Coordination multi-robots terrestre et aérien

Emmanuel RACHELSON

ONERA/DCSD

2 septembre 2005

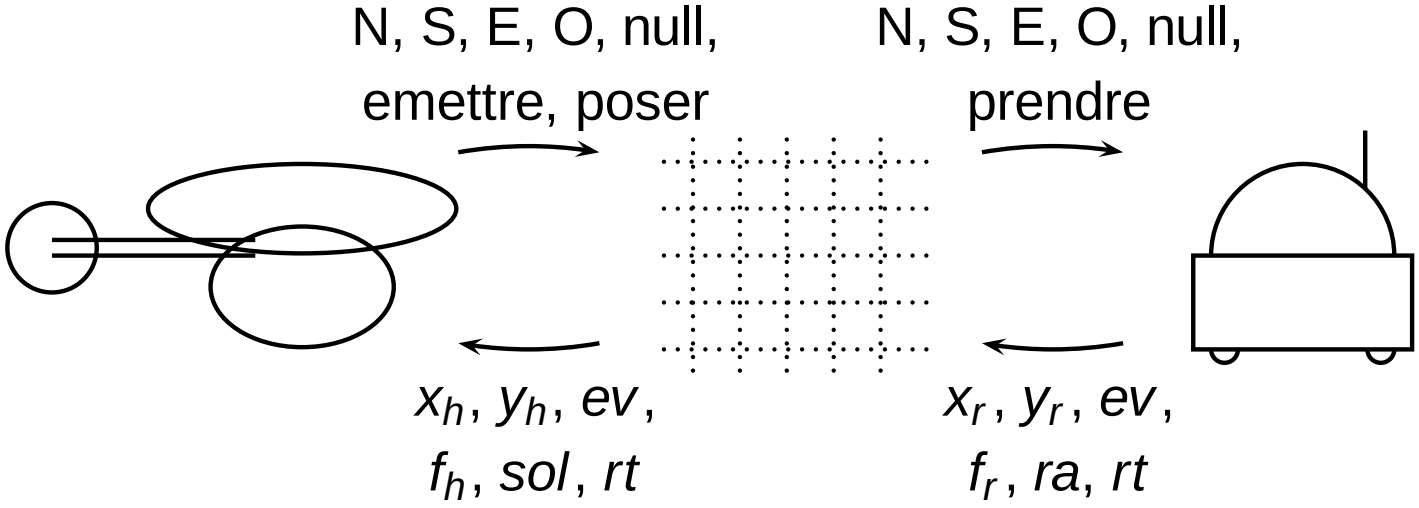
Plan

1. Introduction et problématique
2. Décision mono et multi-agents
3. Le temps, une variable cruciale aux problèmes de coopération
4. L'algorithme de résolution proposé

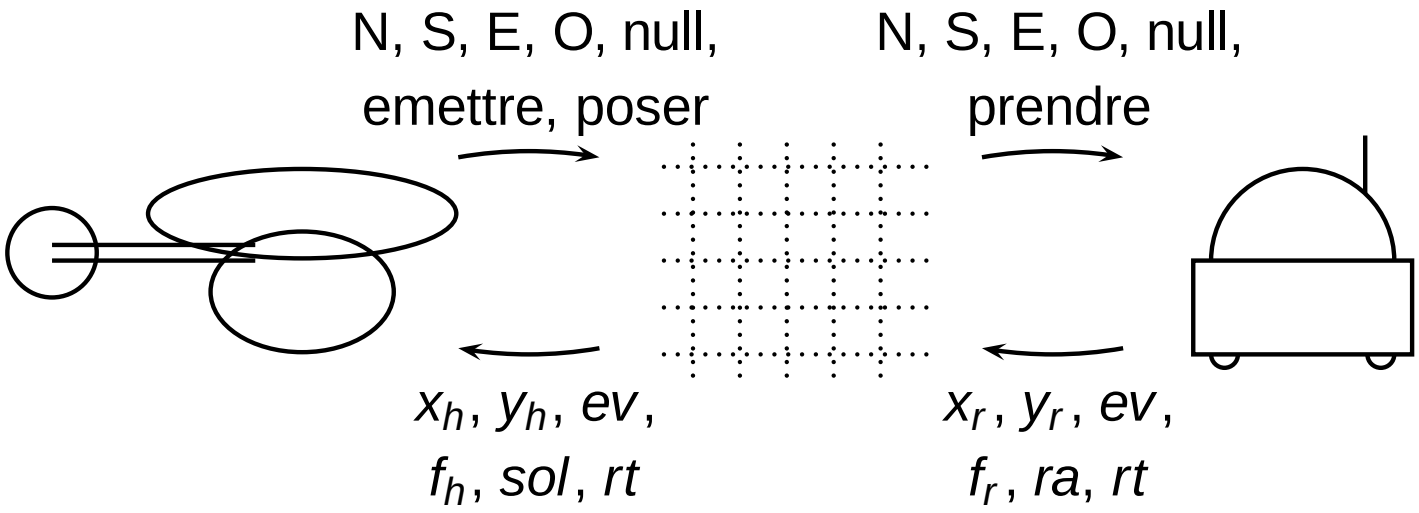
Plan

1. Introduction et problématique
2. Décision mono et multi-agents
3. Le temps, une variable cruciale aux problèmes de coopération
4. L'algorithme de résolution proposé

La problématique du binôme aéro-terrestre

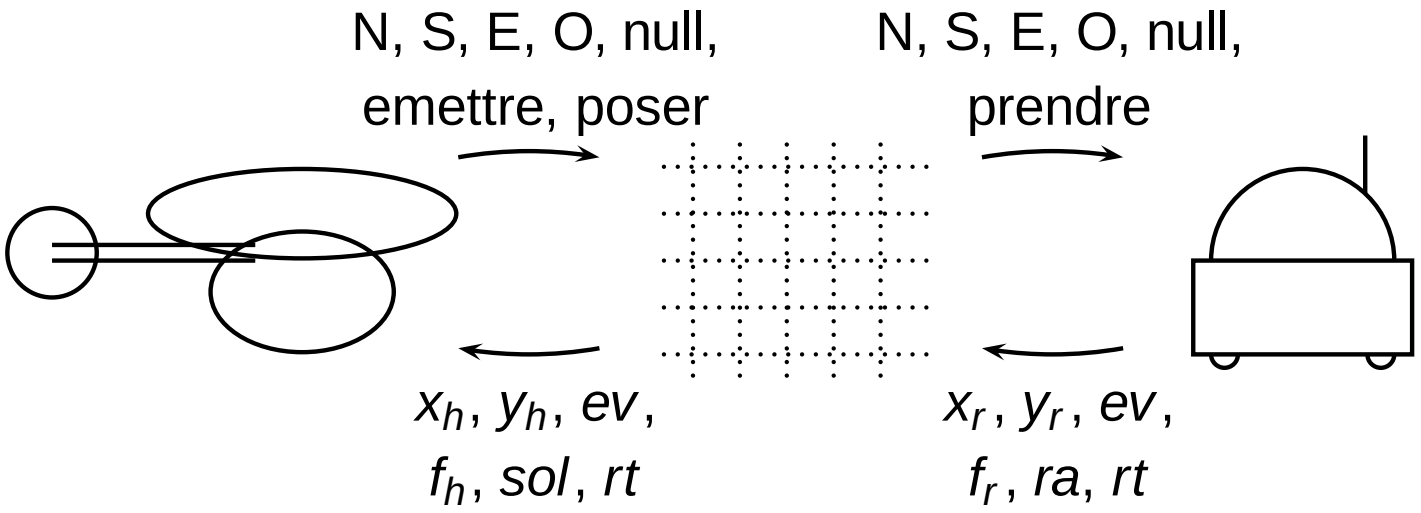


La problématique du binôme aéro-terrestre



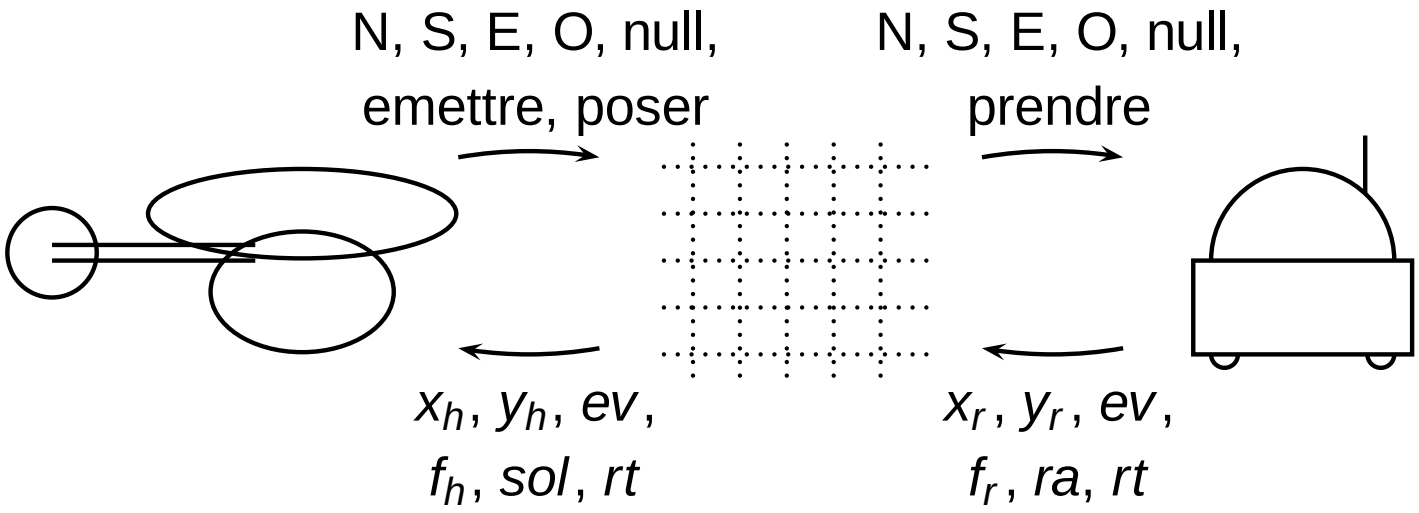
- Coopération

La problématique du binôme aéro-terrestre



- Coopération
- Coûts et récompenses : valeurs communes.

La problématique du binôme aéro-terrestre



- Coopération
- Coûts et récompenses : valeurs communes.
- Approche compatible avec une optimisation de stratégie en boucle fermée
  ↳ incertitudes, temps, multi-agents.

Plan

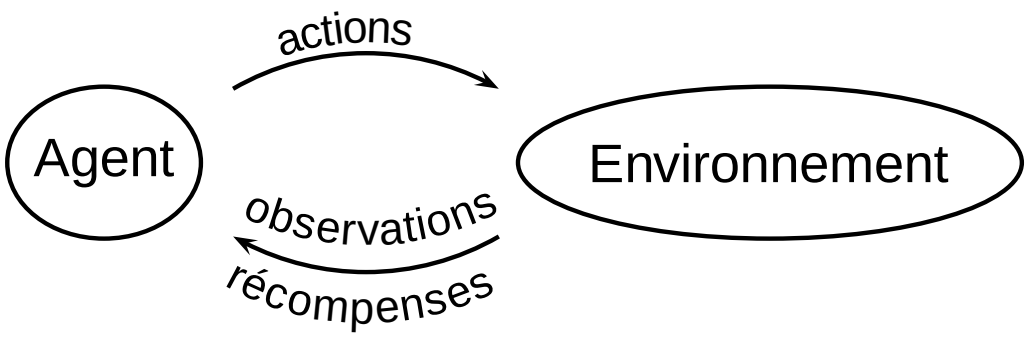
1 Introduction et problématique

2 Décision mono et multi-agents

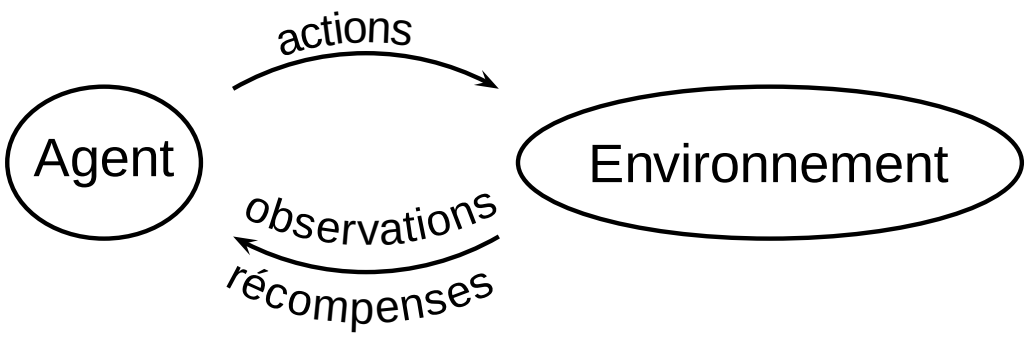
3 Le temps, une variable cruciale aux problèmes de coopération

4 L'algorithme de résolution proposé

Le problème de décision mono-agent

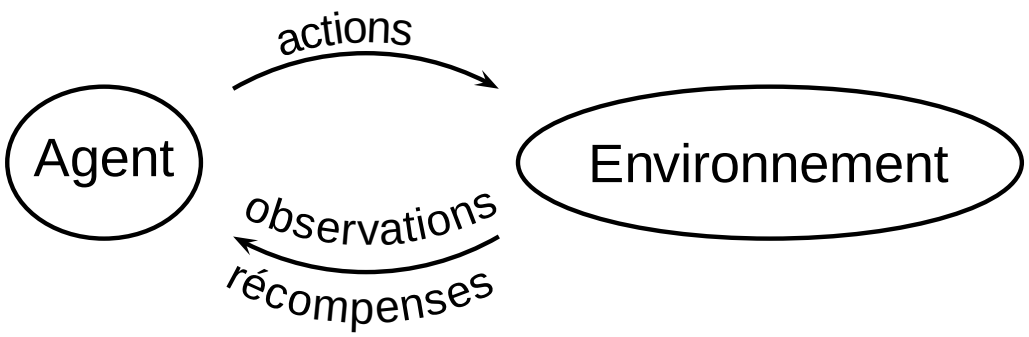


Le problème de décision mono-agent



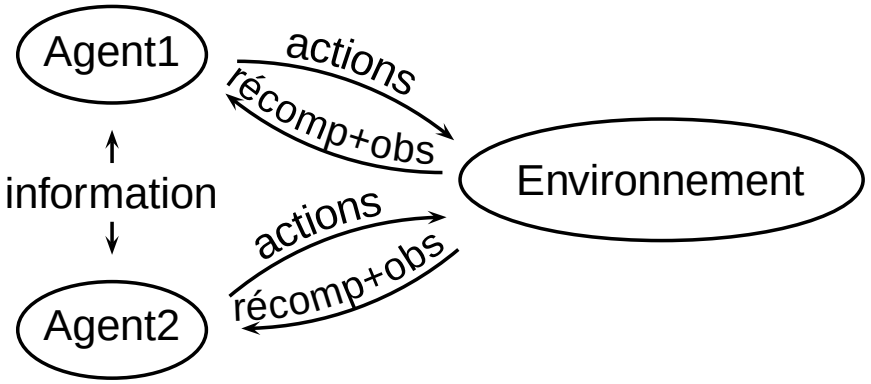
- stratégie en boucle fermée $\longrightarrow$ politique $\pi : s \mapsto a$

Le problème de décision mono-agent

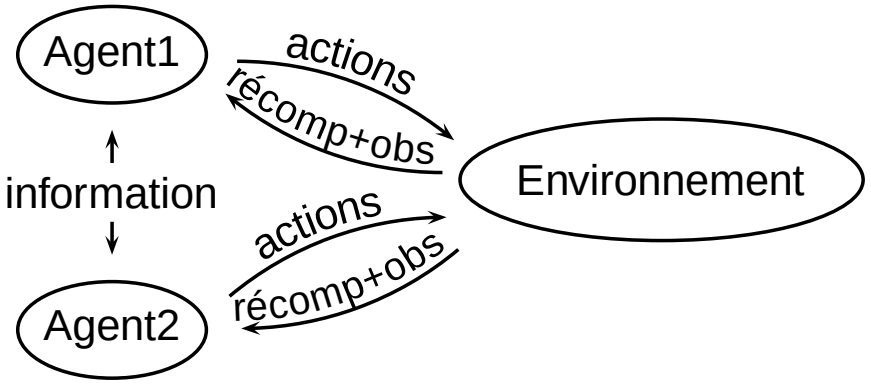


- stratégie en boucle fermée $\longrightarrow$ politique $\pi : s \mapsto a$
- différentes échelles de temps

... les différences avec le multi-agent

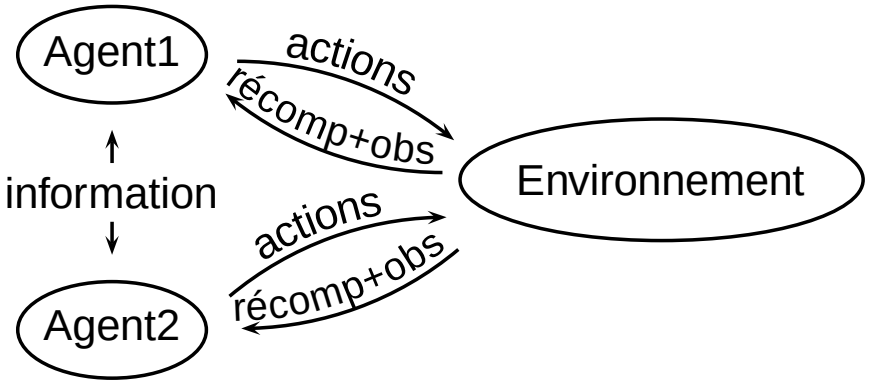


... les différences avec le multi-agent



- Approches réactives et comportements multi-agents
- Approches MDP (Zilberstein, Mouaddib)

... les différences avec le multi-agent



- Approches réactives et comportements multi-agents
- Approches MDP (Zilberstein, Mouaddib)
- Influence de l'action d'un agent sur l'état de l'univers pour l'autre.
- Deux stratégies individuelles optimales $\not\Rightarrow$ une stratégie conjointe optimale

  ↳ synchronisation et communication

Le cadre de résolution

- pas d'approche centralisée $\rightarrow$ chaque agent peut être physiquement isolé

Le cadre de résolution

- pas d'approche centralisée $\longrightarrow$ chaque agent peut être physiquement isolé
- on considère des sous-problèmes individuels $\longrightarrow$ décomposer et paralléliser la résolution (approches décomposées et factorisées, Parr, Boutilier, Hauskrecht, ...)

Le cadre de résolution

- pas d'approche centralisée $\longrightarrow$ chaque agent peut être physiquement isolé
- on considère des sous-problèmes individuels $\longrightarrow$ décomposer et paralléliser la résolution (approches décomposées et factorisées, Parr, Boutilier, Hauskrecht, ...)
- principe exploité : processus itératif d'amélioration de la valeur commune

Le cadre de résolution

- pas d'approche centralisée $\rightarrow$ chaque agent peut être physiquement isolé
- on considère des sous-problèmes individuels $\rightarrow$ décomposer et paralléliser la résolution (approches décomposées et factorisées, Parr, Boutilier, Hauskrecht, ...)
- principe exploité : processus itératif d'amélioration de la valeur commune
- amélioration individuelle de la valeur commune, nécessite info :
  - politique indiduelle
  - information sur les conséquences de la politique de l'autre agent (Parr)

Le cadre de résolution

- pas d'approche centralisée $\rightarrow$ chaque agent peut être physiquement isolé
- on considère des sous-problèmes individuels $\rightarrow$ décomposer et paralléliser la résolution (approches décomposées et factorisées, Parr, Boutilier, Hauskrecht, ...)
- principe exploité : processus itératif d'amélioration de la valeur commune
- amélioration individuelle de la valeur commune, nécessite info :
  - politique indiduelle
  - information sur les conséquences de la politique de l'autre agent (Parr)

Communication des conséquences de la politique du second agent sur le premier $\rightarrow$ utilisation de la notion de chronologie.

Plan

1 Introduction et problématique

2 Décision mono et multi-agents

3 Le temps, une variable cruciale aux problèmes de coopération

4 L'algorithme de résolution proposé

La notion de chronologie

- quantité d'information minimale pour l'agent qui reçoit
- quantité d'information suffisante pour définir comment change le problème de l'agent qui reçoit

La notion de chronologie

- quantité d'information minimale pour l'agent qui reçoit
- quantité d'information suffisante pour définir comment change le problème de l'agent qui reçoit

Principe :

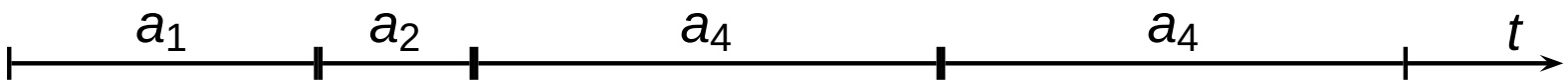
- hiérarchisation de l'espace d'action en niveaux d'abstraction $\rightarrow$ création de tâches ou macro-actions (approches décomposées (Sabbadin, Parr) ou approche MAXQ (Diettrich))

La notion de chronologie

- quantité d'information minimale pour l'agent qui reçoit
- quantité d'information suffisante pour définir comment change le problème de l'agent qui reçoit

Principe :

- hiérarchisation de l'espace d'action en niveaux d'abstraction $\rightarrow$ création de tâches ou macro-actions (approches décomposées (Sabbadin, Parr) ou approche MAXQ (Diettrich))
- chronologie = suite de macro-transitions les plus probables selon la politique courante et en partant de l'état initial.



Chaque macro-action a_i affecte le problème de l'autre agent.
L'information contenue dans une chronologie permet donc au second agent de planifier en tenant compte des intentions du premier.

L'introduction de la variable temps

chronologie $\Rightarrow$ dates

L'introduction de la variable temps

chronologie $\Rightarrow$ dates

différentes approches :

- MDP classique, échelle de temps synchrone connue.

L'introduction de la variable temps

chronologie $\Rightarrow$ dates
différentes approches :

- MDP classique, échelle de temps synchrone connue.
- Travaux de Mouaddib et Beynier, chronologie pré-établie, échéances et dates de décision connues.

L'introduction de la variable temps

chronologie $\Rightarrow$ dates
différentes approches :

- MDP classique, échelle de temps synchrone connue.
- Travaux de Mouaddib et Beynier, chronologie pré-établie, échéances et dates de décision connues.
- modèles à temps continu : SMDP, CTMDP, durées (Puterman)

L'introduction de la variable temps

chronologie $\Rightarrow$ dates
différentes approches :

- MDP classique, échelle de temps synchrone connue.
- Travaux de Mouaddib et Beynier, chronologie pré-établie, échéances et dates de décision connues.
- modèles à temps continu : SMDP, CTMDP, durées (Puterman)
- modèle SMDP avec dates (transparent suivant)

L'introduction de la variable temps

chronologie $\Rightarrow$ dates
différentes approches :

- MDP classique, échelle de temps synchrone connue.
- Travaux de Mouaddib et Beynier, chronologie pré-établie, échéances et dates de décision connues.
- modèles à temps continu : SMDP, CTMDP, durées (Puterman)
- modèle SMDP avec dates (transparent suivant)
- modèle discret avec dates (transparent suivant)

Les différents modèles intégrant le temps

Modèle SMDP : intégrer récompenses et probabilités de transition sur les durées pour chaque transition (départ, action, arrivée).

Les différents modèles intégrant le temps

Modèle SMDP : intégrer récompenses et probabilités de transition sur les durées pour chaque transition (départ, action, arrivée).

SMDP avec dates : enrichir le modèle SMDP d'un temps absolu
$\longrightarrow$ problèmes dont les modèles de récompense dépendent du temps.

Les différents modèles intégrant le temps

Modèle SMDP : intégrer récompenses et probabilités de transition sur les durées pour chaque transition (départ, action, arrivée).

SMDP avec dates : enrichir le modèle SMDP d'un temps absolu
$\longrightarrow$ problèmes dont les modèles de récompense dépendent du temps.

enjeu 1 : Limiter la quantité de calcul $\longrightarrow$ séparabilité des variables continues et discrètes.

Les différents modèles intégrant le temps

Modèle SMDP : intégrer récompenses et probabilités de transition sur les durées pour chaque transition (départ, action, arrivée).

SMDP avec dates : enrichir le modèle SMDP d'un temps absolu
$\rightarrow$ problèmes dont les modèles de récompense dépendent du temps.

enjeu 1 : Limiter la quantité de calcul $\rightarrow$ séparabilité des variables continues et discrètes.

enjeu 2 : Complexité de l'algorithme $\rightarrow$ nombre de fonctions intégrales à calculer

$$\propto \begin{pmatrix} nb\ segments\ temporels \\ dans\ modèle\ récomp. \end{pmatrix} \begin{pmatrix} nb\quad d'itérations \\ d'optimisation \end{pmatrix}$$

Deux cas d'applications des modèles temporels

- cas particulier du modèle continu : les fonctions "exp-aff"
- le modèle discret $\longrightarrow$ pas de temps égal au pgcd des durées de transition

Plan

1. Introduction et problématique
2. Décision mono et multi-agents
3. Le temps, une variable cruciale aux problèmes de coopération
4. L'algorithme de résolution proposé

L'algorithme d'itération de la chronologie

Drone hélicoptère

résolution du pb de base en mono-agent $\longrightarrow \pi_1$

L'algorithme d'itération de la chronologie

Drone hélicoptère

résolution du pb de base en mono-agent $\longrightarrow \pi_1$

$\downarrow$

établissement d'une chronologie $\longrightarrow Q_1$

L'algorithme d'itération de la chronologie

Drone hélicoptère

résolution du pb de base en mono-agent $\longrightarrow \pi_1$

$\downarrow$

établissement d'une chronologie $\longrightarrow Q_1$

~~échange des chronologies~~

L'algorithme d'itération de la chronologie

Drone hélicoptère

résolution du pb de base en mono-agent $\longrightarrow \pi_1$

$\downarrow$

établissement d'une chronologie $\longrightarrow Q_1$

échange des chronologies

modification du pb de base :

- ajout des dates dans l'état
- modification des récompenses
- mise à jour du modèle de transition

L'algorithme d'itération de la chronologie

Drone hélicoptère

résolution du pb de base en mono-agent $\longrightarrow \pi_1$

$\downarrow$

établissement d'une chronologie $\longrightarrow Q_1$

échange des chronologies

modification du pb de base :

- ajout des dates dans l'état
- modification des récompenses
- mise à jour du modèle de transition

$\downarrow$

résolution du pb augmenté $\longrightarrow \pi_2$

L'algorithme d'itération de la chronologie

Drone hélicoptère

résolution du pb de base en mono-agent $\longrightarrow \pi_1$

$\downarrow$

établissement d'une chronologie $\longrightarrow Q_1$

échange des chronologies

modification du pb de base :

- ajout des dates dans l'état
- modification des récompenses
- mise à jour du modèle de transition

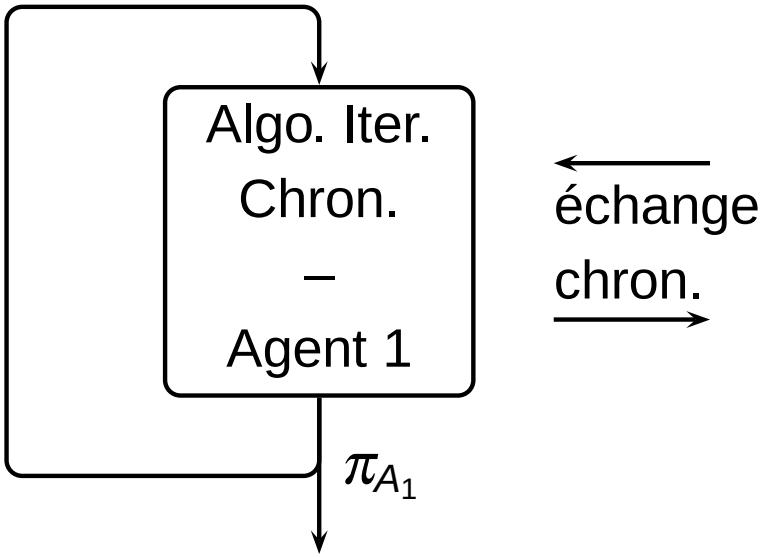
$\downarrow$

résolution du pb augmenté $\longrightarrow \pi_2$

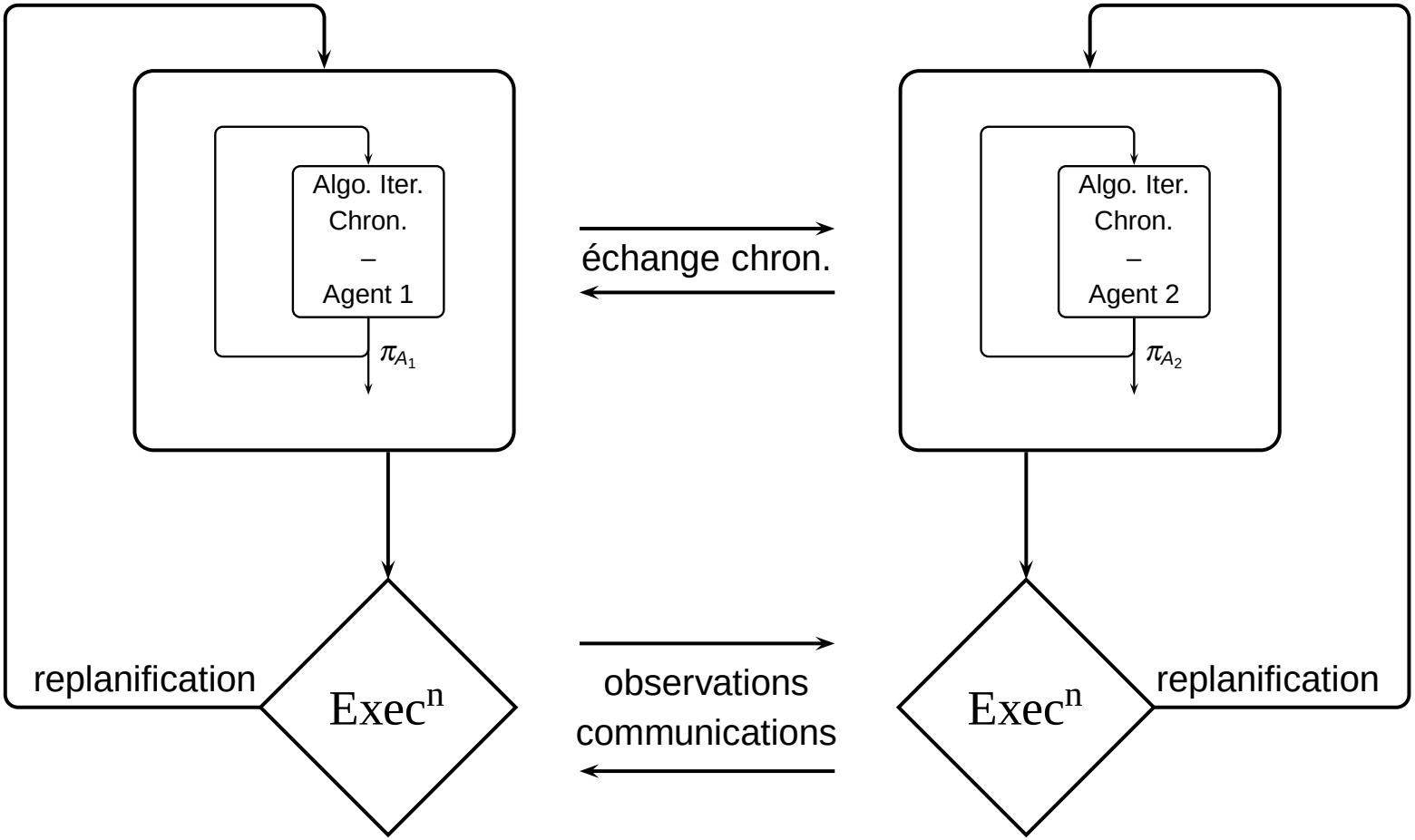
$\downarrow \ldots$

L'algorithme d'itération de la chronologie

- boucle d'optimisation hors-ligne
- algorithme de recherche itérative d'un politique conjointe
- pas de formulation explicite du problème global
- intégration des modèles temporels et des outils mono-agents



L'algorithme dans le contexte d'une exécution en ligne



Compromis : exécution/replanification execution/optimisation

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps
  - Un algorithme de résolution des problèmes de coopération

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps
  - Un algorithme de résolution des problèmes de coopération
  - Une approche de décomposition du problème global pour des résolutions individuelles

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps
  - Un algorithme de résolution des problèmes de coopération
  - Une approche de décomposition du problème global pour des résolutions individuelles
  - Extensible aux problèmes multi-agents en général

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps
  - Un algorithme de résolution des problèmes de coopération
  - Une approche de décomposition du problème global pour des résolutions individuelles
  - Extensible aux problèmes multi-agents en général
  - Permet d'utiliser tous les travaux existants dans le cadre MDP et MDP distribués

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps
  - Un algorithme de résolution des problèmes de coopération
  - Une approche de décomposition du problème global pour des résolutions individuelles
  - Extensible aux problèmes multi-agents en général
  - Permet d'utiliser tous les travaux existants dans le cadre MDP et MDP distribués
- Perspectives
  - Intégrer les méthodes issues des MDPs classiques et en implémenter des versions "automatiques" pour l'exécution de l'algorithme.

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps
  - Un algorithme de résolution des problèmes de coopération
  - Une approche de décomposition du problème global pour des résolutions individuelles
  - Extensible aux problèmes multi-agents en général
  - Permet d'utiliser tous les travaux existants dans le cadre MDP et MDP distribués
- Perspectives
  - Intégrer les méthodes issues des MDPs classiques et en implémenter des versions "automatiques" pour l'exécution de l'algorithme.
  - Apprentissage par renforcement et extensions

Conclusion

- Contribution
  - Deux nouveaux modèles de prise en compte du temps
  - Un algorithme de résolution des problèmes de coopération
  - Une approche de décomposition du problème global pour des résolutions individuelles
  - Extensible aux problèmes multi-agents en général
  - Permet d'utiliser tous les travaux existants dans le cadre MDP et MDP distribués
- Perspectives
  - Intégrer les méthodes issues des MDPs classiques et en implémenter des versions "automatiques" pour l'exécution de l'algorithme.
  - Apprentissage par renforcement et extensions
  - Compromis exécution/replanification et exécution/optimisation