2,4}	{2,4}	choose-3	$1 - p_3 = 0.2$	0
{2,4}	{2,3,4}	choose-3	$p_3 = 0.8$	$r_3 = 3$
{2,4}	{2,4}	choose-1	$1 - p_1 = 0.7$	0
{2,4}	{1,2,4}	choose-1	$p_1 = 0.3$	$r_1 = 4$
{3,4}	{3,4}	choose-1	$1 - p_1 = 0.7$	0
{3,4}	{1,3,4}	choose-1	$p_1 = 0.3$	$r_1 = 4$
{3,4}	{3,4}	choose-2	$1 - p_2 = 0.5$	0
{3,4}	{2,3,4}	choose-2	$p_2 = 0.5$	$r_2 = 6$
{1,2,3}	{1,2,3}	choose-4	$1 - p_4 = 0.8$	0
{1,2,3}	{1,2,3,4}	choose-4	$p_4 = 0.2$	$r_4 = 9$
{1,2,4}	{1,2,4}	choose-3	$1 - p_3 = 0.2$	0
{1,2,4}	{1,2,3,4}	choose-3	$p_3 = 0.8$	$r_3 = 3$
{1,3,4}	{1,3,4}	choose-2	$1 - p_2 = 0.5$	0
{1,3,4}	{1,2,3,4}	choose-2	$p_2 = 0.5$	$r_2 = 6$
{2,3,4}	{2,3,4}	choose-1	$1 - p_1 = 0.7$	0
{2,3,4}	{1,2,3,4}	choose-1	$p_1 = 0.3$	$r_1 = 4$
{1,2,3,4}	{1,2,3,4}	do-nothing	1	0

2 Task

In order to model the risk of not passing the exam we can find out probability of not passing the exam. Fail means < 50% of points (N=4 - number of attempts). According to the points that can be got for every task total number of achievable points is 22, half - 11. Student gets ≥ 11 if he does any three tasks or task 4 and any of the remaining. Then

$$P\{not-passing-the-exam\} = 1 - P\{passing-the-exam\} =$$

$$= 1 - P\{solve-4-tasks\} - P\{solve-3-tasks\} - P\{solve-4^{th}-and-any\} =$$

$$= 1 - \prod_{i=1}^{4} p_i - C_4^3 \prod_{i=j_1}^{j_3} p_{j_i} - C_3^1 p_i p_4$$

.

3 Task

Policy π_A - to work in order of increasing difficulty is chain of actions starting in state $\varnothing$:

$$choose-3 \rightarrow choose-2 \rightarrow choose-1 \rightarrow choose-4$$

In order to find expected return we will use formula

$$V^\pi(s) = \sum_a \pi(s,a) \sum_{s'} P_{ss'}^a [R_{ss'}^a + \gamma V^\pi(s')]$$

We put $\gamma = 1$ as the student will always finish in {1,2,3,4} and get the following value and we suppose that the student will return to the same task until solving it:

$$V^{\pi_A}(s) = p_3(r_3 + p_2(r_2 + p_1(r_1 + p_4 r_4))) = 0.8(3 + 0.5(6 + 0.3(4 + 0.2*9))) = 5.496$$

For policy π_B - to work in order of decreasing gain, by the same considerations we get:

$$choose-4 \rightarrow choose-2 \rightarrow choose-1 \rightarrow choose-3$$

$$V^{\pi_B}(s) = p_4(r_4 + p_2(r_2 + p_1(r_1 + p_3 r_3))) = 0.2(9 + 0.5(6 + 0.3(4 + 0.8*3))) = 2.592$$

4 Task

Improved policy π_C

$$choose-3 \rightarrow choose-2 \rightarrow choose-4 \rightarrow choose-1$$

$$V^{\pi_C}(s) = p_3(r_3 + p_2(r_2 + p_4(r_4 + p_1 r_1))) = 0.8(3 + 0.5(6 + 0.2(9 + 0.3*4))) = 5.616$$

5 Task

As an example of a process model where Markov assumption is not justified can be considered a mechanism that can be broken and once when it is broken it begins to repair. There are two states 'working' and 'repairing' and we do not know the possibility of changing the states. But if we will introduce one more state 'checking', that will fix probabilities of transitions, because the check will be hold with some known distribution and after checking the mechanism will either be repairing or working with known probability. Hence we will have set of states, transition probabilities and this makes the process where Markov assumption is justified.