

# Correction Exam IA/APP

①

## Apprentissage par renforcement

1)  $S = \{P_1, P_2, P_3, P_5, P_5, P_6\}$

• Pour l'ensemble des actions : on considère qu'il y a une action possible entre chaque couple d'états  $P_i \rightarrow P_j$  mais que si l'action est illicite elle n'aboutit à rien et l'agent reste sur place (avec une sanction)

### • Récompense

- action licite -1 (consommation d'énergie)
- action interdite -10 (sanction)
- action  $\rightarrow P_6$  10 (nourriture)

2) Voir programme

3) Paramètre  $\gamma$  : prise en compte du long terme

«  $\epsilon$  (taux expls du programme) : répartition exploration/exploitation