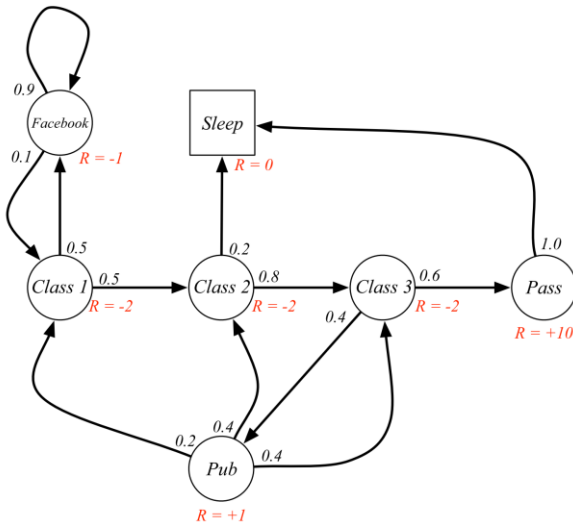


# 1 Exercices

## 1.1 Exercice 1

Soit l'exemple « Student Markov Chain » avec le MDP :



1. Donner des épisodes en prenant l'état « Class 1 » comme état initial et « Sleep » comme état final.
2. Donner la matrice de transition de ce réseau (sans considérer les récompenses).
3. Evaluer la récompense des épisodes donnés de la question 1.

## 1.2 Exercice 2

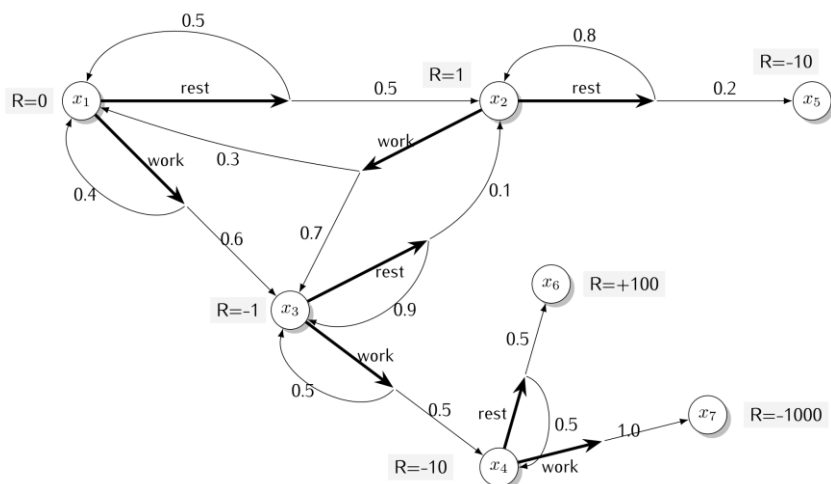
Soit l'environnement régissant les états d'un véhicule :

- trois états du moteur : normal, chaud, surchauffe
- deux actions : lent, vite
- en allant plus vite, on double la récompense +1 pour les actions lent, +2 pour action vite, -10 pour atteindre un état surchauffe.

Donner une description formelle sous forme d'un MDP ?

## 1.3 Exercice 3

Soit le réseau MDP :



1. Si un agent effectue la séquence d'actions <work,work,rest> à partir de l'état  $x_1$ , dans quels états peut-il se retrouver et avec quelle probabilité ?
2. Définissez une politique  $\pi$  et écrivez le système linéaire qui définit  $v_\pi$  pour cette politique. Votre réponse finale devrait être prête pour faire l'application numérique si on avait accès à un solveur de système linéaire.