

□ **Consistance de la relaxation d'heuristiques** – Soit P_{rel} une relaxation d'un problème de recherche. Par théorème, on a :

$$h(s) = \text{FutureCost}_{\text{rel}}(s) \implies h(s) \text{ constante}$$

□ **Compromis lors du choix d'heuristique** – Le choix d'heuristique se repose sur un compromis entre :

- Complexité de calcul : $h(s) = \text{FutureCost}_{\text{rel}}(s)$ doit être facile à calculer. De manière préférable, cette fonction peut s'exprimer de manière explicite et elle permet de diviser le problème en sous-parties indépendantes.
- Approximation adéquate : l'heuristique $h(s)$ devrait être assez proche de $\text{FutureCost}(s)$ ce qui veut dire qu'il ne faudrait pas enlever trop de contraintes.

□ **Heuristique max** – Soient $h_1(s)$, $h_2(s)$ deux heuristiques. On a la propriété suivante :

$$h_1(s), h_2(s) \text{ constante} \implies h(s) = \max\{h_1(s), h_2(s)\} \text{ constante}$$

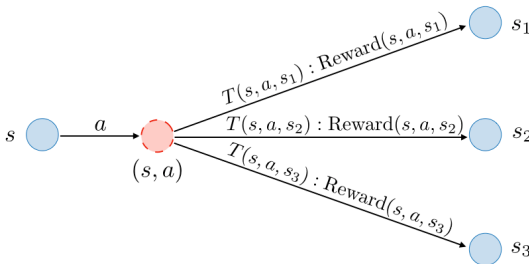
2.2 Processus de décision markovien

Dans cette section, on suppose qu'effectuer l'action a à partir de l'état s peut mener de manière probabiliste à plusieurs états s'_1, s'_2, \dots . Dans le but de trouver ce qu'il faudrait faire entre un état initial et un état final, on souhaite trouver une stratégie maximisant la quantité des récompenses en utilisant un outil adapté à l'imprévisibilité et l'incertitude : les processus de décision markoviens.

2.2.1 Notations

□ **Définition** – l'objectif d'un processus de décision markovien (en anglais *Markov decision process* ou *MDP*) est de maximiser la quantité de récompenses. Un tel problème est défini par :

- un état de départ s_{start}
- l'ensemble des actions $\text{Actions}(s)$ pouvant être effectuées à partir de l'état s
- la probabilité de transition $T(s, a, s')$ de l'état s vers l'état s' après avoir pris l'action a
- la récompense $\text{Reward}(s, a, s')$ pour être passé de l'état s à l'état s' après avoir pris l'action a
- la connaissance d'avoir atteint ou non un état final $\text{IsEnd}(s)$
- un facteur de dévaluation $0 \leq \gamma \leq 1$



□ **Probabilités de transition** – La probabilité de transition $T(s, a, s')$ représente la probabilité de transitionner vers l'état s' après avoir effectué l'action a en étant dans l'état s . Chaque $s' \mapsto T(s, a, s')$ est une loi de probabilité :