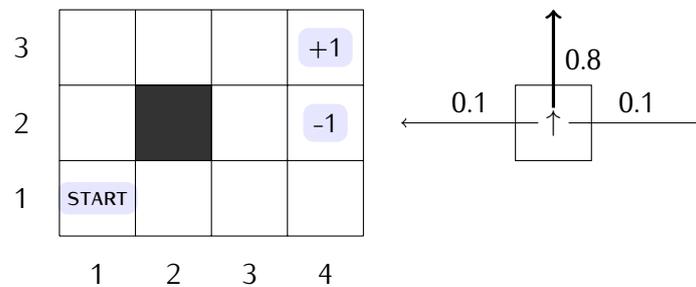


## Résoudre un PDM



- **modèle de transition** : Il y a un état initial noté START et deux états finaux en (4,3) et (4,2). L'action désirée est réalisée avec une probabilité de 0.8, mais avec une probabilité de 0.2, l'agent se déplace à angle droit de la direction souhaitée. Si l'action de l'agent l'emporte vers un mur (bord de la grille ou obstacle en (2,2), l'agent reste à sa place.
- **modèle de récompenses** : la récompense est de -0.04 pour chaque état sauf pour les deux états terminaux qui donnent +1 ou -1.

1. Quels états peuvent être atteints depuis l'état initial avec la séquence  $[\uparrow, \uparrow, \rightarrow, \rightarrow, \rightarrow]$  ?
2. Donnez la probabilité d'être dans chacun des états après la séquence séquence  $[\uparrow, \uparrow, \rightarrow]$  (donnez l'expression).
3. Ecrivez l'équation de Bellman optimale pour l'état (1,1). Evidemment, on ne va pas résoudre l'équation à la main.

En dehors du cours :

4. Ecrivez un programme pour estimer l'utilité des états non terminaux avec un  $\gamma$  donné.
5. Selon la valeur  $r$  donnée à chaque état non-terminal, on peut observer différents comportements. A l'aide d'un programme, calculez les valeurs seuils et donnez les différentes politiques optimales.