

Les modèles de Markov dans la 'vraie' vie

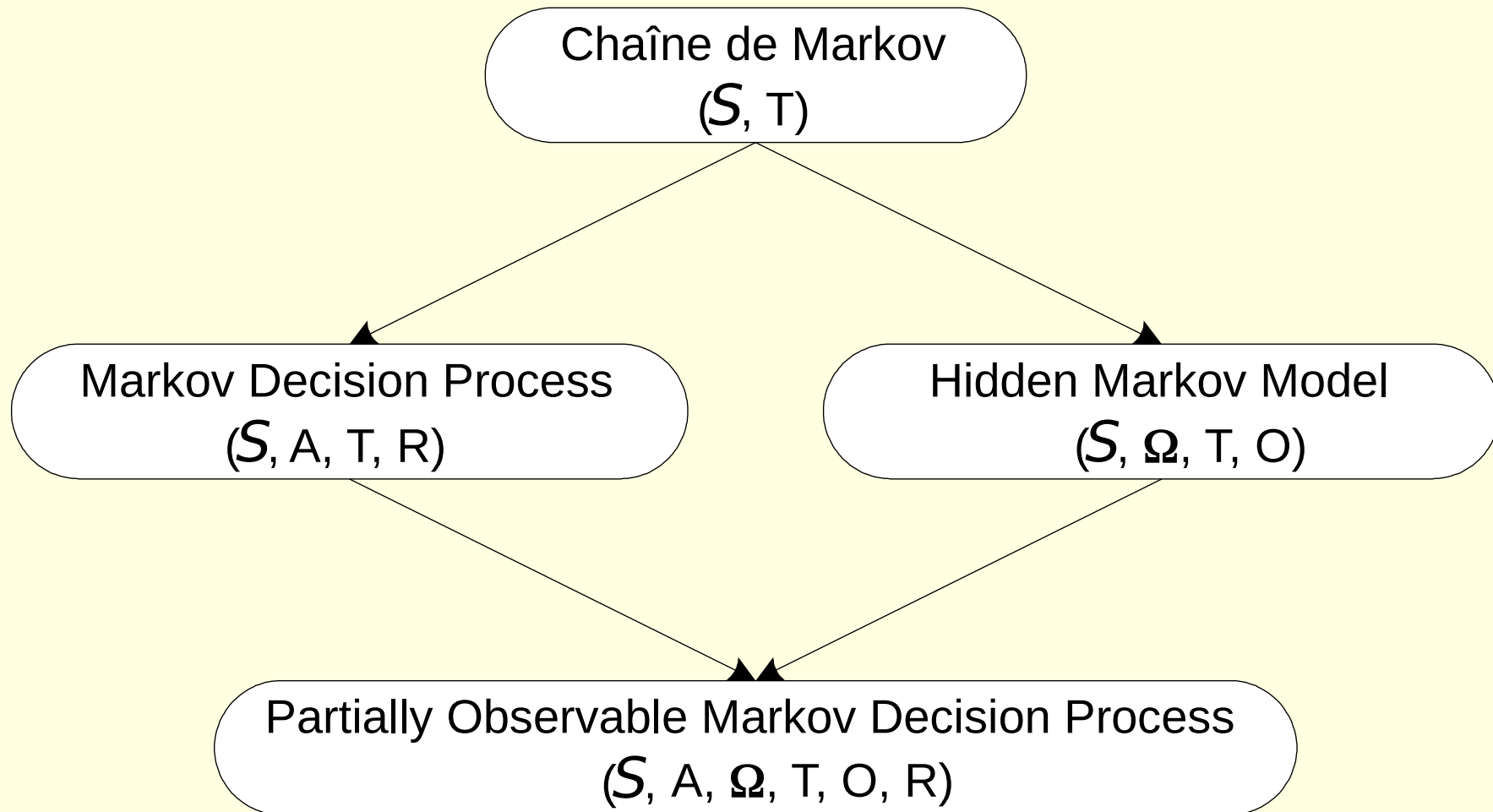
Algorithmes topologiques

Laurent Jeanpierre
(laurent.jeanpierre@unicaen.fr)

Plan du cours

- Rappels sur les modèles de Markov
- Un exemple
- Les algorithmes topologiques
- Les approches hiérarchiques

Rappels sur les modèles de Markov



Rappels sur les modèles de Markov

Un POMDP est un n -uplet $\{S, A, \Omega, O, T, R\}$

- S : Ensemble fini d'états
- A : Ensemble fini d'actions
- Ω : Ensemble fini d'observations
- O : Fonction d'observation
 $O : \Omega \times S \times A \times S \rightarrow [0; 1] = P(o|s, a, s')$
- T : Loi de transition probabiliste
 $T : S \times A \times S \rightarrow [0; 1] = P(s'|s, a)$
- R : Fonction de récompense
 $R : S \times A \times S \rightarrow R = R(s, a, s')$

Plan du cours

- Rappels sur les modèles de Markov
- Un exemple
- Les algorithmes topologiques
- Les approches hiérarchiques

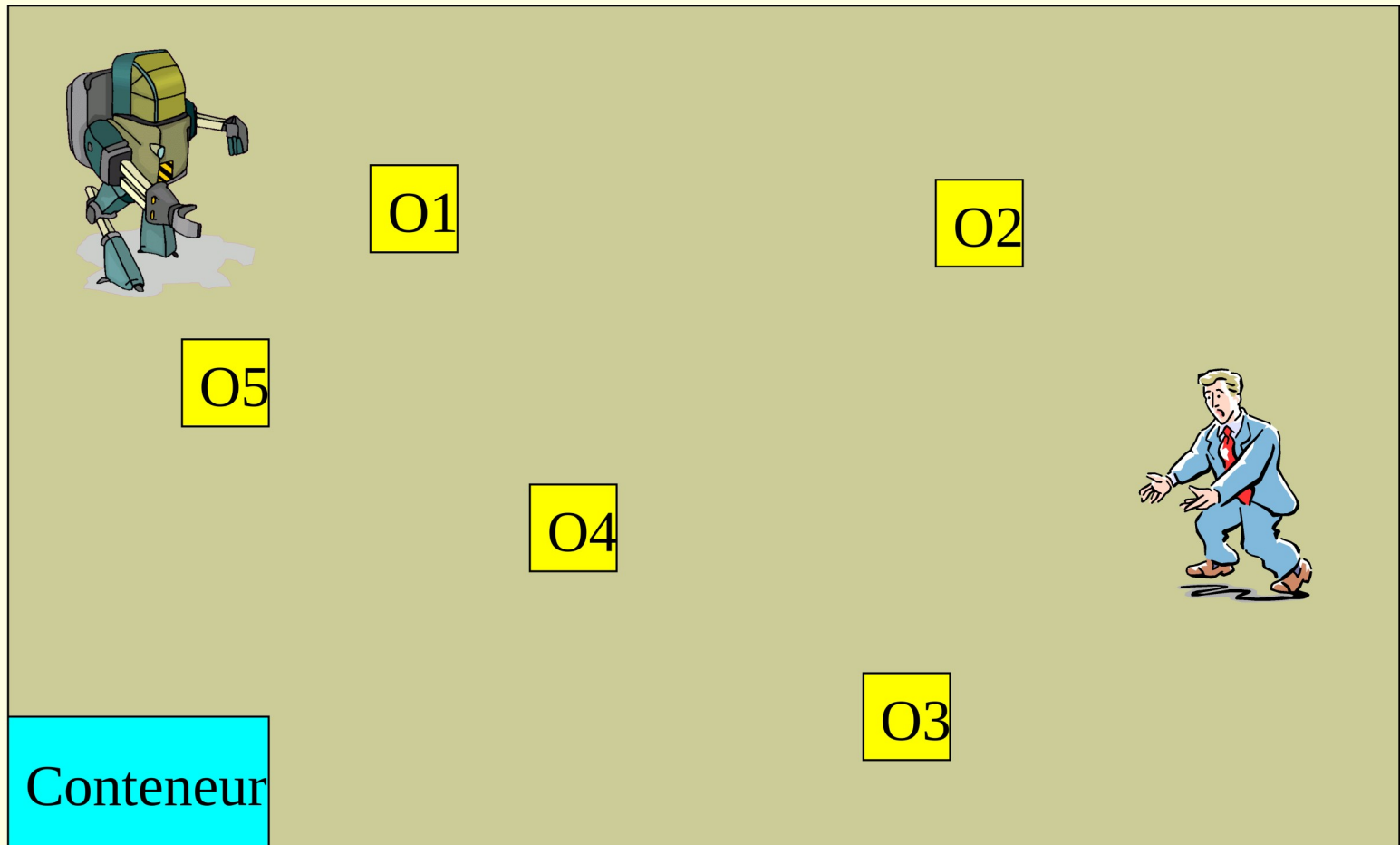
Exemple applicatif

L'interaction Homme-Robot

- Système multi-agent hétérogène
 - 1 ou plusieurs robots
 - 1 ou plusieurs êtres humains
- Une mission commune
 - Plusieurs tâches
 - Peuvent être réalisées
 - Par l'homme
 - Par le robot
 - Par l'un ou l'autre
 - Par les deux
 - Avec ou sans communication explicite
- Projet Amorçe

Exemple d'application

Stockage d'objets dans un conteneur



Exemple d'application (suite)

Les contraintes

- Le robot doit :
 - Apporter son aide
 - Sans gêner l'homme
- Une tâche :
 - Ramasser un objet au sol
 - Se déplacer jusqu'au conteneur
- Objets homogènes
 - Pas de préférence *a priori*
 - Pas de contrainte d'impossibilité/interdit
- Pas de communication directe
- Observation totale de l'environnement

Les limites du discret

Généralités

- Notre monde est continu (quoique...)
 - Distances, vitesses
- Nos ordinateurs sont discrets
 - Notion de bit, de cases mémoires
- Les flottants : un leurre crédible
 - Échange la précision contre l'amplitude
 - Petits nombres → très précis
 - Grands nombres → très approximatifs
 - ***Pas d'itération 'correcte' possible !***
- ⇒ Il faut discrétiser

Etats continus

- Un problème très étudié
 - Depuis 2000 environ...
- Plusieurs approches différentes
 - Discrétisation environnement
 - Préalable
 - Adaptative (Fonction de valeur)
 - Adaptative (Politique)
 - Discrétisation de la politique
 - Approximateurs mathématiques

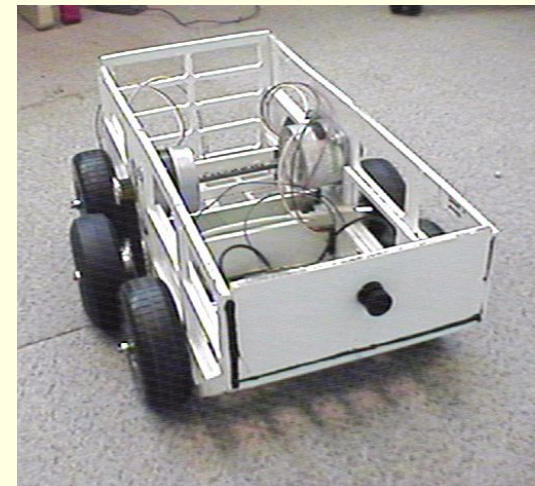
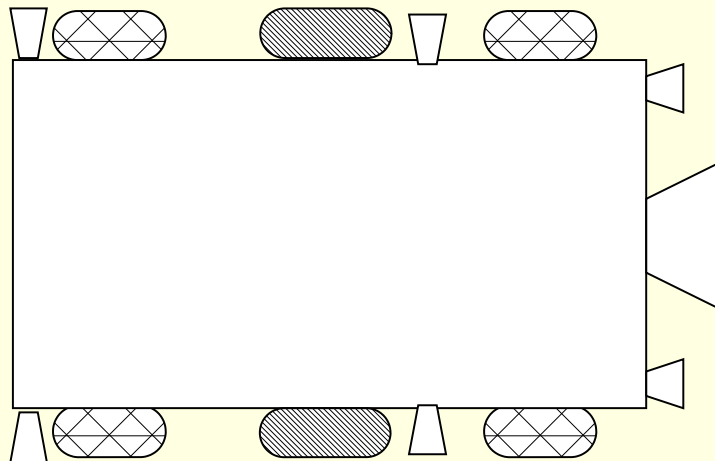
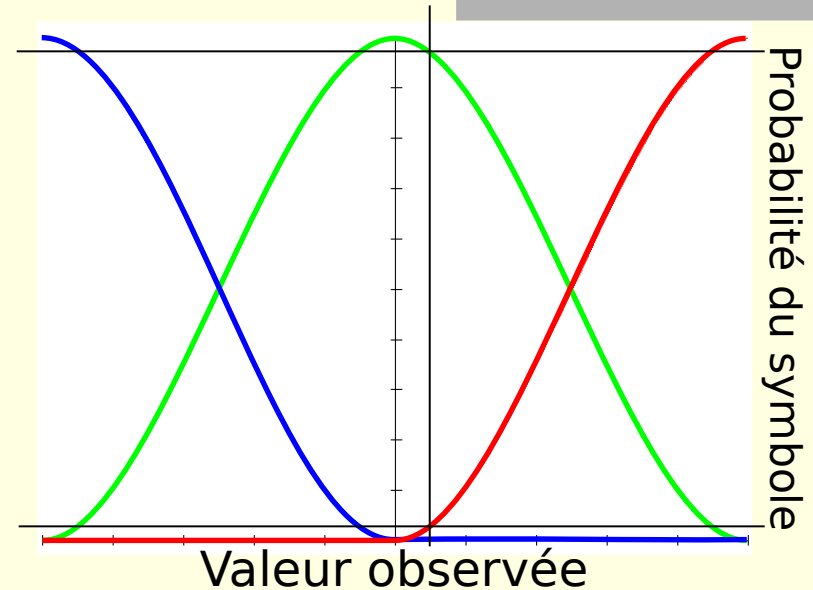
Discrétisations possibles

- Par zone géographique
 - Maillage +/- régulier
 - Ex : précision de 10cm \Rightarrow échiquier
 - Valeur des capteurs ?
 - \Rightarrow Distribution de probabilités
 - États gros \Rightarrow Probabilités uniformes
 - Plusieurs états \Rightarrow Même observation
- Par observations
 - Un état \rightarrow Une observation
 - \Rightarrow Nombreux états
 - \Rightarrow Souvent répartis
 - \Rightarrow Pas d'information géographique

Discrétisation préalable

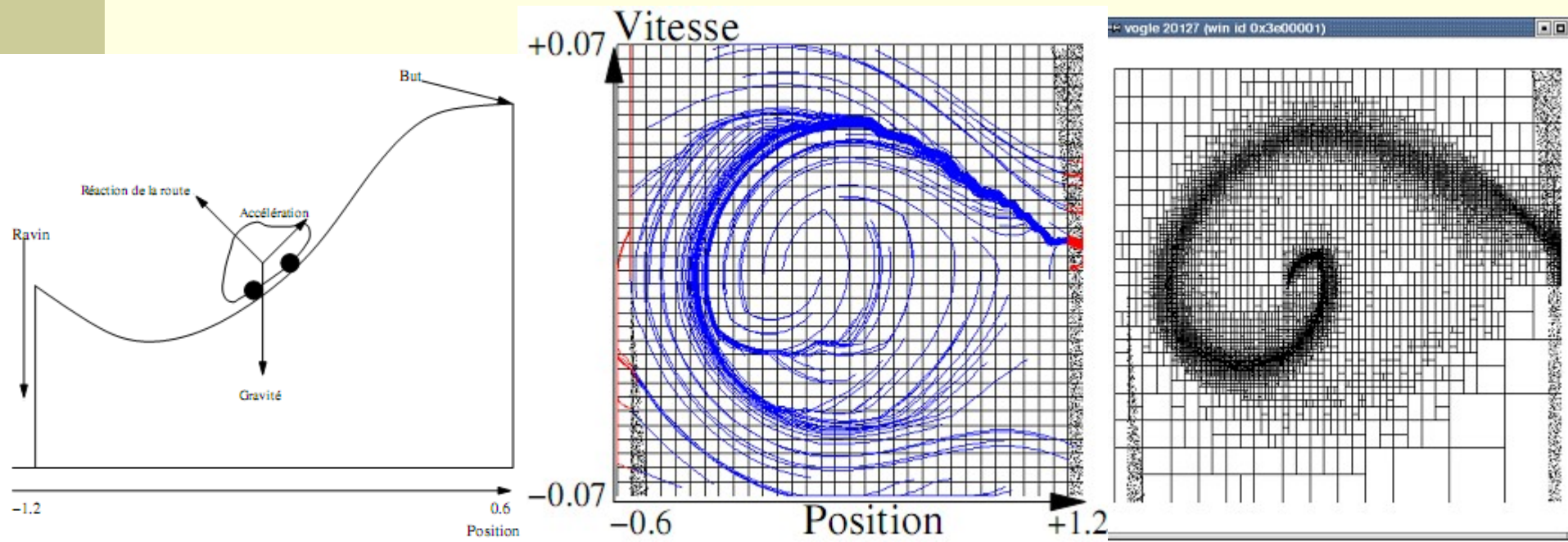
Démonstration 1/2

- Perception floue
 - Proche ($<20\text{cm}$)
 - Distant ($\sim 50\text{cm}$)
 - Loin ($>80\text{cm}$)
- 6 Capteurs



Discrétisation adaptative

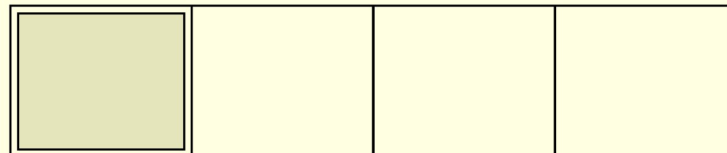
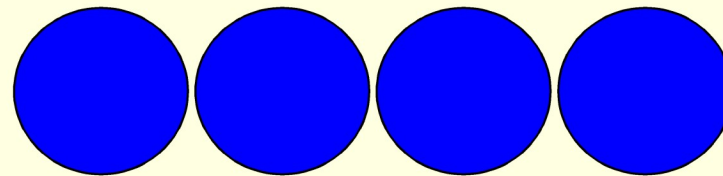
- Comment mettre des états là où il en faut ?
 - Grandes variations de la fonction de valeur
 - Variations de la politique
 - Modification pendant le calcul de politique
- Ex : 'Car on the hill' (© Bruno Scherrer 2003)



Les limites du discret

Augmentation de l'indéterminisme

- Action déterministe : Avancer de 33cm.
- Environnement maillé par pas de 1m
- État initial = état 0, 100%
 - T=1 : 67%, 33%, 0%, 0%
 - T=2 : 44%, 44%, 12%, 0%
 - T=3 : 30%, 44%, 22%, 4%



Non-discrétisation possible ?

Il est possible de NE PAS discrétiser les états :

- Discrétisation du belief-state : MC-POMDP
 - Travaux de S. Thrun *et al* depuis 2000
- Discrétisation de la fonction de valeur : PB-VI
 - Travaux récents, multiples chercheurs
 - Pineau, Gordon, Thrun, Spaan, Vlassis, Porta, Poupart

Modèle discret

L'environnement

R			O2				
		O1					
						O4	
			O3	H			
B							

Espace d'états :

- Position Robot
 - 8x8
- Position Homme
 - x8x8
- Position O1..O4
 - x{au_sol, humain, robot, rangé}
- ...

Total :

$$64^2 \times 4^4 = 262\ 144 \text{ états}$$

Modèle discret

Observabilité partielle

- Ajout des intentions de l'homme
 - Ramasser un objet (lequel ?)
 - Retourner au conteneur
 - Ne rien faire d'utile
 - → Non-observables directement
- Le MDP devient un POMDP
- Observation des actions de l'homme
 - Comparaison avec une politique rationnelle
 - 1 MDP mono-agent par tâche
 - → Politique rationnelle
 - → Fonction de valeur & Q-Values
 - Transformation des Q-Values en probabilité d'observation des actions

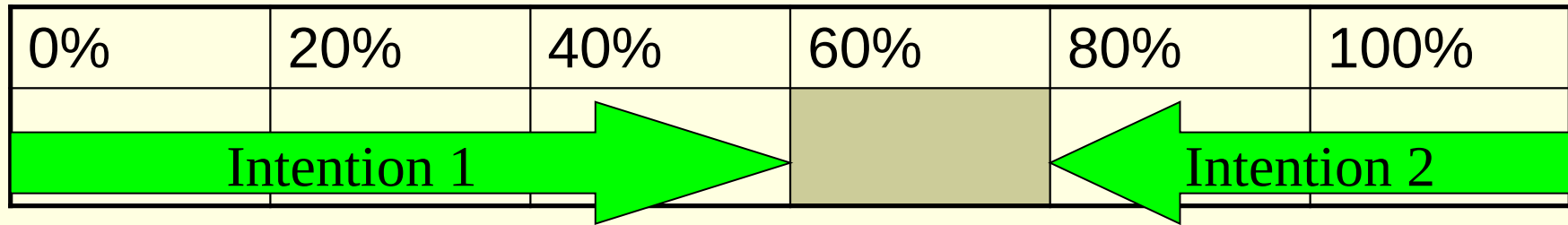
Modèle discret

Observabilité partielle (suite)

- Passage au Belief-MDP
 - Discrétisation des intentions
 - Exemple : 5 tranches de 20%
- Conséquences :
 - Retour à un MDP
 - Espace d'états augmente
 - Exemple précédent : $262\,144 \times C_5^{10} > 132 \cdot 10^6$
 - \Rightarrow éventuellement solvable
 - Élagage des états
 - États cohérents : 4 179 540 seulement (3% dans ce cas)

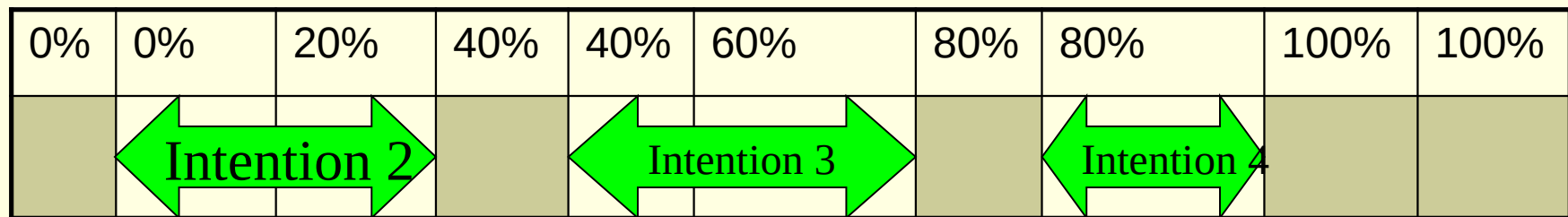
Apparté sur les distributions de probas

Pourquoi C_5^{10} ?



2 Intentions, 5 probas \Rightarrow 1 choix parmi 6

Ici : 60%, 40%



6 Intentions, 5 probas \Rightarrow 5 choix parmi 10

Ici : 0%, 40%, 40%, 20%, 0%, 0%

Exercice : Construire toutes les distributions possibles...

Plan du cours

- Rappels sur les modèles de Markov
- Un exemple
- Les algorithmes topologiques
- Les approches hiérarchiques

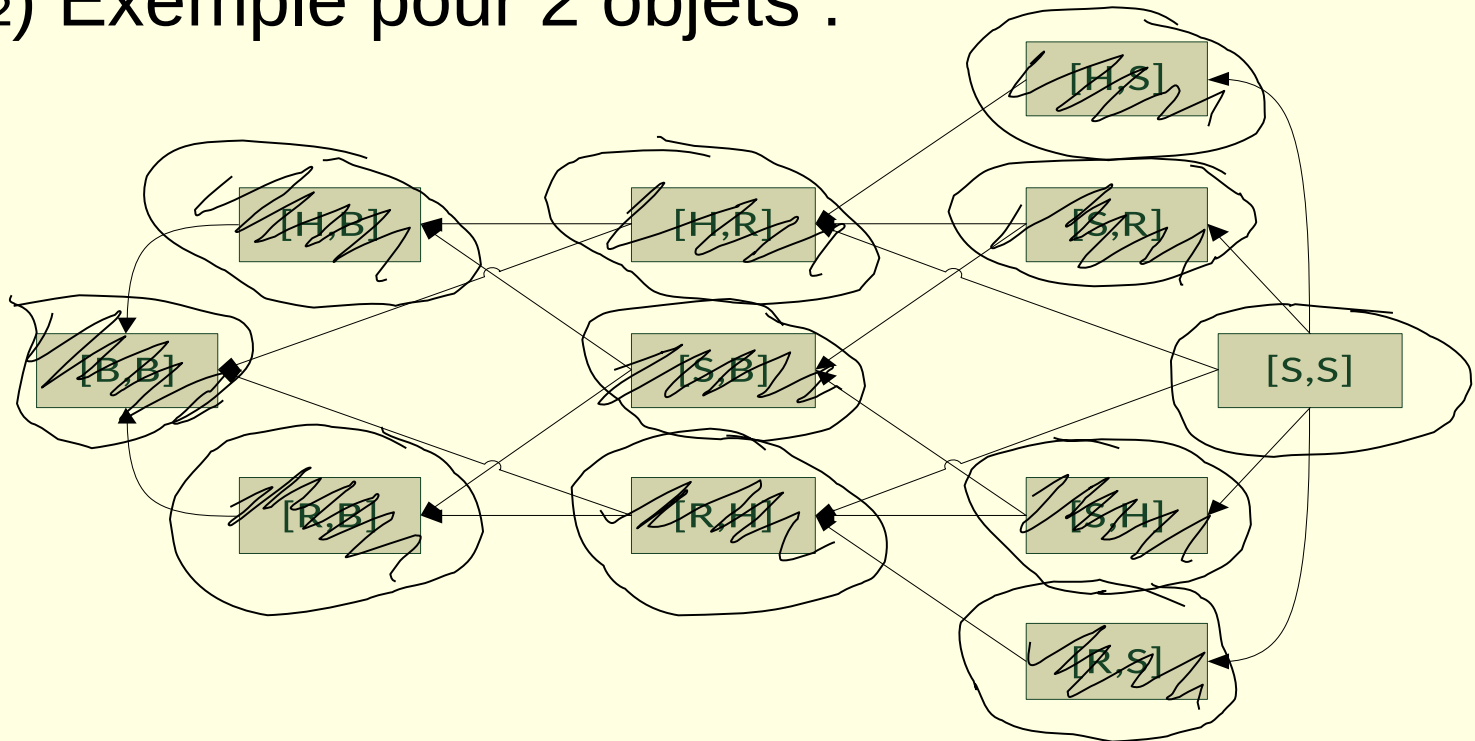
Motivation

- *Value iteration* propage les valeurs
 - À partir des récompenses
 - Vers les états qui permettent de les atteindre
 - Et ainsi de suite
- Tirer parti de la structure du problème
 - Séquences (étapes) obligatoires
 - En ordre connu ou variable
- ⇒ Résoudre des modèles de décision
 - Avec espace d'état important
 - De façon efficace

Modèle topologique

Application à l'exemple

- Récompense quand tous les objets sont rangés
- Chaque agent ne peut porter qu'un seul objet
- Un objet porté ne peut être mis que dans la boîte
- (1/2) Exemple pour 2 objets :

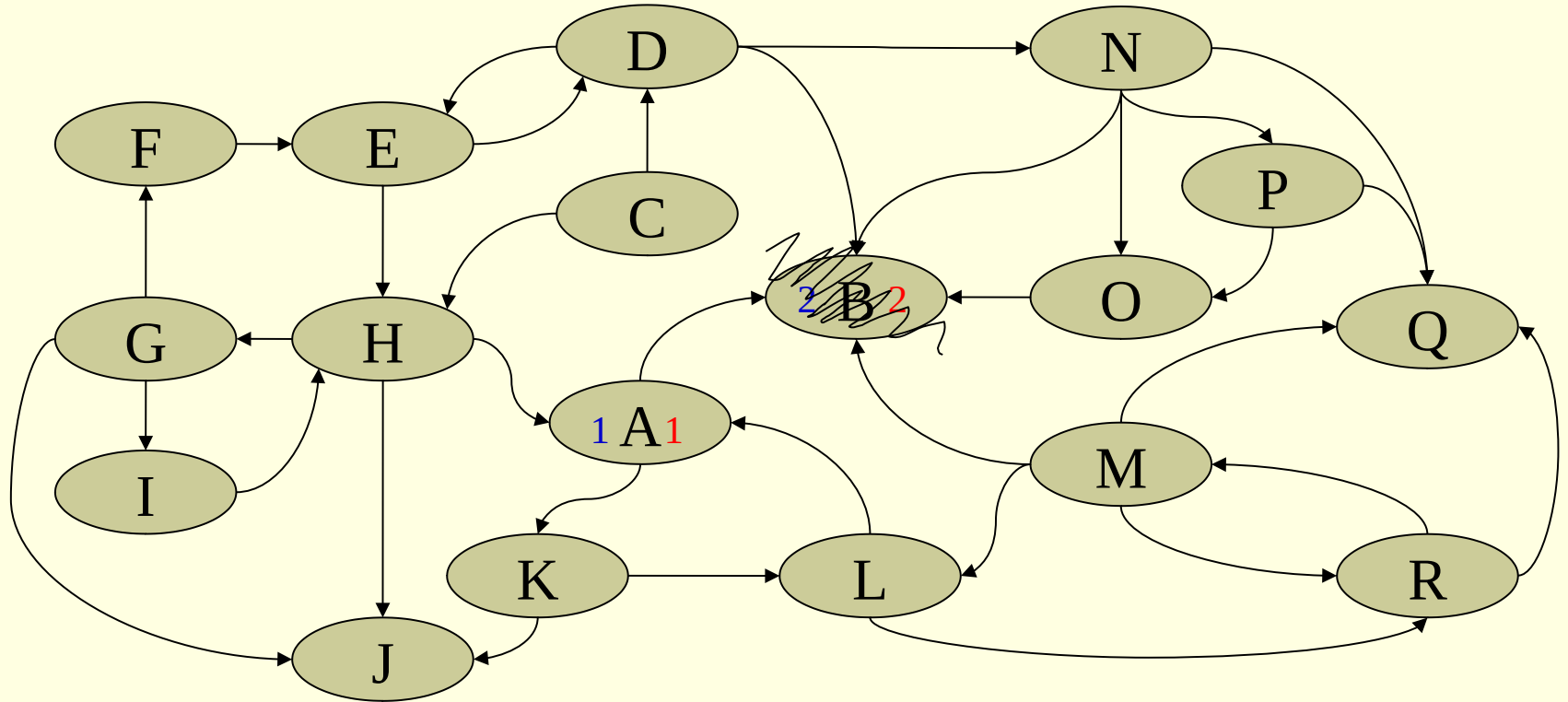


Algorithmique Topologique

- Recherche de la topologie
 - Composantes fortement connexes
 - Groupe d'états accessibles à partir de tout le groupe
 - Algorithmes
 - Kosaraju (Rao S. Kosaraju. Unpublished. 1978)
 - Tarjan (Robert E. Tarjan. *Depth first search and linear graph algorithms*. **SIAM Journal on Computing**, Volume 1, Number 2, pp. 146-160, **1972**.)
- Résolution
 - TVI (MDP, Dai P., Goldsmith J. *Topological Value Iteration Algorithm for Markov Decision Processes*. **IJCAI, 2007**)
 - TOP(POMDP, Dibangoye J.S., Shani G., Chaib-draa B., Mouaddib A-I., *Topological Order Planner for POMDPs*. **IJCAI, 2009**)

Algorithme de Tarjan

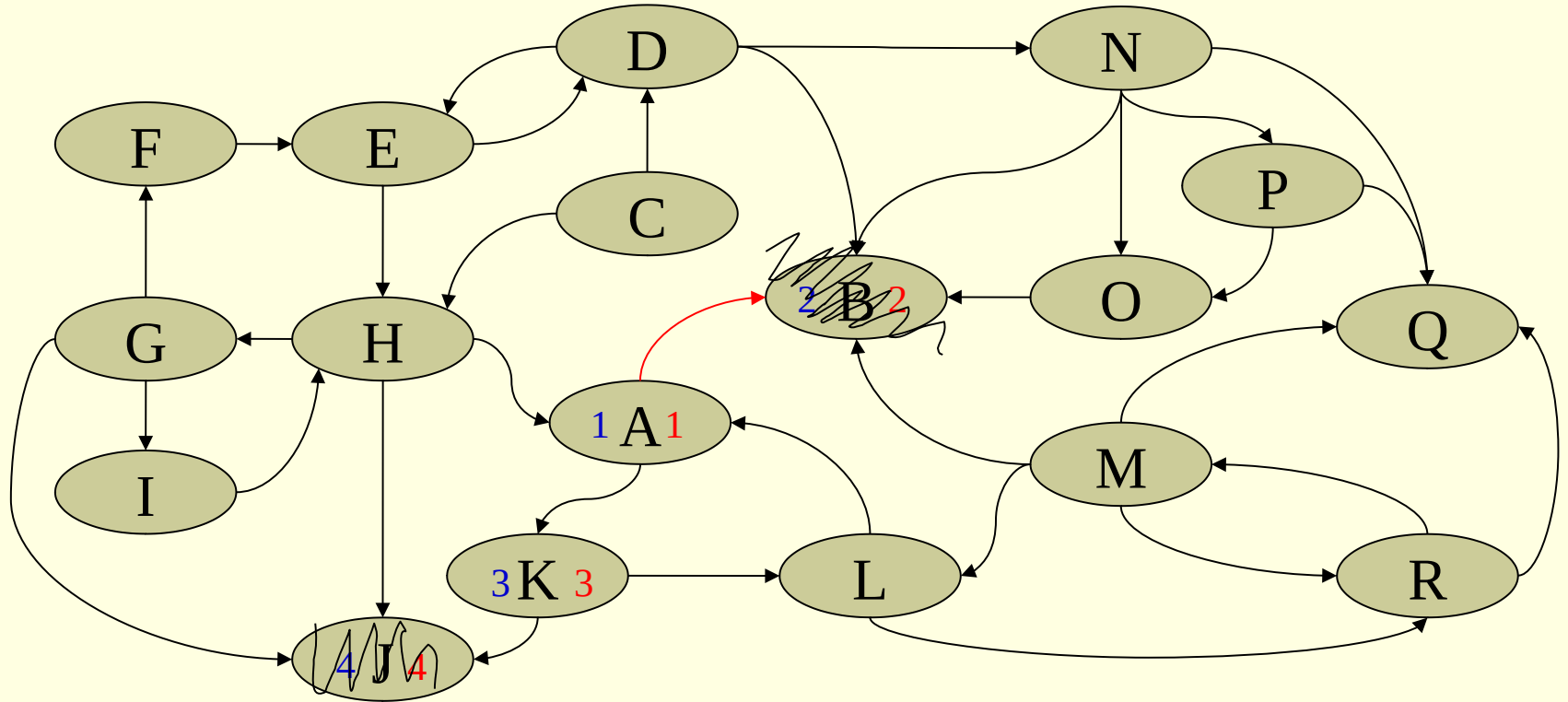
En images



B

Algorithme de Tarjan

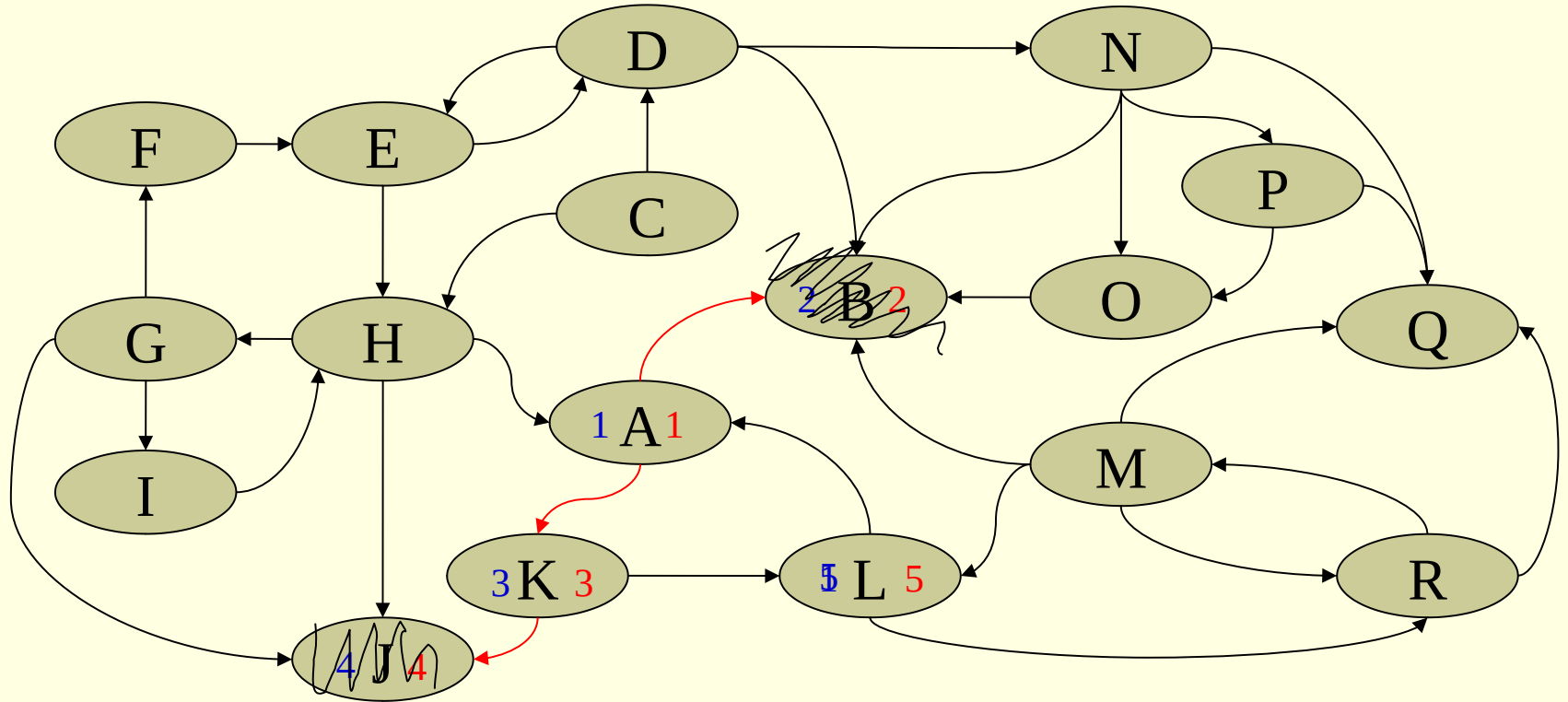
En images



B J

Algorithme de Tarjan

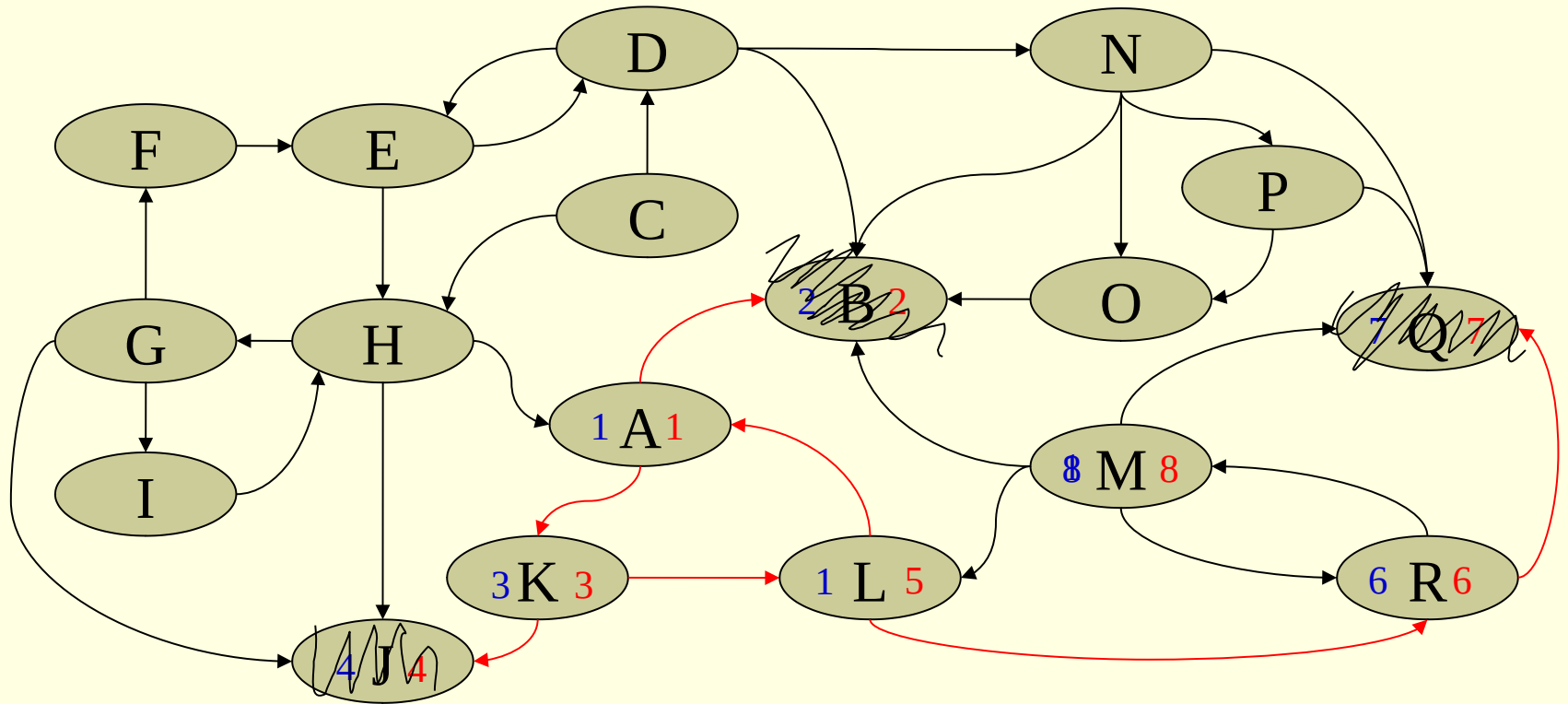
En images



B J

Algorithme de Tarjan

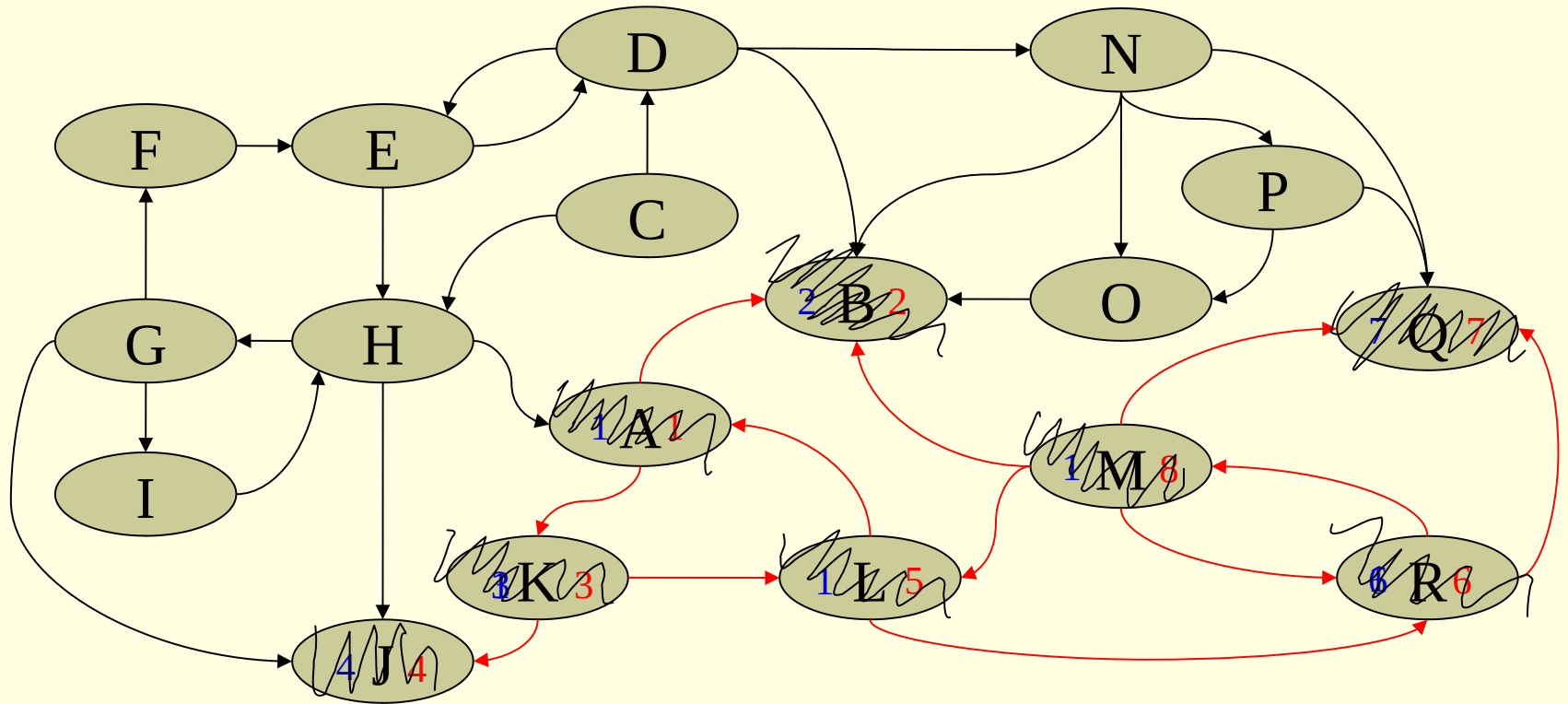
En images



B J Q

Algorithme de Tarjan

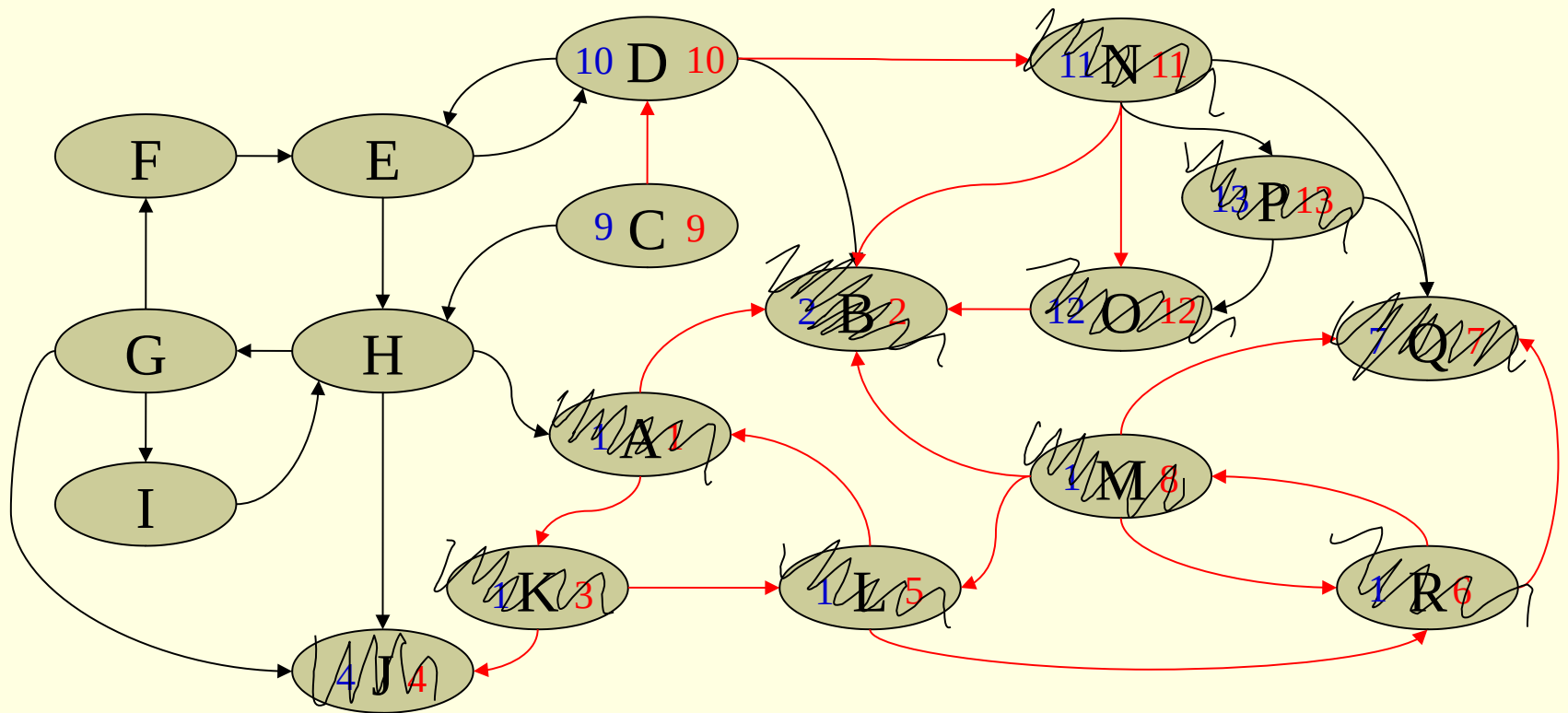
En images



B J Q M,R,L,K,A

Algorithme de Tarjan

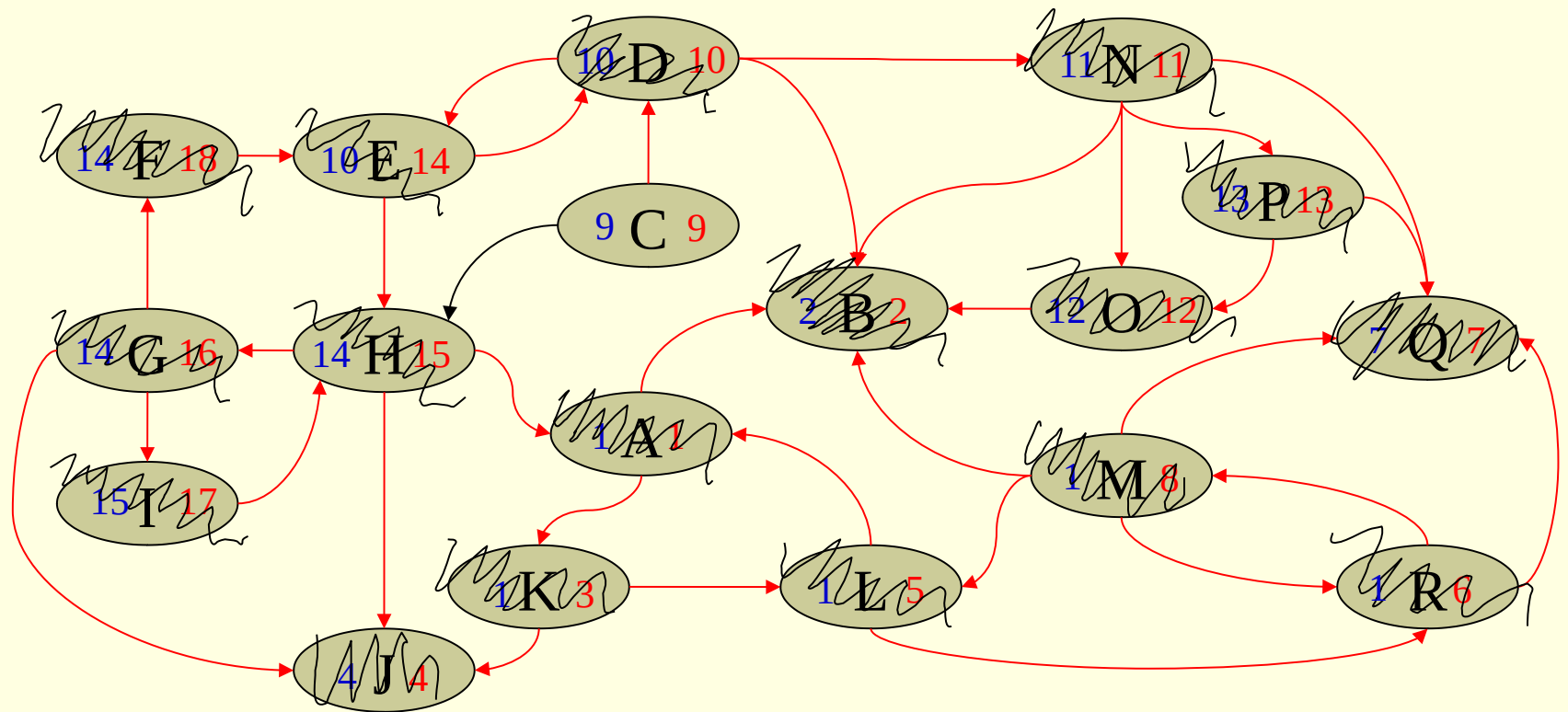
En images



- | | | | | | | |
|---|---|---|-----------|---|---|---|
| B | J | Q | M,R,L,K,A | O | P | N |
|---|---|---|-----------|---|---|---|

Algorithme de Tarjan

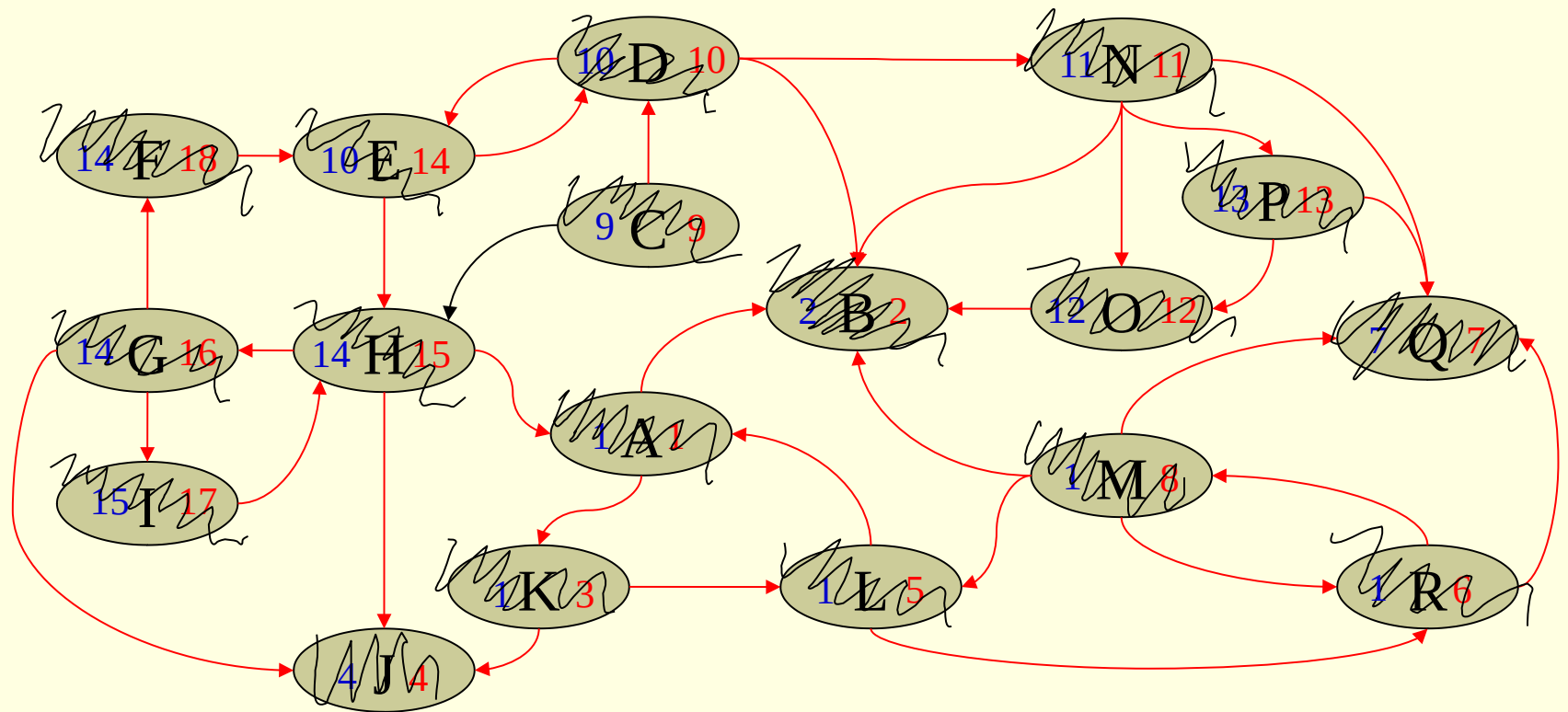
En images



[B] [J] [Q] [M,R,L,K,A] [O] [P] [N] [F,I,G,H,E,D]

Algorithme de Tarjan

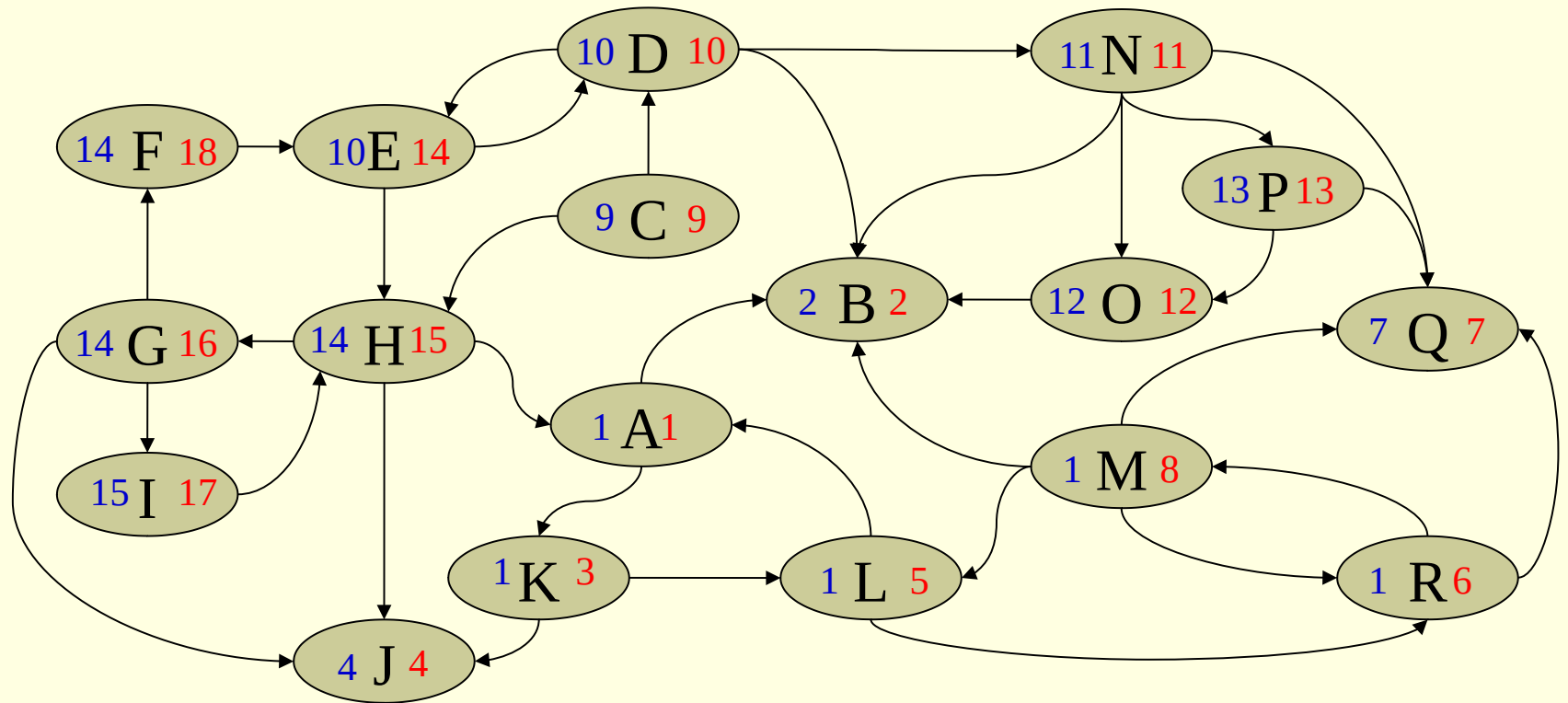
En images



[B] [J] [Q] [M,R,L,K,A] [O] [P] [N] [F,I,G,H,E,D] [C]

Algorithme de Tarjan

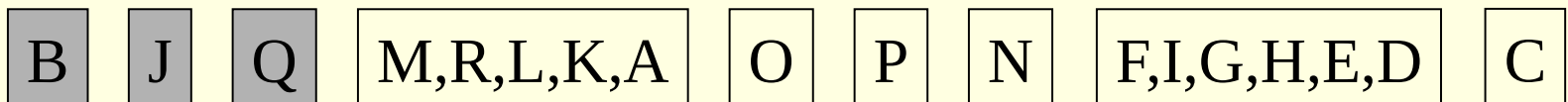
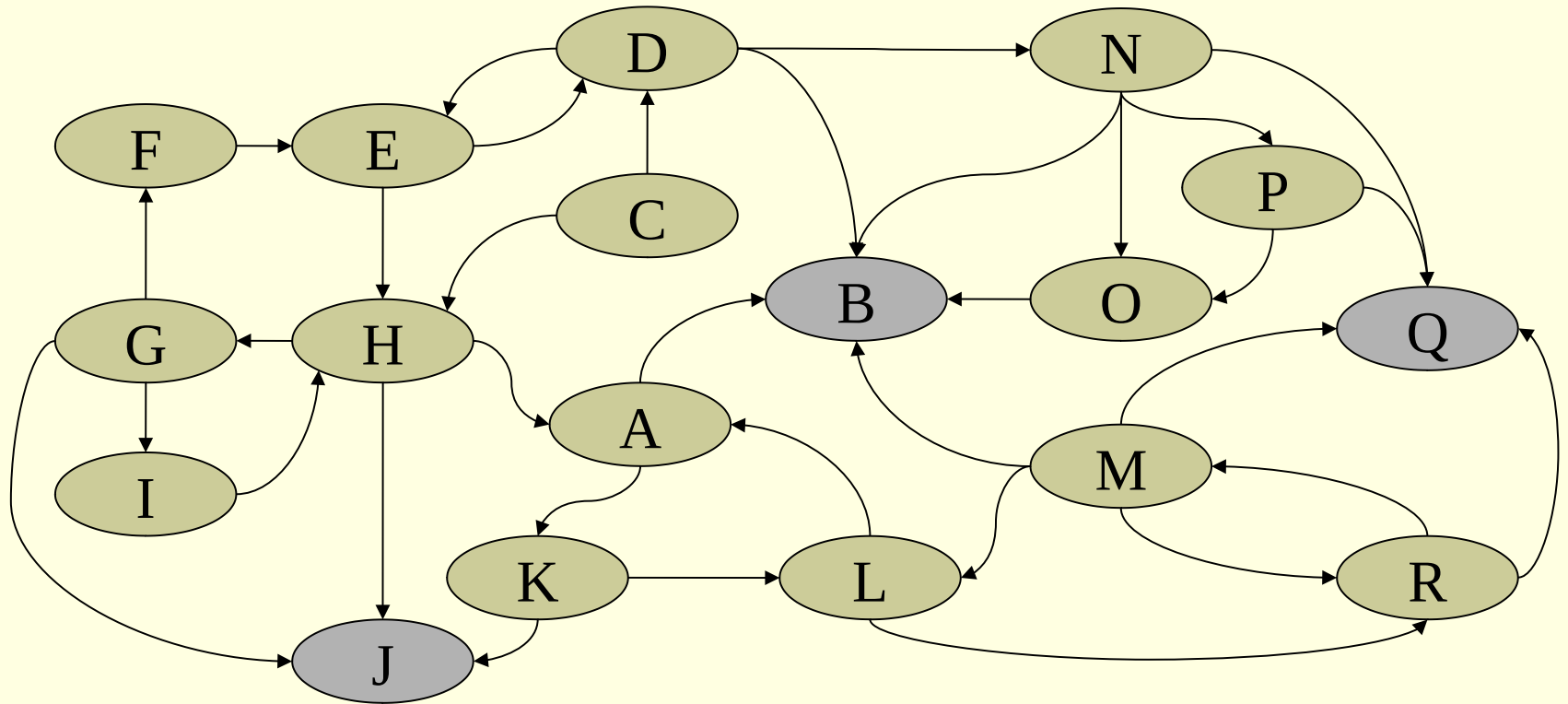
En images



B J Q M,R,L,K,A O P N F,I,G,H,E,D C

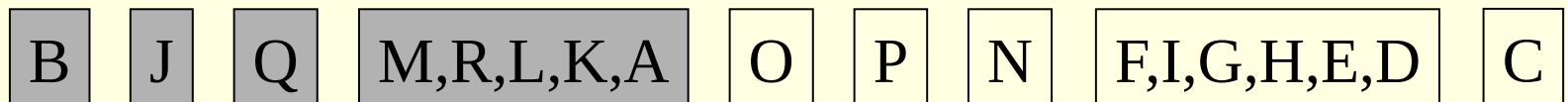
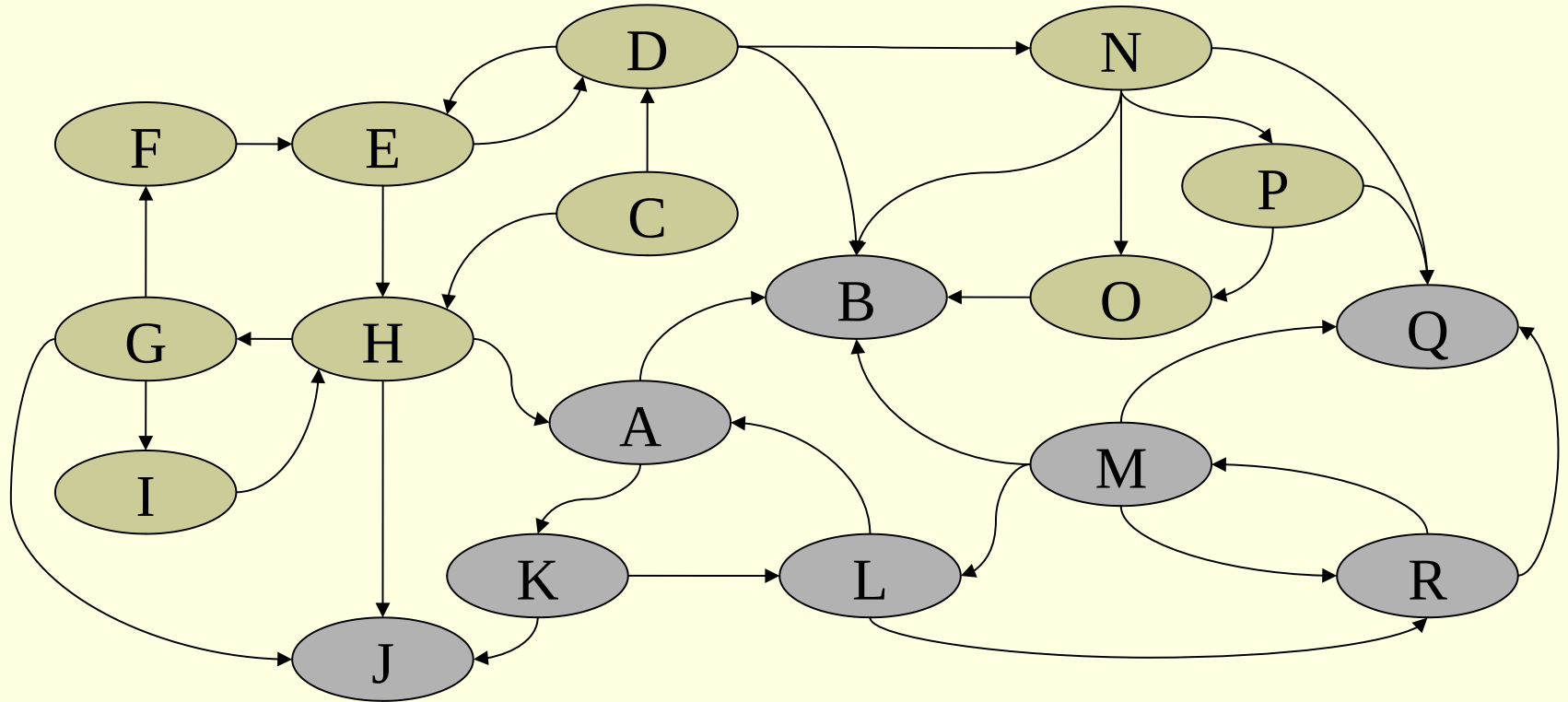
Topological Value Iteration

En images



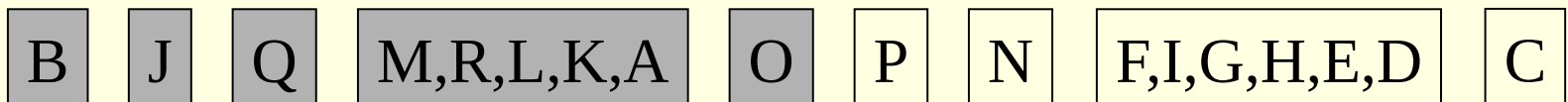
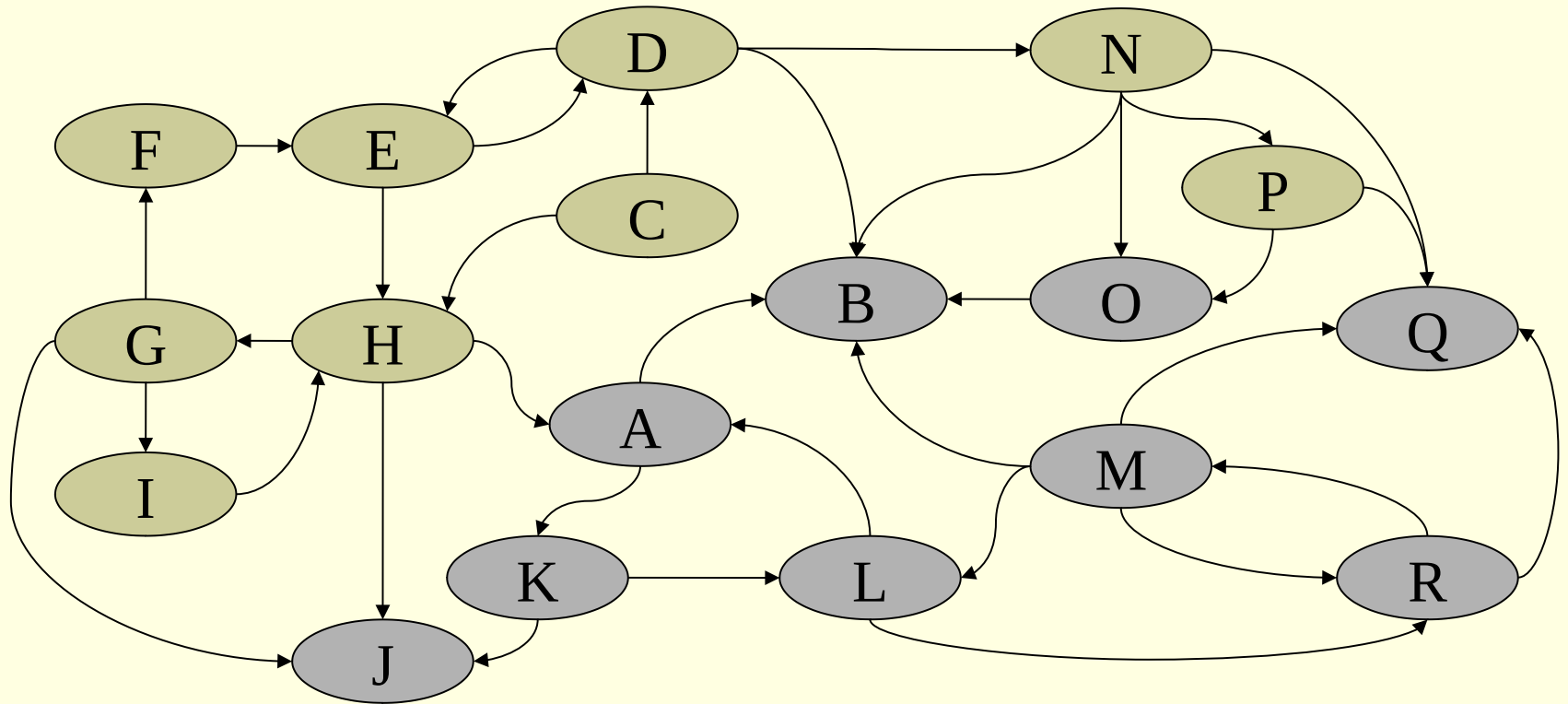
Topological Value Iteration

En images



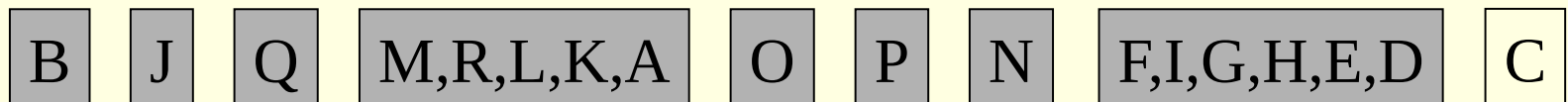
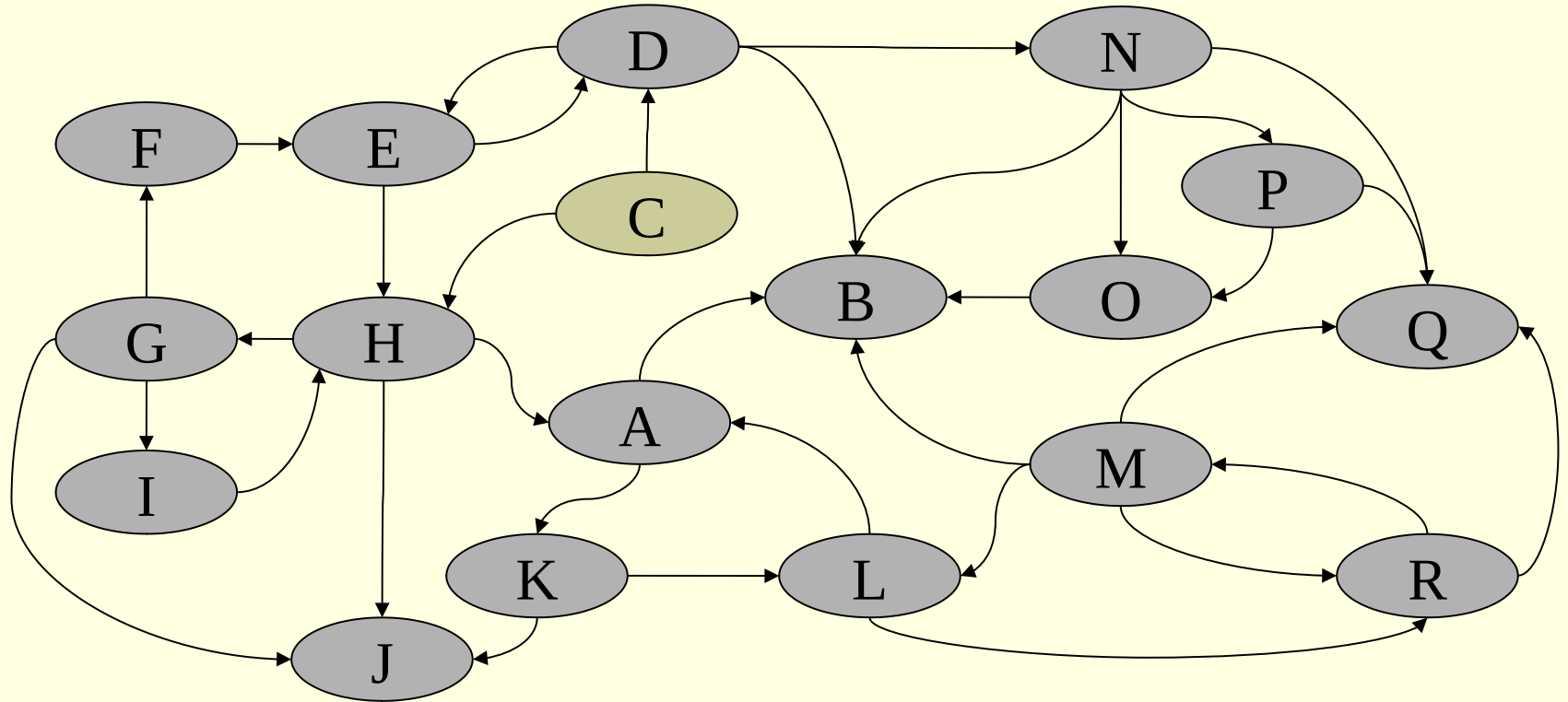
Topological Value Iteration

En images



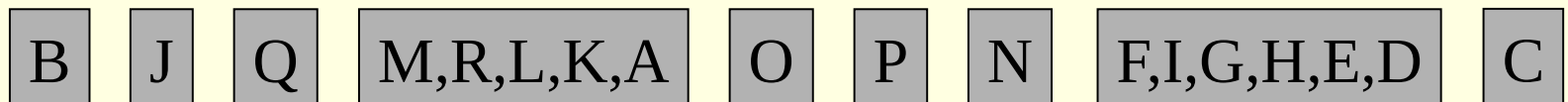
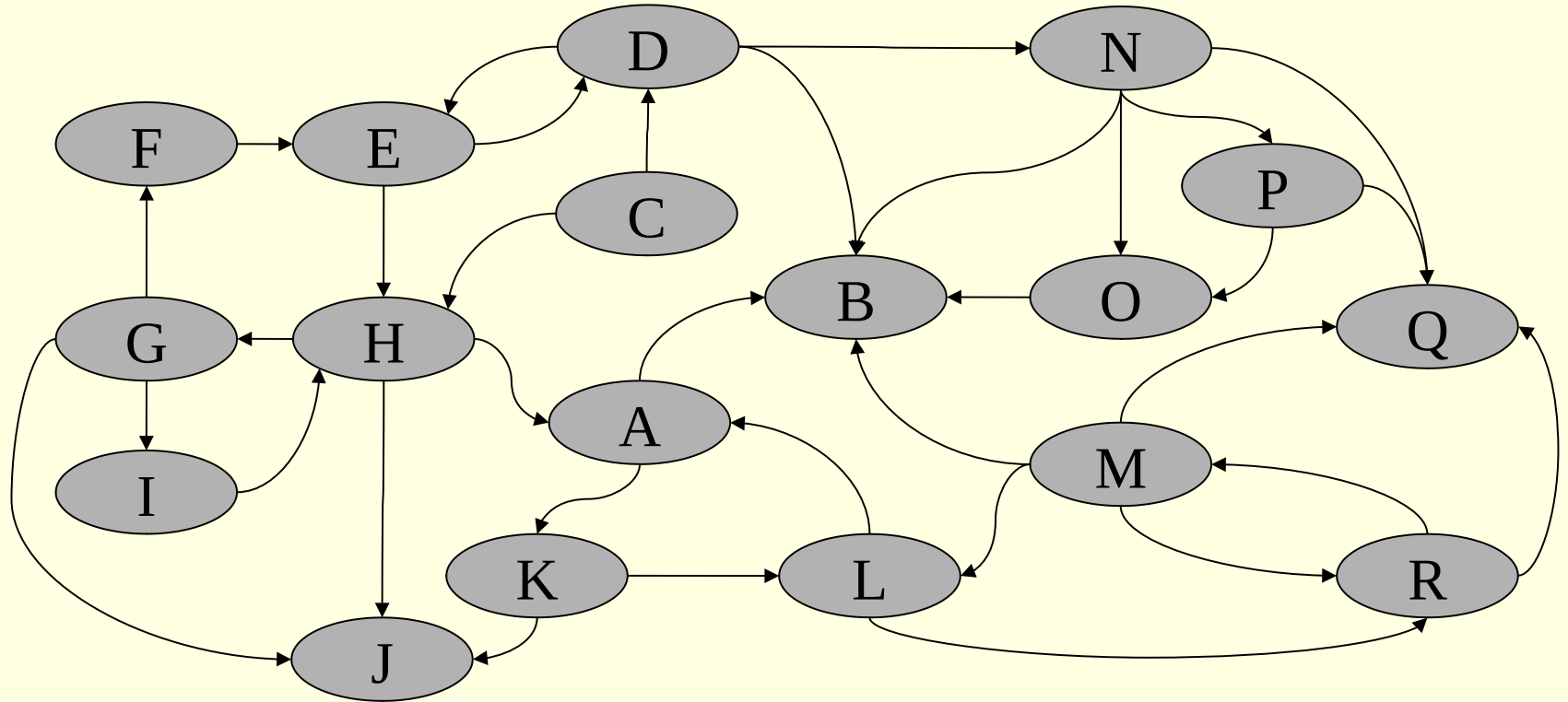
Topological Value Iteration

En images



Topological Value Iteration

En images



Topological Value Iteration

Bilan

- Sur l'exemple :
 - 18 états
 - 9 Value Iterations successifs
 - Au plus 6 états simultanés
- Rappel : Complexité de Value Iteration
 - $= O(s^2.a)$
 - Pire cas : s iterations (pour avoir la bonne politique)
 - $\Rightarrow O(s^3)$
- Complexité de l'exemple :
 - $18^3 = 5832$
 - $1 + 6^3 + 1 + 1 + 1 + 5^3 + 1 + 1 + 1 = 348$ (6% environ)

Topological Value Iteration

Fin

- Démonstration

Plan du cours

- Rappels sur les modèles de Markov
- Un exemple
- Les algorithmes topologiques
- Les approches hiérarchiques

Motivation

- TVI fonctionne très bien
 - Résolution exacte et rapide
 - S'il existe des composantes connexes
- Sinon ?
 - 1 seule composante
 - 1 seul Value Iteration
 - Aucun gain
- Que faire ?
 - Simuler des composantes connexes
 - Résolution souvent approximative

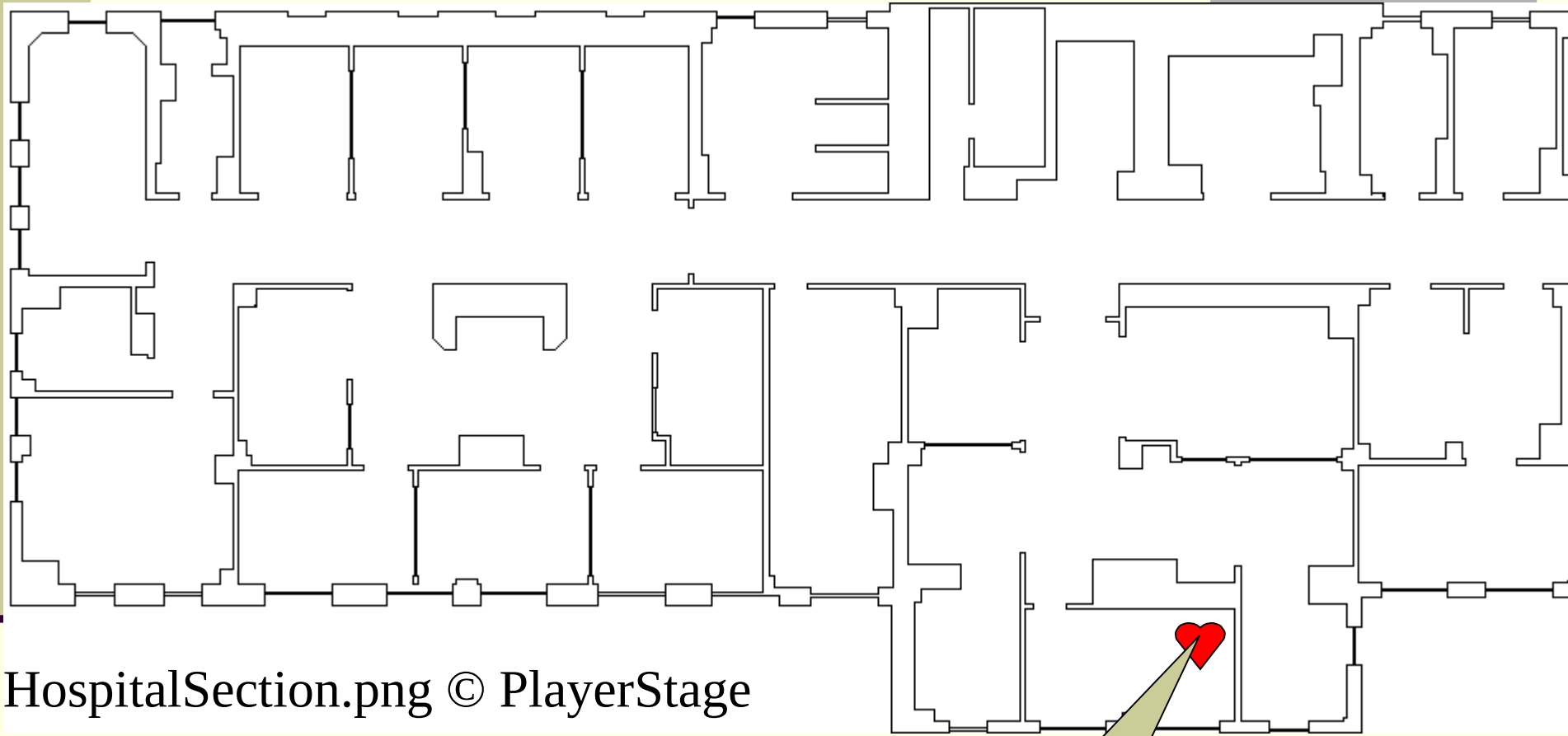
Algorithmes hiérarchiques

Principe général

- Créer un macro-modèle
 - Graphe de macro-états
 - Action = changer de macro-état
 - 1 macro-état = 1 MDP indépendant
- Résolution :
 - Créer le macro-modèle
 - Résoudre le macro-modèle
 - \Rightarrow Séquence de macro-états
 - Résoudre chaque macro-état
 - But local = Atteindre le macro-état suivant

Algorithmes hiérarchiques

Illustration



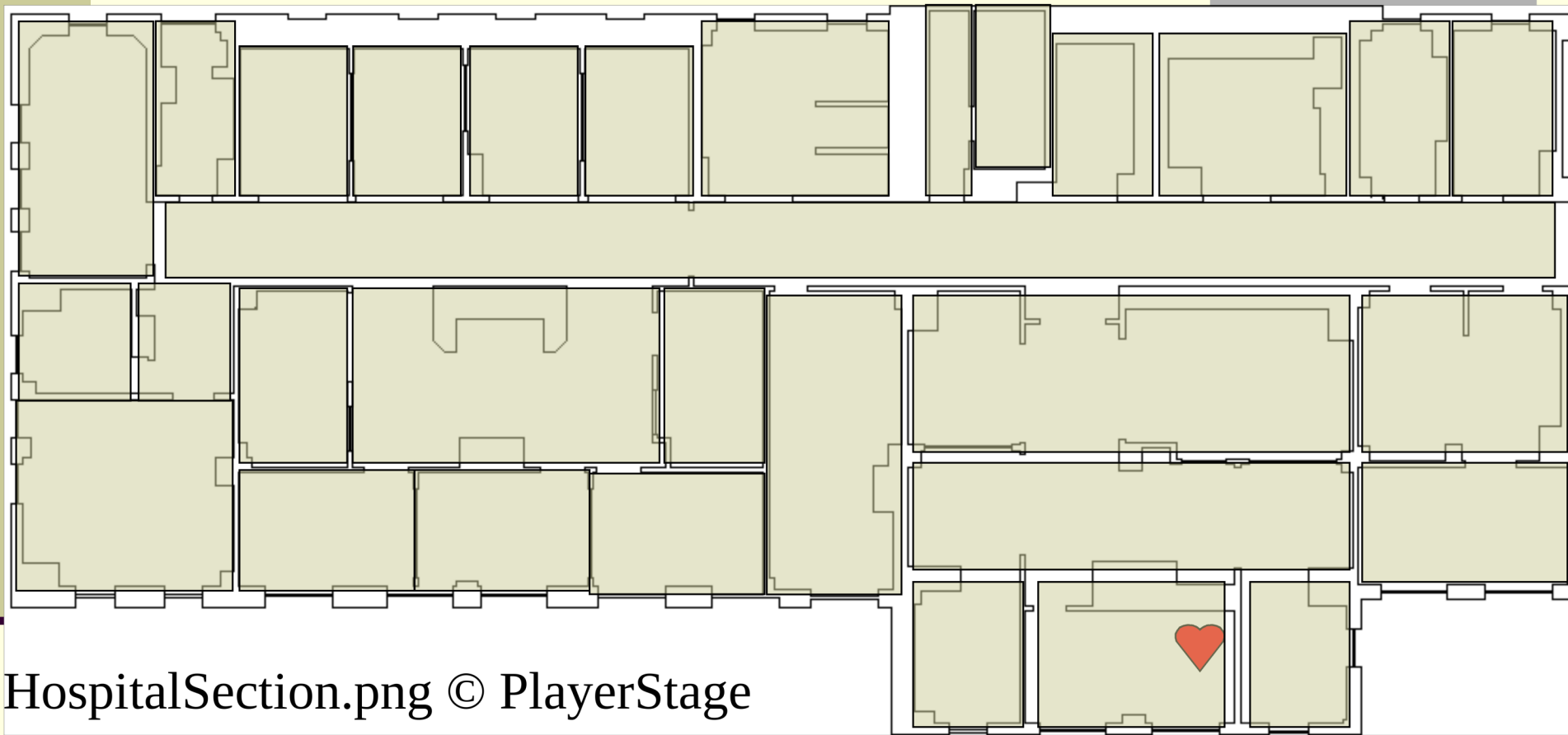
HospitalSection.png © PlayerStage

Modèle brut : plusieurs milliards d'états

Objectif

Algorithmes hiérarchiques

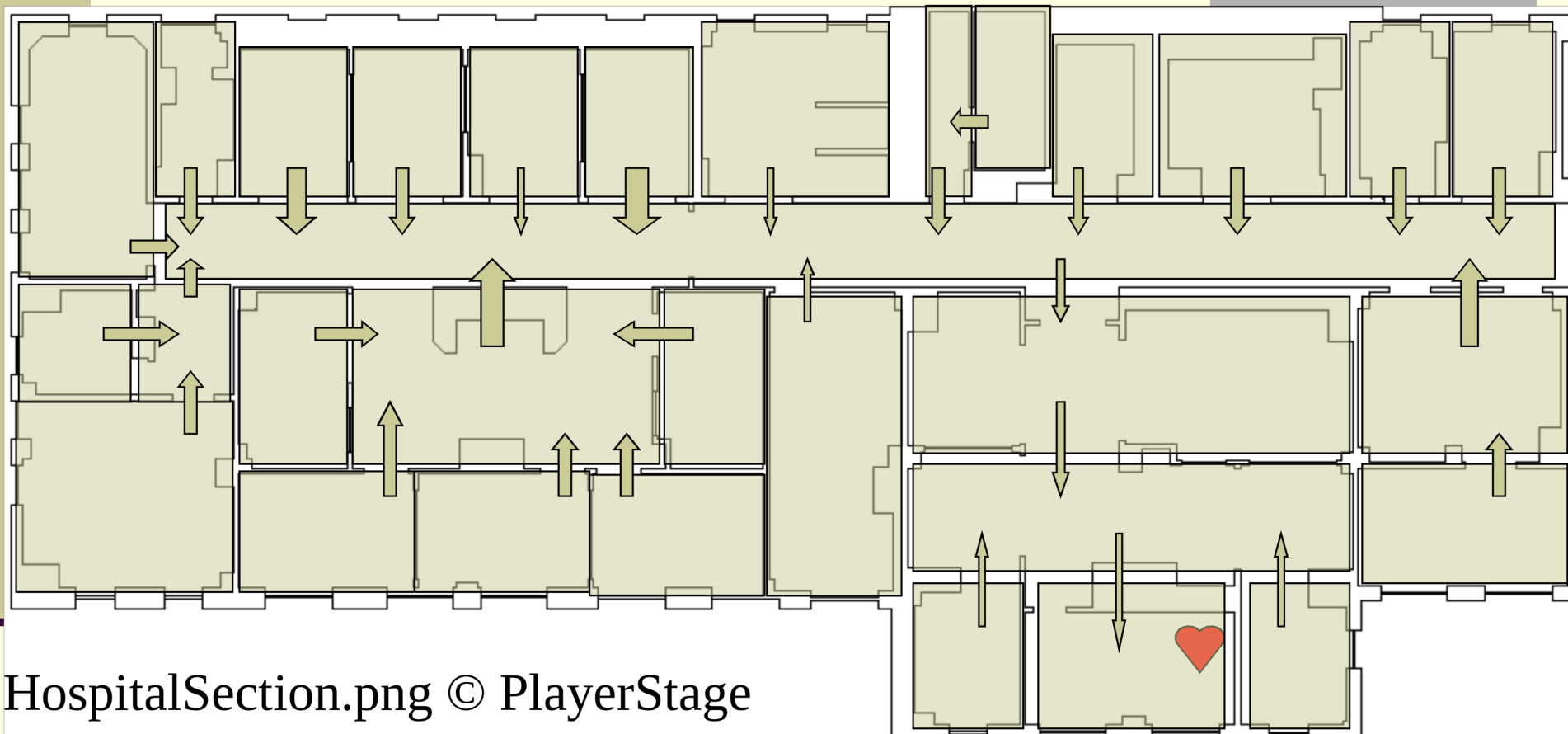
Illustration – Création du macro-modèle



⇒ 34 sections ou « macro-états »

Algorithmes hiérarchiques

Illustration – Résolution du macro-modèle

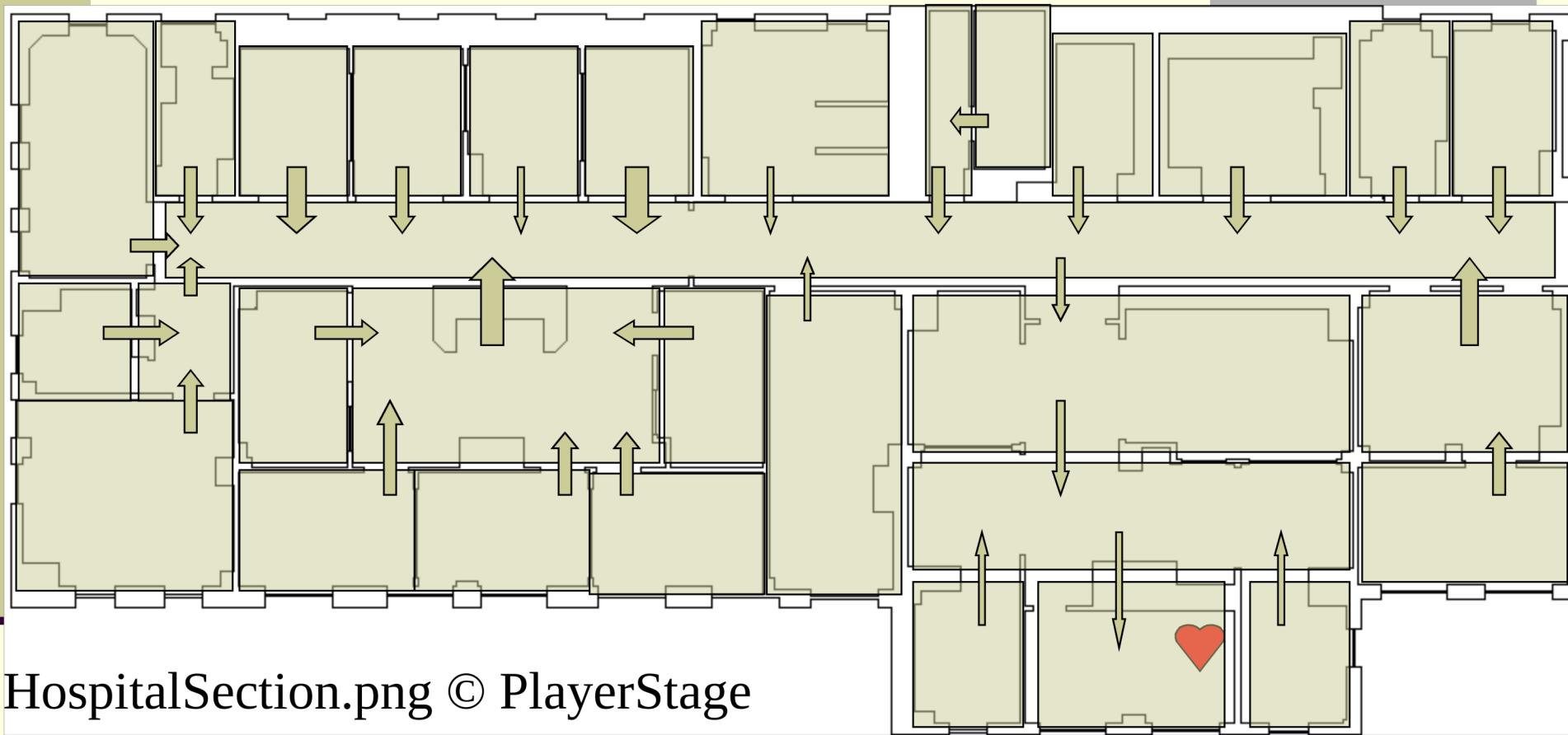


HospitalSection.png © PlayerStage

⇒ Atteindre le macro-état but

Algorithmes hiérarchiques

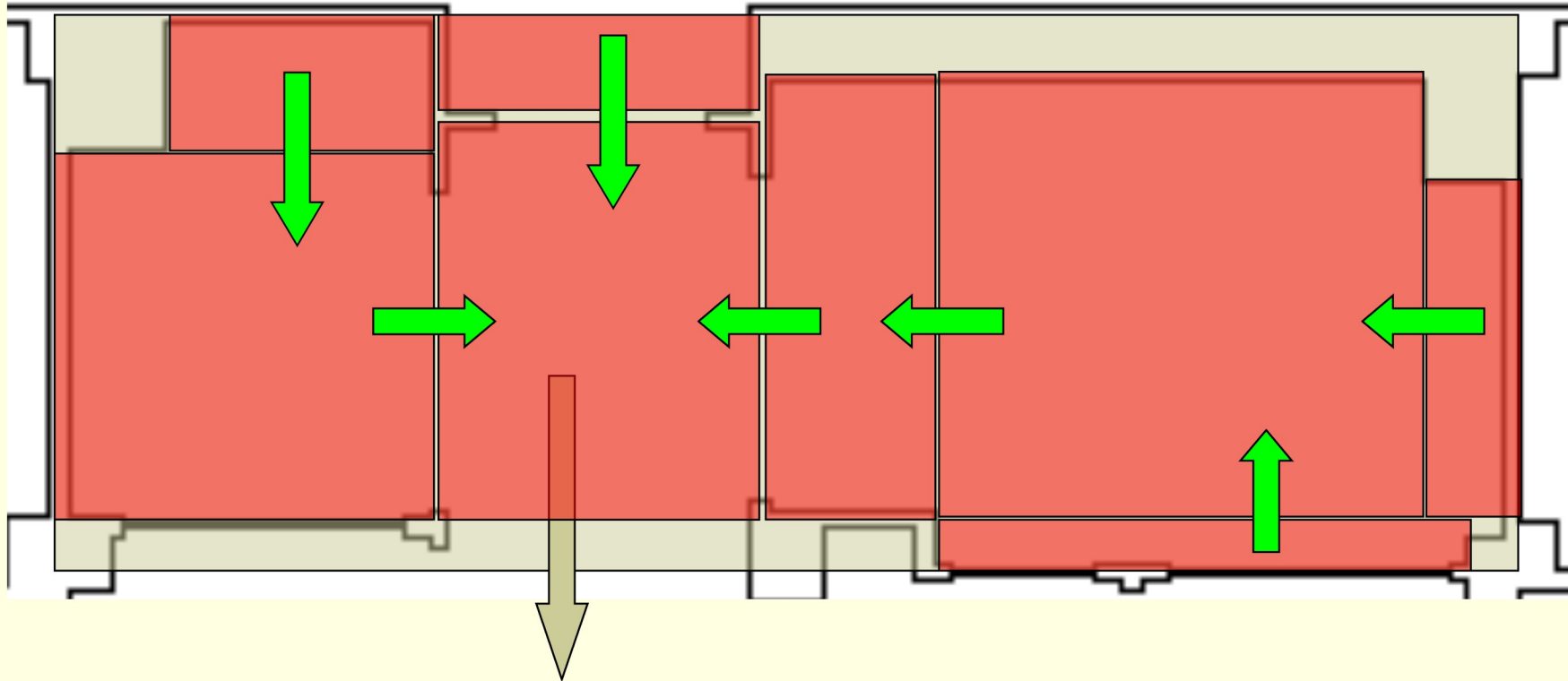
Illustration – Résolution des sous-modèles



HospitalSection.png © PlayerStage

Algorithmes hiérarchiques

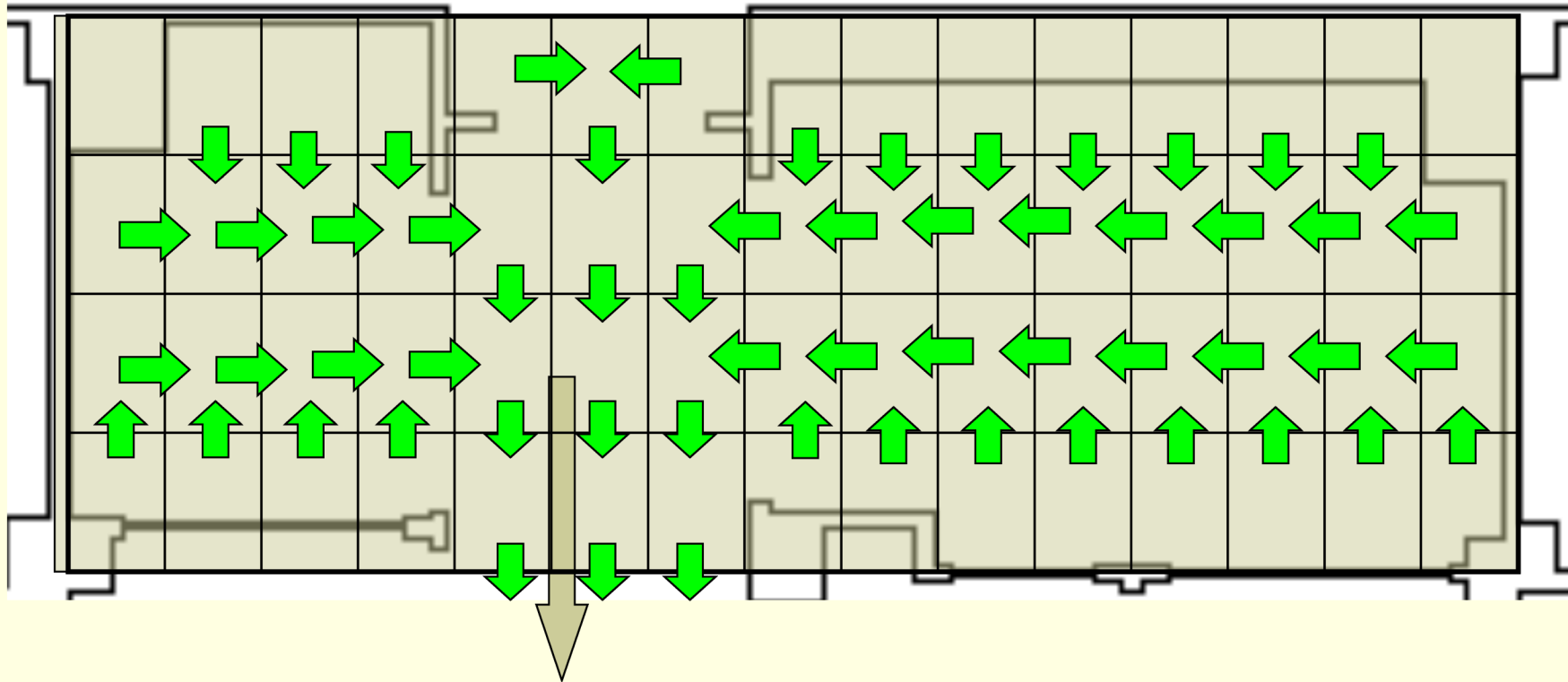
Illustration – Résolution d'un sous-modèles



Nouveau niveau hiérarchique possible...

Algorithmes hiérarchiques

Illustration – Résolution d'un sous-modèles

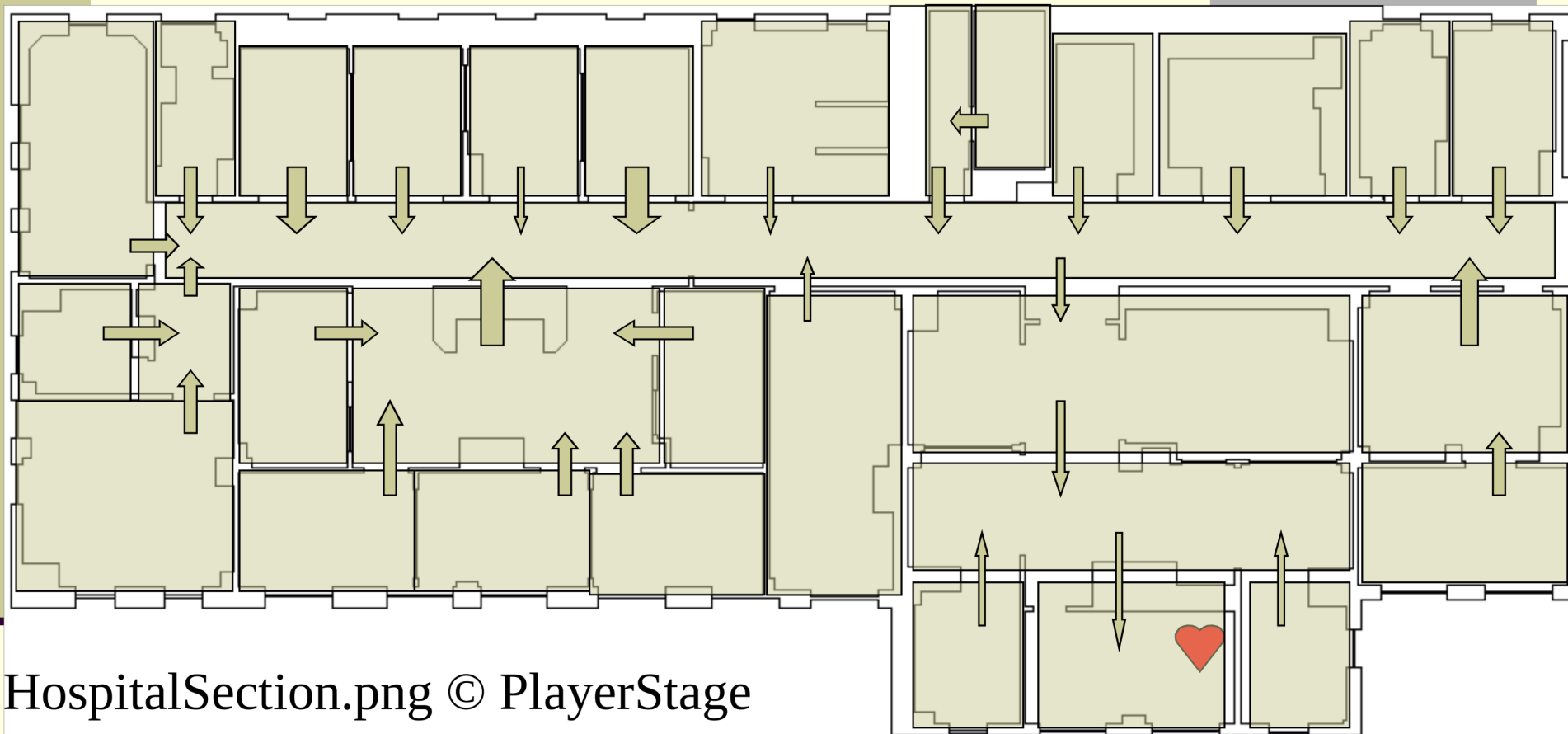


Ou modèle primitif...

Ici $15 \times 4 = 60$ états ... éventuellement $\times 4$ orientations $\rightarrow 240$ états

Algorithmes hiérarchiques

Illustration – Résolution des sous-modèles



→ La politique finale est la composée des sous-politiques

Algorithme hiérarchique

© Jenifer Barry, 2010

■ Hypothèses

- États buts absorbants, avec récompense nulle
- Toutes autres récompenses négatives

■ Principes

- Construit des macro-états
 - A partir des cycles du graphe original
- Construit un macro-modèle déterministe
 - Résolution = plus court chemin
- Résoud les sous-modèles
 - Récompense = vers le bon macro-état
 - Pénalité si vers un mauvais macro-état

Algorithme hiérarchique

© Jenifer Barry, 2010 (Construction des macro-états)

- Recherche cycles du modèle original
 - Recherche de boucles
 - Ex. Algorithme de Johnson
- Agrège les cycles en macro-états
 - → ensemble connexe
 - Tout sous-état peut atteindre tous les autres
 - Sans sortir du macro-état
 - → pas FORTEMENT connexe
 - Un état peut appartenir à un autre cycle
 - ⇒ pas de relation « propre » comme dans TVI
- Arrêt quand
 - Macro-états trop gros
 - Trop de macro-états

Algorithme hiérarchique

© Jenifer Barry, 2010 (Construction du macro-modèle)

- Approximation déterministe
 - Probabilité de changement de macro-état = 1
 - Récompense = moyenne($R \cdot P$)
 - Coût moyen pour changer d'état
- Coût de macro-transition
 - Moyenne des coûts de chemin
 - Moyenne sur tous les sous-états
 - Plus court chemin vers macro-état but
 - Utilise algorithme de Dijkstra

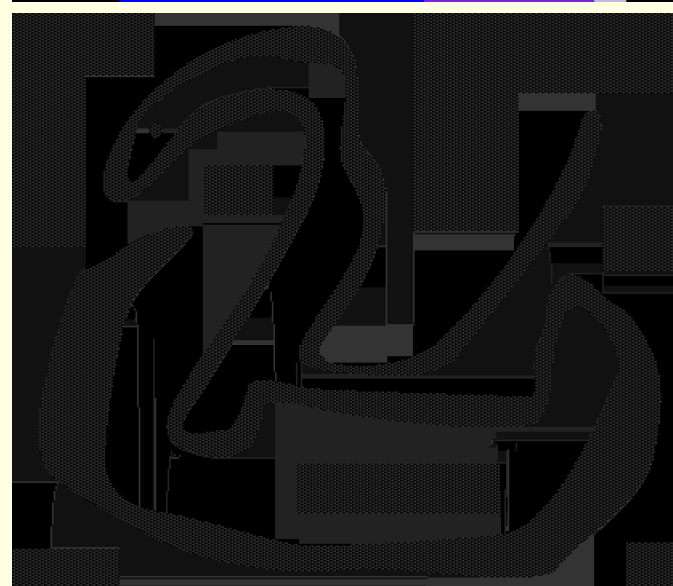
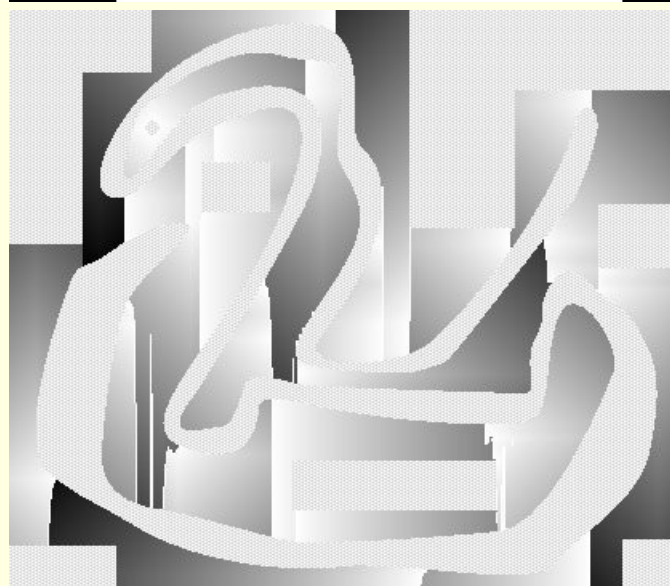
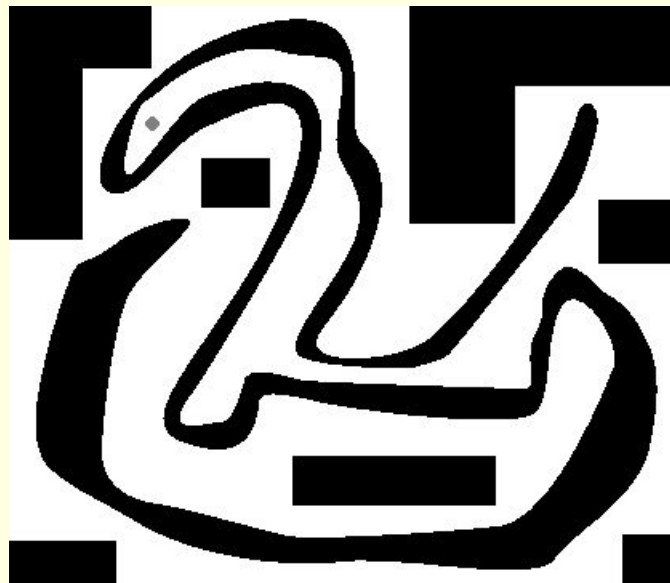
Algorithme hiérarchique

© Jenifer Barry, 2010 (Résolution du macro-modèle)

- Calcul du plus court chemin
 - Vers le macro-état but
 - Déterministe
 - \Rightarrow algorithme de Dijkstra
- \Rightarrow Politique déterministe
 - Les macro-niveaux
 - Recommencent avec le but du niveau supérieur
 - Le niveau primitif
 - Résout un MDP avec le but du niveau supérieur
 - Retour au modèle stochastique
 - Retour aux actions « normales »

Algorithme hiérarchique

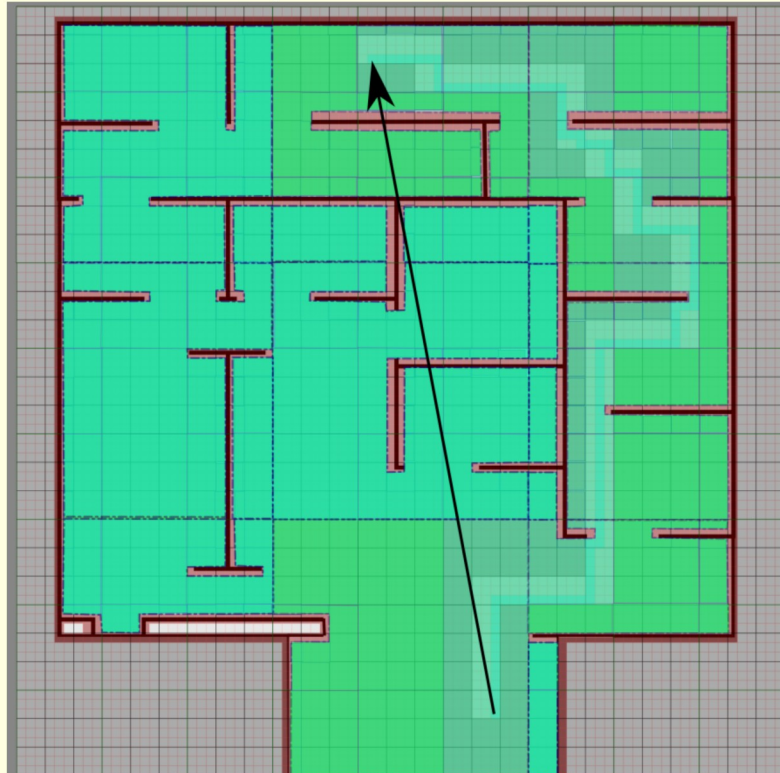
© Jenifer Barry, 2010 (Illustration)



Algorithme hiérarchique

Discussion

- Travaux en cours (cartographie)
 - Représentation hiérarchique de l'environnement
 - ⇒ Carte extensible à volonté
 - ⇒ Clusters « simples » à construire



Algorithme hiérarchique

Démonstration

Algorithmes Topologique

Exercice de modélisation

- On souhaite ajouter des données non-géographiques ?
 - Ajout d'une clé
 - Ajout d'une porte
 - L'agent ne peut franchir la porte que s'il a la clé

- Exercice : Modéliser ce problème !

Exercice de Modélisation 2

La vaccination multi-troupeaux

- Un problème d'éleveurs de vaches
- Un troupeau peut être
 - Sain
 - Vacciné
 - Infecté
- On se place dans un groupement d'éleveurs
 - 50 troupeaux ou +
- Un troupeau s'infecte par
 - Contact avec les voisins Infectés
 - Achats d'animaux

La vaccination multi-troupeaux

Dynamique

- Un troupeau Sain devient :
 - Vacciné dans 90% des cas si vaccination
 - Infecté dans 5% des cas (achats)
 - +10% des cas si un voisin est infecté
 - \Rightarrow 15% si un voisin infecté
 - \Rightarrow 5+10+9% si deux voisins infectés
 - \Rightarrow 5+10+9+8,1% si trois voisins infectés
 - ...
- Un troupeau Vacciné devient :
 - Sain au bout de 12 mois environ
- Un troupeau Infecté devient :
 - Vacciné dans 70% des cas si vaccination

La vaccination multi-troupeaux

Modélisez !

- Proposez une modélisation
 - Espace d'états principalement

- Discutez la taille du modèle
 - Nombre d'états ?
 - Technique de résolution ?