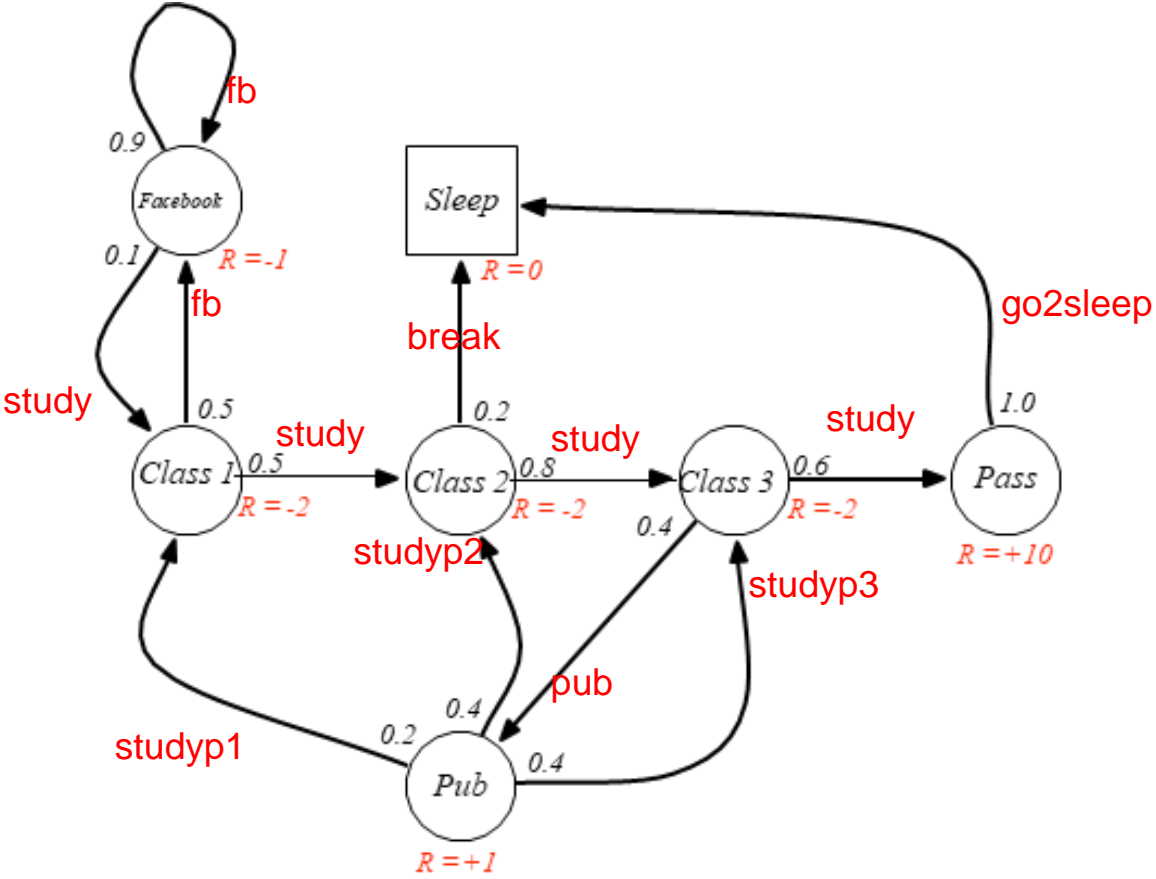


Assignment 1. : student MRP example (10 points)

- 1. Evaluate value of each State(i.e. $V(s)$) given below.
- 2. Obtain action values, $q(s,a)$ for each arrow in the given example.
- 3. How many iterations do we require to obtain the final value both for $V(s)$ and $q(s,a)$.

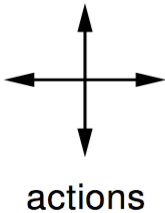


Assignment 2. : Value iteration in the Small Gridworld (10 points)

Write algorithm code for a synchronous Value iteration agent.

Value iteration computes k-step estimates of the state values, $V(k)$ for $k=0, 1, 5, 10$ and ∞ .

			+1 Goal with (award)
			-1 Goal with (penalty)



- Initial policy : random
- Tabular MDP with 16 states
- Action : agent allow to move 4 normal direction
- Discount = none (gamma=1.0)
- reward = - 0.1 on all transition
- Two terminal states : one +1 award, the other -1 penalty

Assignment 2 : Evaluating a Random Policy in the Small Gridworld

V k=0

0	0	0	+1
0	0	0	-1
0	0	0	0
0	0	0	0

V k=1

			+1
			-1

V k=5

			+1
			-1

V k=10

			+1 Goal with (award)
			-1 Goal with (penalty)

V k= ∞

			+1 Goal with (award)
			-1 Goal with (penalty)