



Analyse syntaxique par SMA adaptatifs : apport à la construction d'ontologies

Kévin Ottens

DEA RCFR - IRIT

Equipes

Systemes Multi-Agents Coopératifs

et

Conception de Systemes Coopératifs

Plan

- Positionnement du sujet
- Domaines concernés
- Système implémenté
- Bilan & Perspectives

Présentation du projet

- Construction et maintenance d'ontologies à partir de textes
- DEA : se limiter à un étiquetage simple
- Déterminer la pertinence de l'approche Adaptative Multi-Agent Systems (AMAS) au niveau syntaxique
 - Travail d'état de l'art à réaliser
 - a donné lieu à une communication lors d'une journée d'étude de l'Association pour le Traitement Automatique des Langues (ATALA)
 - Travail expérimental à partir d'une implémentation

A l'intersection de trois domaines

- La construction d'ontologies
 - Domaine actif de l'Ingénierie des Connaissances (IC)
 - Maintenance souvent oubliée...
 - Besoin d'améliorations pour réduire les coûts

A l'intersection de trois domaines

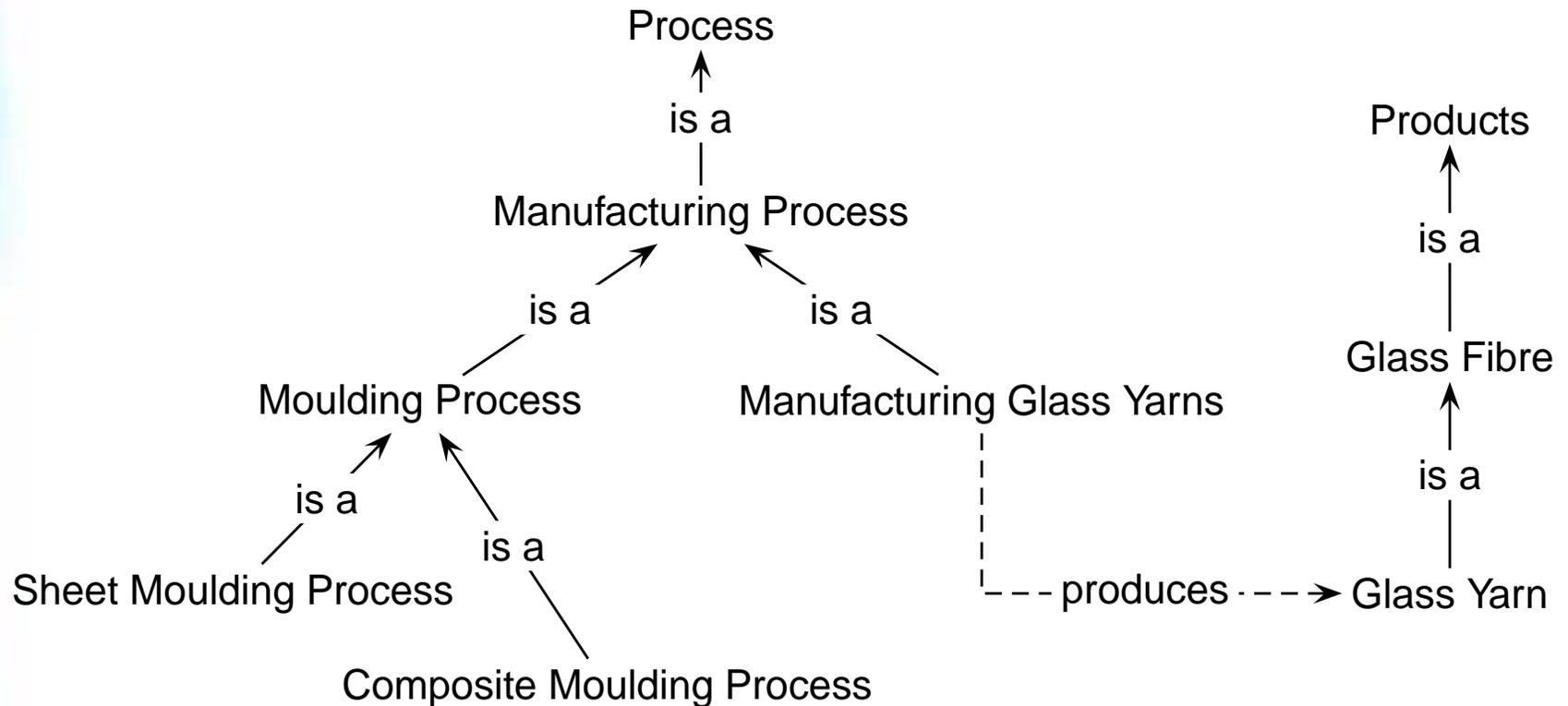
- La construction d'ontologies
 - Domaine actif de l'Ingénierie des Connaissances (IC)
 - Maintenance souvent oubliée...
 - Besoin d'améliorations pour réduire les coûts
- Le Traitement Automatique de la Langue (TAL)
 - Construction d'ontologies à partir de textes
 - Systèmes distribués d'analyse syntaxique

A l'intersection de trois domaines

- La construction d'ontologies
 - Domaine actif de l'Ingénierie des Connaissances (IC)
 - Maintenance souvent oubliée...
 - Besoin d'améliorations pour réduire les coûts
- Le Traitement Automatique de la Langue (TAL)
 - Construction d'ontologies à partir de textes
 - Systèmes distribués d'analyse syntaxique
- Les Systèmes Multi-Agents (SMA)
 - Un SMA au service du TAL
 - Un SMA au service de la construction d'ontologies
 - Intérêt de l'approche AMAS

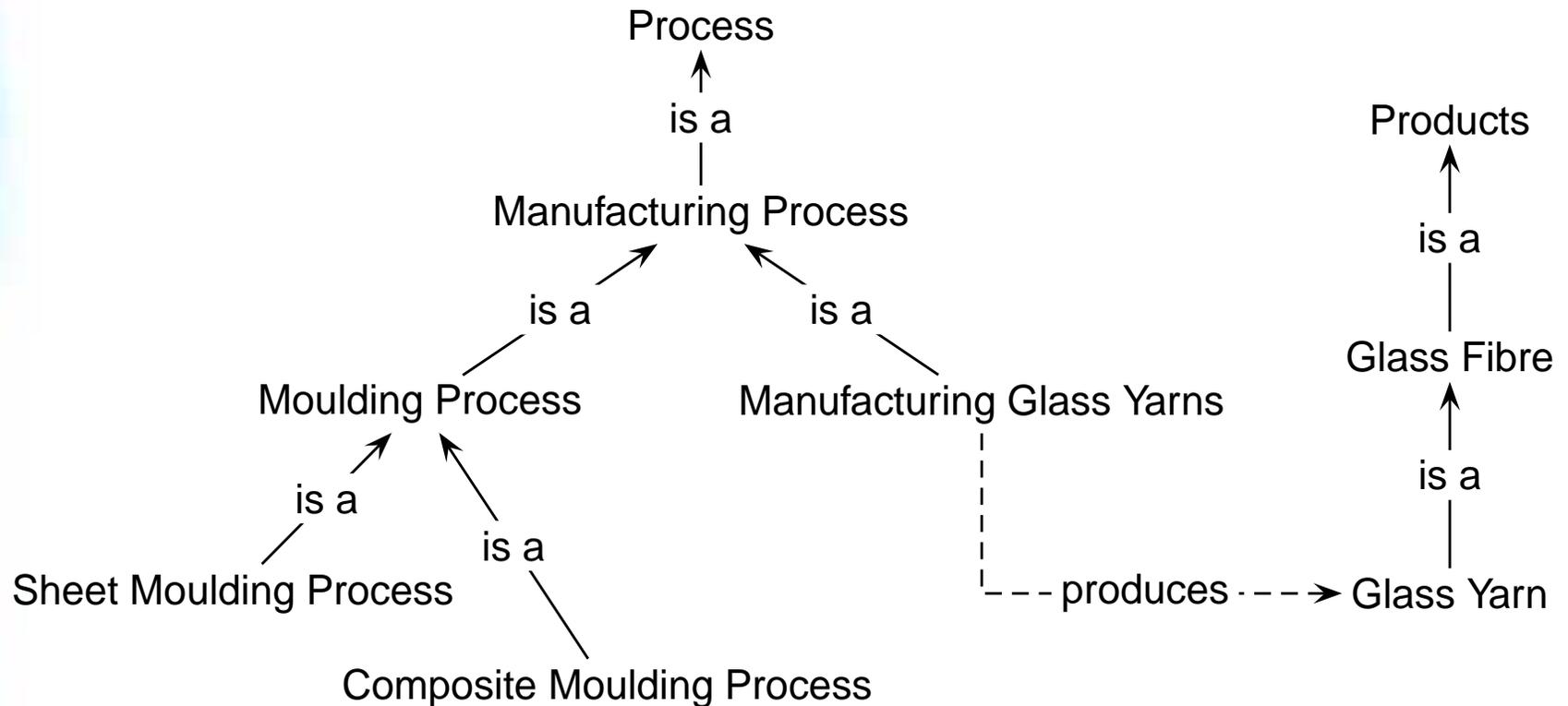
Construction d'ontologies

- « ontologies » → objets à construire
- Structuration conceptuelle



Construction d'ontologies

- « ontologies » → objets à construire
 - Structuration conceptuelle
 - Interopérabilité
 - Diversité de ces objets



Construction d'ontologies

- « ontologies » → objets à construire
 - Structuration conceptuelle
 - Interopérabilité
 - Diversité de ces objets
- Construction difficile
 - Nécessité de méthodes
 - Biais omniprésent

Construction d'ontologies

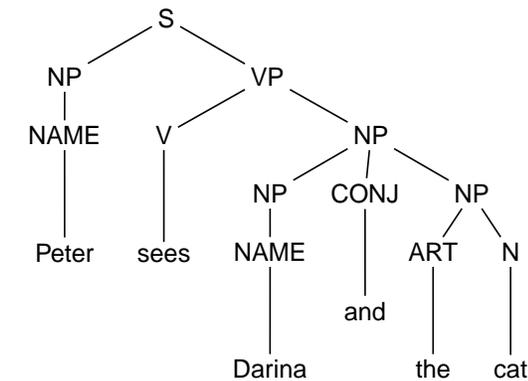
- « ontologies » → objets à construire
 - Structuration conceptuelle
 - Interopérabilité
 - Diversité de ces objets
- Construction difficile
 - Nécessité de méthodes
 - Biais omniprésent
 - Construction à partir de textes
 - Richesse terminologique
 - Formalisation plus faible

Analyse syntaxique

- Segmenter + Identifier + Relier

Analyse syntaxique

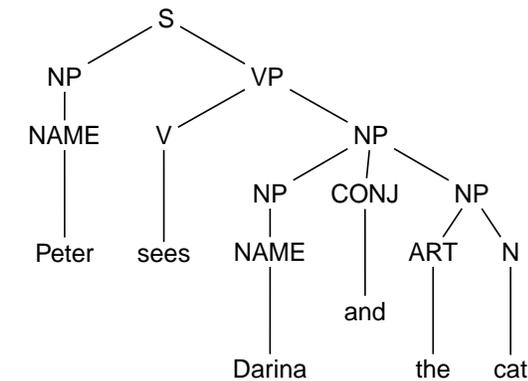
- Segmenter + Identifier + Relier
- Deux principales approches à base de grammaires
 - Grammaires syntagmatiques (Chomsky)



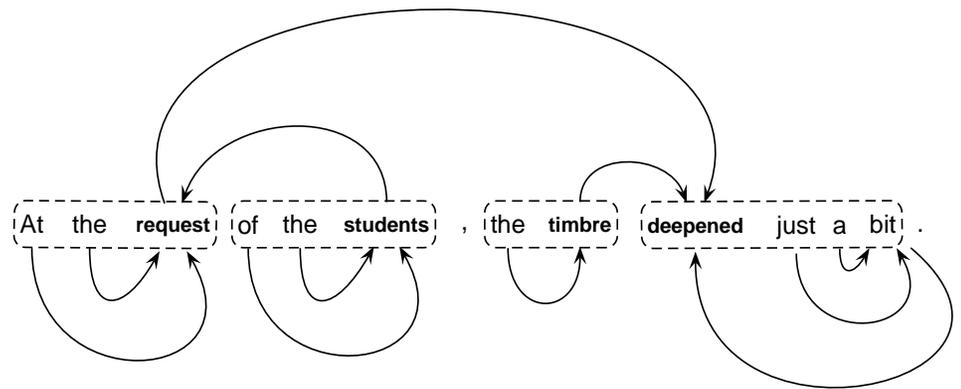
(Syntagmatique)

Analyse syntaxique

- Segmenter + Identifier + Relier
- Deux principales approches à base de grammaires
 - Grammaires syntagmatiques (Chomsky)
 - Grammaires de dépendances (Tesnière)



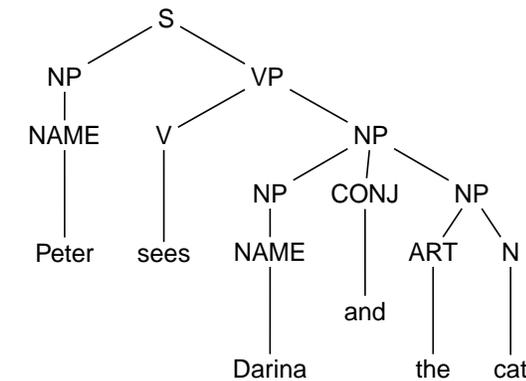
(Syntagmatique)



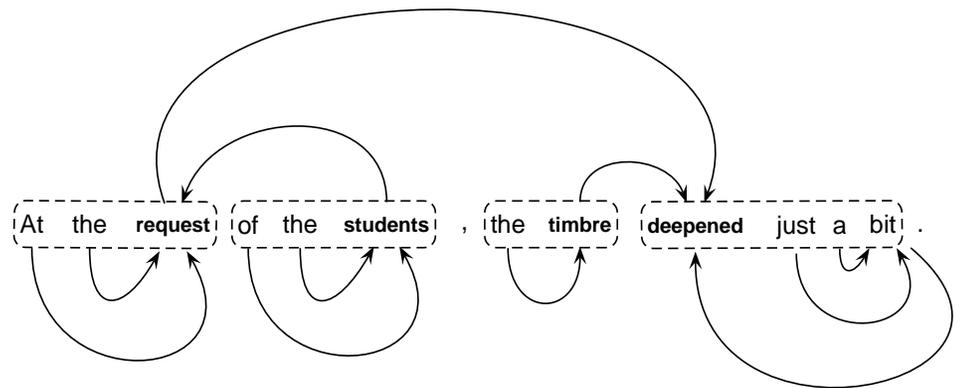
(Dépendances)

Analyse syntaxique

- Segmenter + Identifier + Relier
- Deux principales approches à base de grammaires
 - Grammaires syntagmatiques (Chomsky)
 - Grammaires de dépendances (Tesnière)



(Syntagmatique)



(Dépendances)

- Approches sans ressource très peu courantes
 - Extraction de termes
 - Regroupement en classes conceptuelles
- Principalement étude des cooccurrences

TAL pour la construction d'ontologies

- Repérage de formes linguistiques révélatrices de connaissances

TAL pour la construction d'ontologies

- Repérage de formes linguistiques révélatrices de connaissances
- Plusieurs types d'outils utilisés
 - Concordanciers
 - Extracteurs de termes
 - ...

TAL pour la construction d'ontologies

- Repérage de formes linguistiques révélatrices de connaissances
- Plusieurs types d'outils utilisés
 - Concordanciers
 - Extracteurs de termes
 - ...
- Utilisation de connaissances linguistiques spécifiques
 - Lexiques
 - Listes de morphèmes
 - Règles d'analyse grammaticale, syntaxique ou lexicale

Systeme Multi-Agent

- Environnement du systeme
- Agents autonomes

Systeme Multi-Agent

- Environnement du systeme
- Agents autonomes
 - Competences
 - Représentation de soi et de l'environnement
 - Perceptions
 - Actions

Systeme Multi-Agent

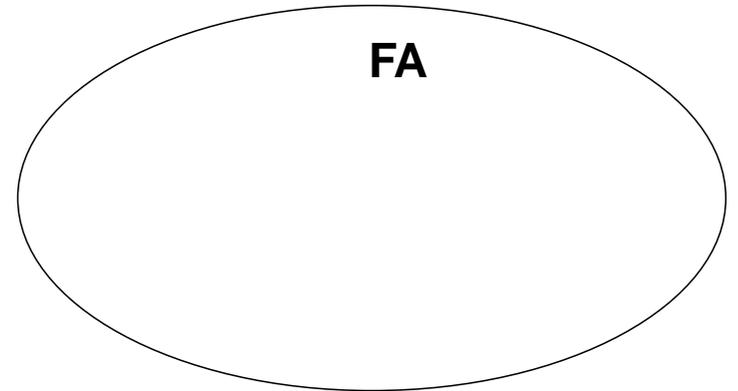
- Environnement du systeme
- Agents autonomes
 - Competences
 - Représentation de soi et de l'environnement
 - Perceptions
 - Actions
- Interactions entre agents
 - Communication par messages
 - Communication par l'environnement physique
- Comportement collectif

Systeme Multi-Agent

- Environnement du systeme
- Agents autonomes
 - Competences
 - Représentation de soi et de l'environnement
 - Perceptions
 - Actions
- Interactions entre agents
 - Communication par messages
 - Communication par l'environnement physique
- Comportement collectif
- Analyses syntaxiques distribuées
 - Multi-Experts : HEARSAY II, CARMEL
 - Découpage fonctionnel : TALISMAN
 - Découpage objet : HACTAR

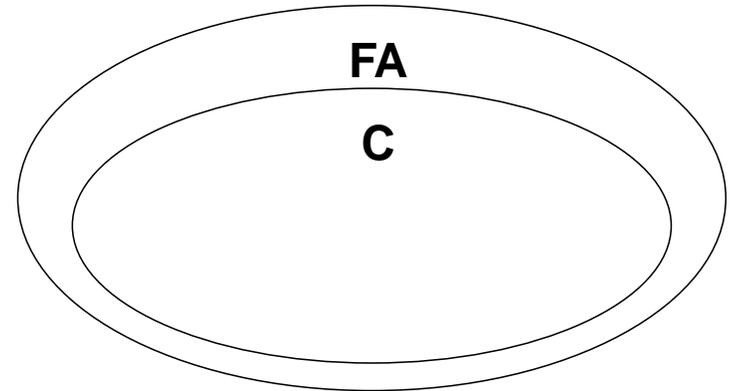
Théorie des AMAS

- Trois classes de systèmes :
 - Fonctionnellement adéquats



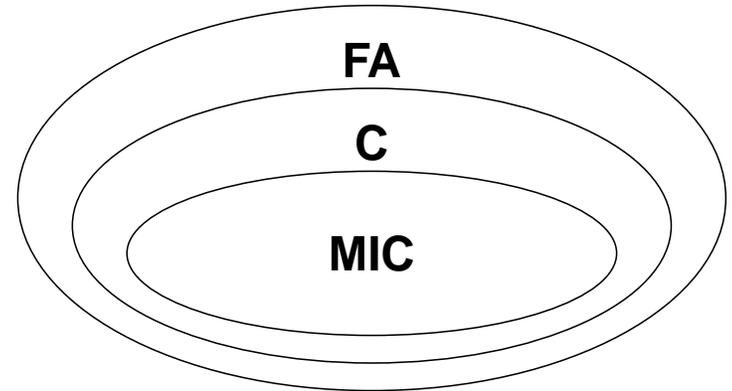
Théorie des AMAS

- Trois classes de systèmes :
 - Fonctionnellement adéquats
 - Coopératifs



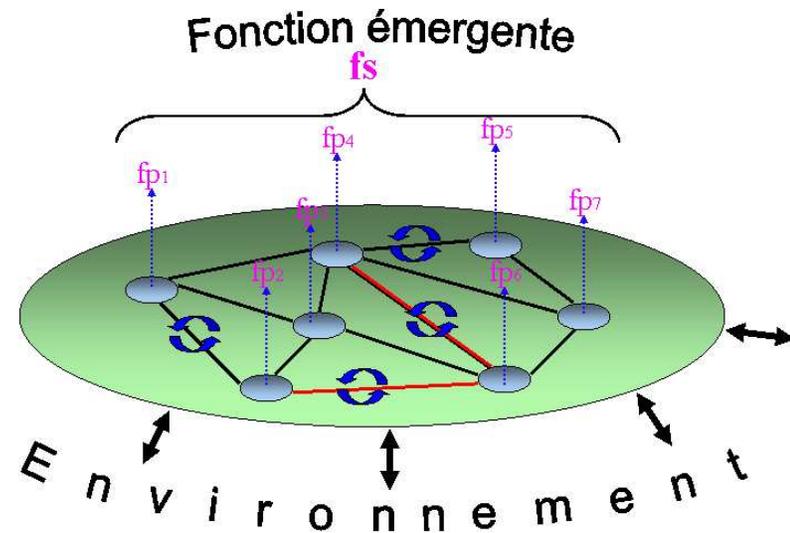
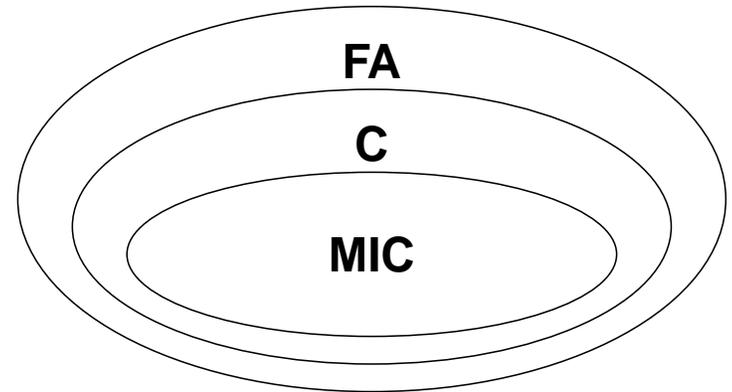
Théorie des AMAS

- Trois classes de systèmes :
 - Fonctionnellement adéquats
 - Coopératifs
 - A milieu intérieur coopératif



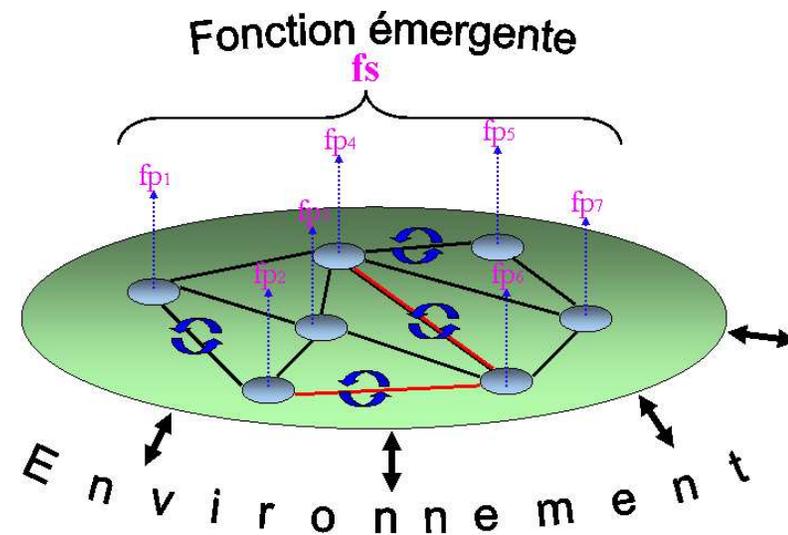
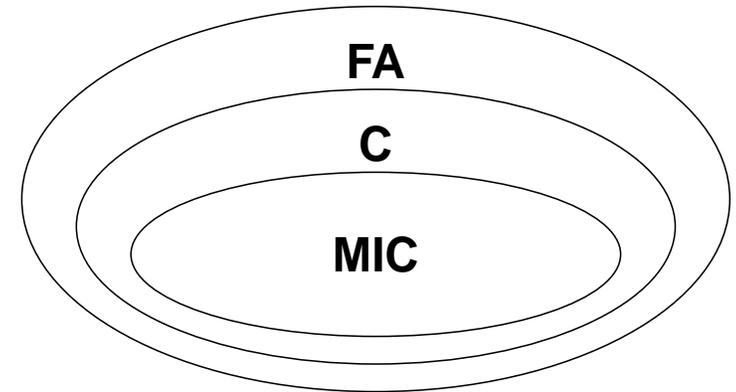
Théorie des AMAS

- Trois classes de systèmes :
 - Fonctionnellement adéquats
 - Coopératifs
 - A milieu intérieur coopératif
- "Adaptative Multi-Agent Systems"



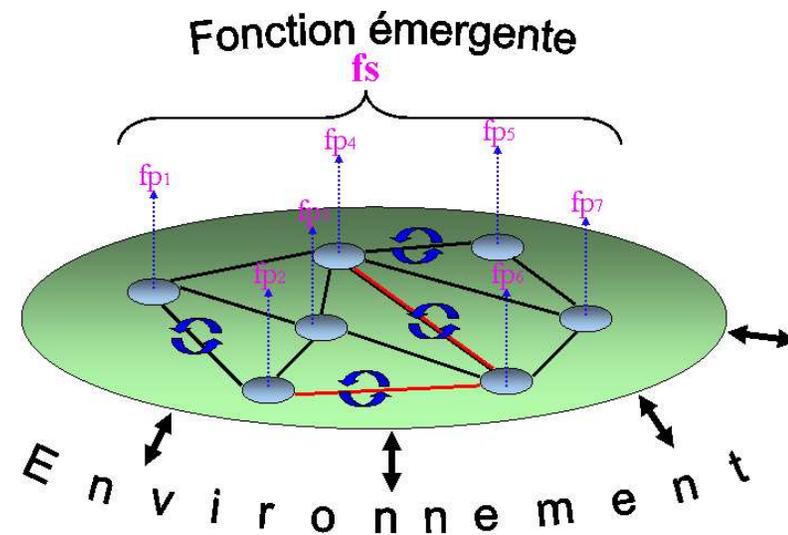
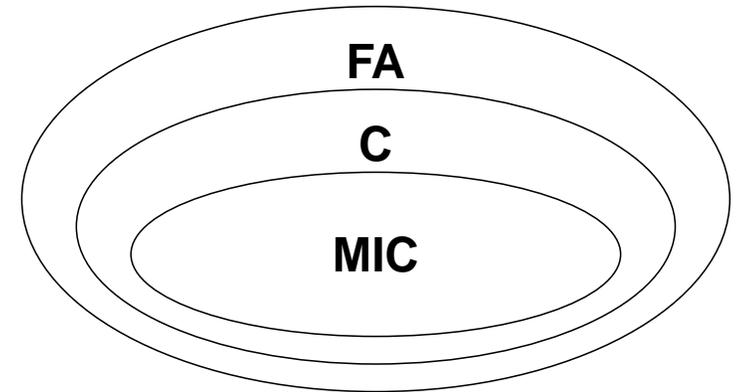
Théorie des AMAS

- Trois classes de systèmes :
 - Fonctionnellement adéquats
 - Coopératifs
 - A milieu intérieur coopératif
- "Adaptative Multi-Agent Systems"
 - Chaque agent produit une fonction f_{P_i}



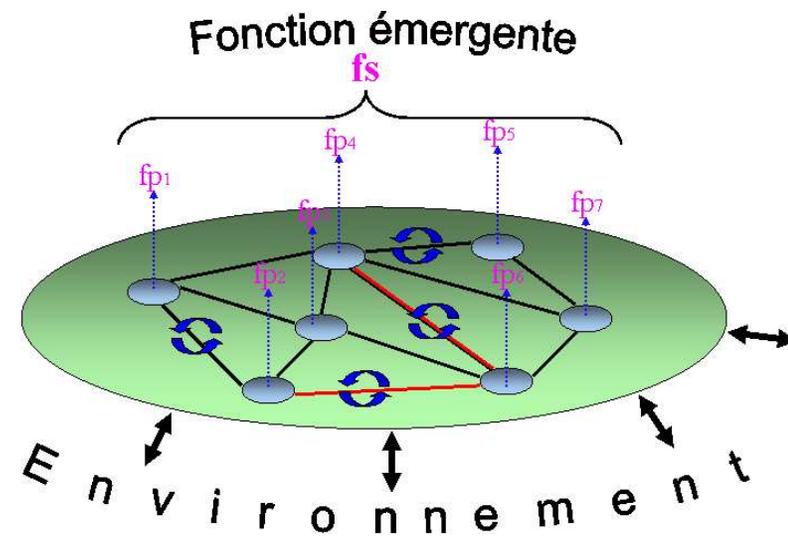
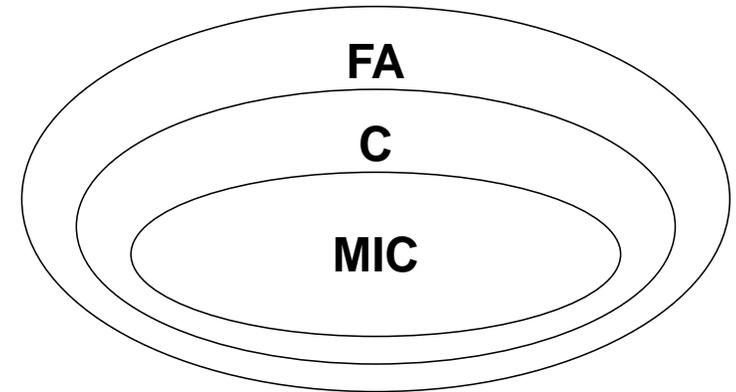
Théorie des AMAS

- Trois classes de systèmes :
 - Fonctionnellement adéquats
 - Coopératifs
 - A milieu intérieur coopératif
- "Adaptative Multi-Agent Systems"
 - Chaque agent produit une fonction f_{P_i}
 - Leurs liens engendrent la fonction f_S



Théorie des AMAS

- Trois classes de systèmes :
 - Fonctionnellement adéquats
 - Coopératifs
 - A milieu intérieur coopératif
- "Adaptative Multi-Agent Systems"
 - Chaque agent produit une fonction f_{P_i}
 - Leurs liens engendrent la fonction f_S
 - Adaptation par auto-organisation



Principes du système

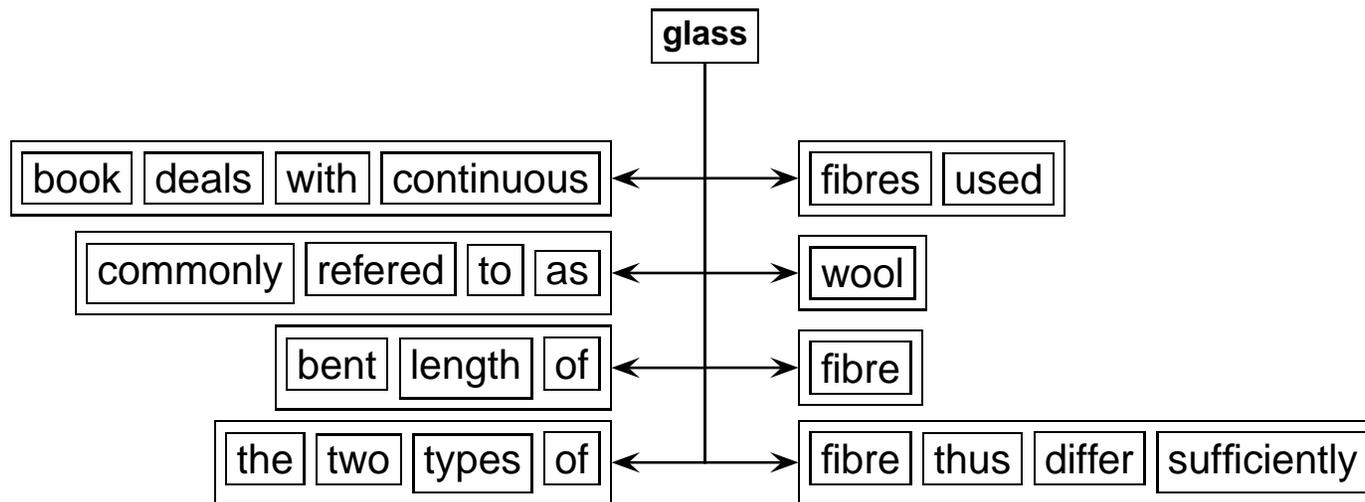
- Système multi-agent composé d'agents "mots"

Principes du système

- Système multi-agent composé d'agents "mots"
- Utilisation de confiances, trois rôles (Verbe, Nom, Fonctionnel)

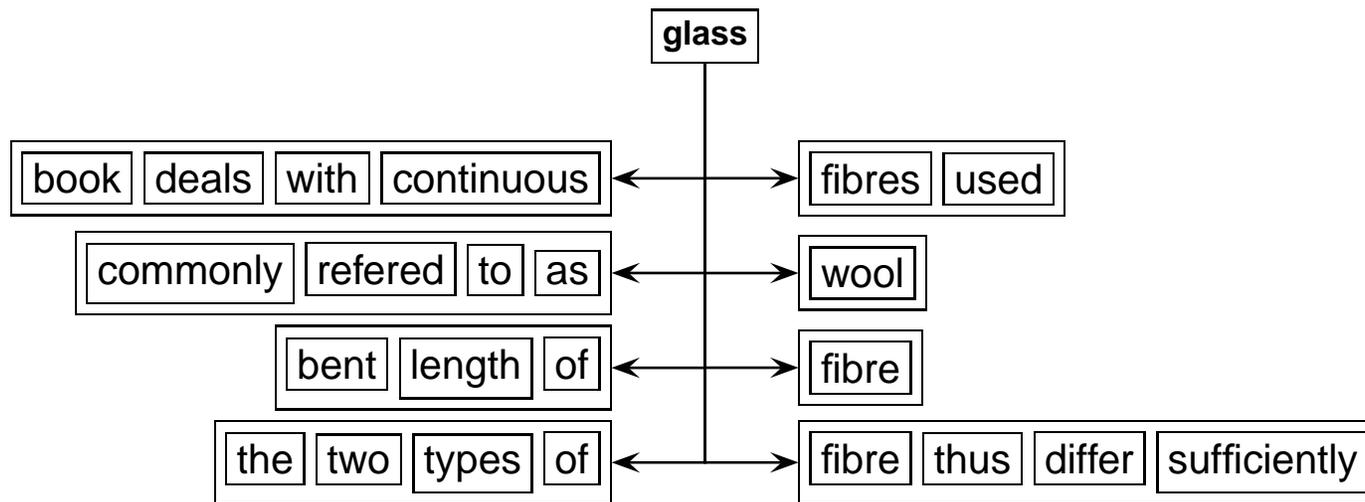
Principes du système

- Système multi-agent composé d'agents "mots"
- Utilisation de confiances, trois rôles (Verbe, Nom, Fonctionnel)
- Environnement uniquement social



Principes du système

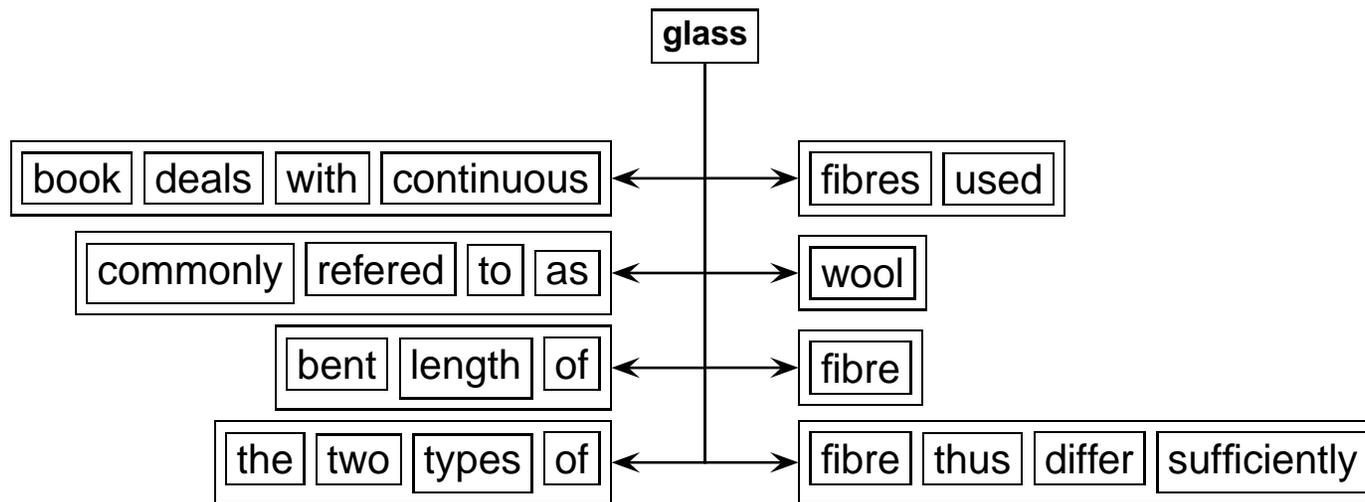
- Système multi-agent composé d'agents "mots"
- Utilisation de confiances, trois rôles (Verbe, Nom, Fonctionnel)
- Environnement uniquement social



- Fonctionnement sans ressource linguistique

Principes du système

- Système multi-agent composé d'agents "mots"
- Utilisation de confiances, trois rôles (Verbe, Nom, Fonctionnel)
- Environnement uniquement social



- Fonctionnement sans ressource linguistique
- Approches statistiques

Compétences des agents

■ Analyse contextuelle

The possibility of drawing heat-softened glass into fine fibres was known to glassmakers in antiquity and is older than the technique of glass blowing.

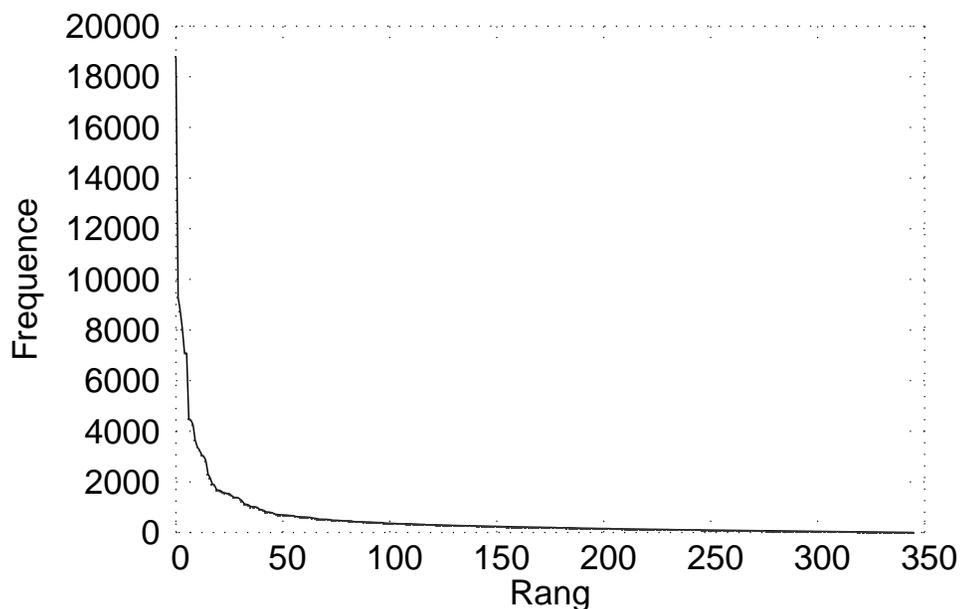
Compétences des agents

■ Analyse contextuelle

The possibility of drawing heat-softened glass into fine fibres was known to glassmakers in antiquity and is older than the technique of glass blowing.

■ Détermination des confiances initiales (deux lois)

■ Basée sur la loi de Zipf



Compétences des agents

- Analyse contextuelle

The possibility of drawing heat-softened glass into fine fibres was known to glassmakers in antiquity and is older than the technique of glass blowing.

- Détermination des confiances initiales (deux lois)
- Régulation locale de population

Compétences des agents

■ Analyse contextuelle

The possibility of drawing heat-softened glass into fine fibres was known to glassmakers in antiquity and is older than the technique of glass blowing.

■ Détermination des confiances initiales (deux lois)

■ Régulation locale de population

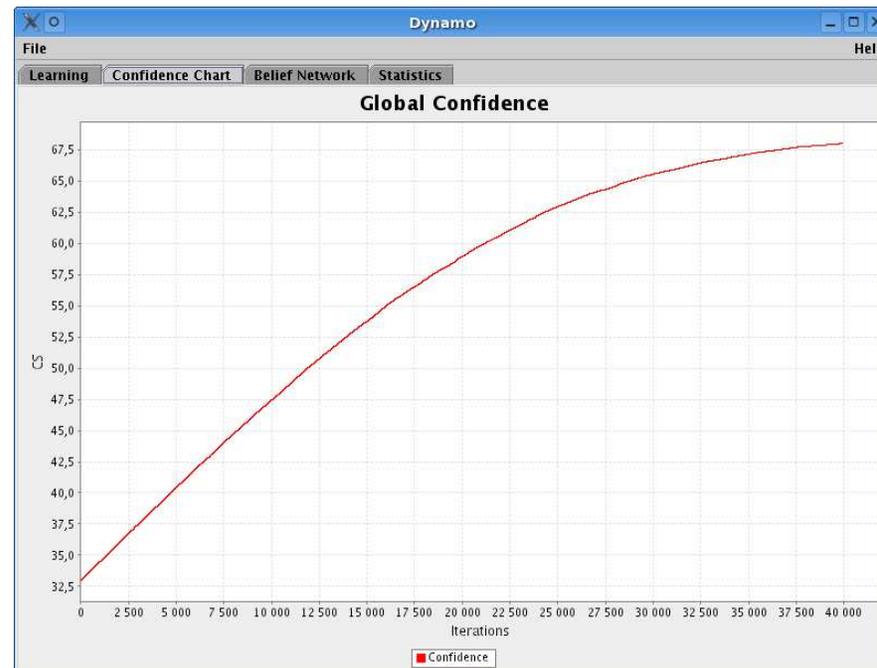
■ Analyse de frontières (deux versions)

■ Basée sur la loi de Zipf

	a	pattern
fréquent		rare
court		long

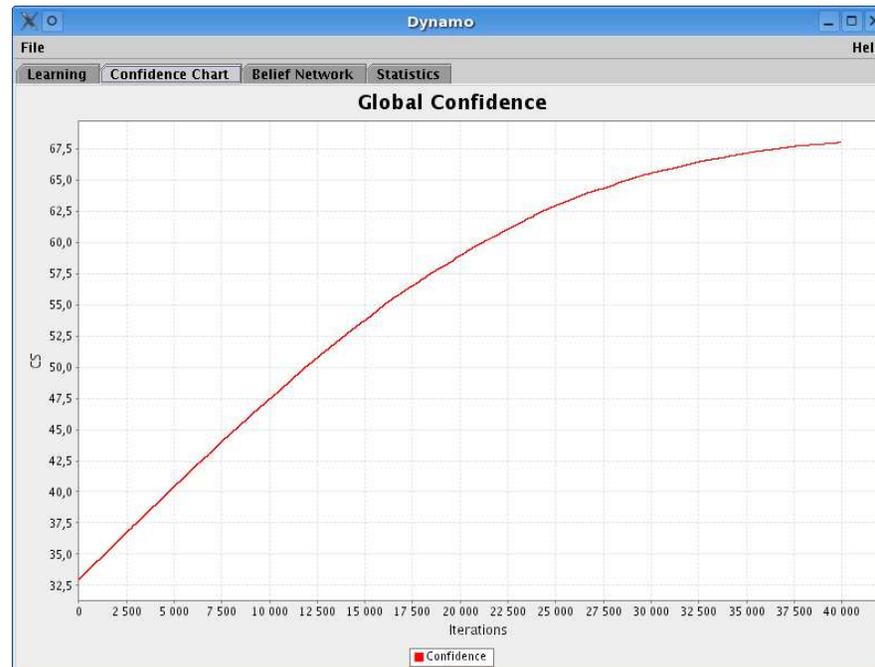
Expérimentations

- Système implémenté en Java
- Douze analyses en utilisant différents ensembles de compétences



Expérimentations

- Système implémenté en Java
- Douze analyses en utilisant différents ensembles de compétences
- Analyse des résultats
 - Qualitatif, examen d'un échantillon de mots
 - Quantitatif, par comparaison avec un étiqueteur (TreeTagger)



Evaluation

- Qualitatif
 - Système aveugle aux verbes
 - Bons résultats sur les noms en particulier avec l'analyse de frontières
 - La régulation de population aide pour les fonctionnels
 - Performance dégradée par l'analyse de frontières

Evaluation

■ Qualitatif

- Système aveugle aux verbes
- Bons résultats sur les noms en particulier avec l'analyse de frontières
- La régulation de population aide pour les fonctionnels
- Performance dégradée par l'analyse de frontières

■ Quantitatif

- Utilisation d'un tirage aléatoire
- Gain maximum par rapport à l'aléatoire environ 10%
- Le meilleur compromis sur le plan qualitatif a un gain de 9%

Bilan

- Résultats à peine meilleurs que ceux d'un tirage aléatoire
- Qualité trop faible pour construire des niveaux d'analyse supplémentaires

Bilan

- Résultats à peine meilleurs que ceux d'un tirage aléatoire
- Qualité trop faible pour construire des niveaux d'analyse supplémentaires
- Raisons de ce problème
 - Qualité des statistiques utilisées dans l'analyse contextuelle?
 - Pas de critère statistique suffisant ✓
 - Coopération intérieur du système discutable ✓
 - Agents à tendance altruiste
 - Difficulté d'évaluation de la coopération

Bilan

- Résultats à peine meilleurs que ceux d'un tirage aléatoire
- Qualité trop faible pour construire des niveaux d'analyse supplémentaires
- Raisons de ce problème
 - Qualité des statistiques utilisées dans l'analyse contextuelle?
 - Pas de critère statistique suffisant ✓
 - Coopération intérieur du système discutable ✓
 - Agents à tendance altruiste
 - Difficulté d'évaluation de la coopération
- Faut-il fonctionner à tout prix sans ressource ?
- Enjeu du système à un niveau supérieur...

Perspectives

- Recentrer les objectifs sur la construction d'ontologies

Perspectives

- Recentrer les objectifs sur la construction d'ontologies
- Adapter une grammaire de dépendance pour un AMAS
- Intégrer la notion de confiance aux grammaires de dépendances?

Perspectives

- Recentrer les objectifs sur la construction d'ontologies
- Adapter une grammaire de dépendance pour un AMAS
- Intégrer la notion de confiance aux grammaires de dépendances?
- Changer le grain des agents "mot" → "mot/catégorie" ou "lemme/catégorie"

Perspectives

- Recentrer les objectifs sur la construction d'ontologies
- Adapter une grammaire de dépendance pour un AMAS
- Intégrer la notion de confiance aux grammaires de dépendances?
- Changer le grain des agents "mot" → "mot/catégorie" ou "lemme/catégorie"
- Focaliser sur l'identification des structures porteuses de sens

Perspectives

- Recentrer les objectifs sur la construction d'ontologies
- Adapter une grammaire de dépendance pour un AMAS
- Intégrer la notion de confiance aux grammaires de dépendances?
- Changer le grain des agents "mot" → "mot/catégorie" ou "lemme/catégorie"
- Focaliser sur l'identification des structures porteuses de sens
- Tenir compte de l'interaction avec l'utilisateur

Perspectives

- Recentrer les objectifs sur la construction d'ontologies
- Adapter une grammaire de dépendance pour un AMAS
- Intégrer la notion de confiance aux grammaires de dépendances?
- Changer le grain des agents "mot" → "mot/catégorie" ou "lemme/catégorie"
- Focaliser sur l'identification des structures porteuses de sens
- Tenir compte de l'interaction avec l'utilisateur
- Travaux potentiels pour une thèse...







La coopération

- Attitude sociale des agents, ils cherchent donc à éliminer les situations non coopératives
- **NON COOPERATION** = $\neg C_1 \vee \neg C_2 \vee \neg C_3$
 - $\neg C_1$: Incompréhension
 - $\neg C_2$: Défaut de raisonnement
 - $\neg C_3$: Inutilité de l'action
- Système plongé dans un environnement dynamique
 - → situations imprévues
 - → SNC au niveau des agents
 - → réorganisation
 - → modification de f_S
 - → prise en charge de l'imprévu