

Dr. IMEN BOUDALI

Imen.boudali@gmail.com

Maître Assistante en Informatique

Institut Supérieur d'Informatique-Université Tunis-ElManar

Année Universitaire 2014-2015

Plan du Cours

- Introduction à l'IC
- Les SBC
- Représentation des connaissances
- Les ontologies
- Méthodologies d'IC : CommonKADS

Chapitre I:

Introduction à l'ingénierie des connaissances

Qu'est ce que l'IC ?

- Ingénierie ?
- Connaissance ?
- Ingénierie des Connaissances ?

Qu'est ce que l'IC ? Ingénierie ?

- Par Ingénierie on entend les activités de :
 - o Analyse : étude des besoins des utilisateurs,
 - Modélisation : formalisation des besoins
 - Simulation
 - Contrôle
 - Optimisation

mettant à contribution les diverses disciplines de l'ingénieur;

• Ce terme se substitue parfois au terme génie d'origine latin genius (produire, créer) désignant l'art de l'ingénieur.

Qu'est ce que l'IC ? Ingénierie ?

 Ingénierie : un <u>ensemble de techniques et de</u> <u>méthodes</u>, appliquées pour la résolution de problèmes complexes.

Qu'est ce que l'IC ? Ingénierie ?

- **ingénieur** (*ingenium*) : personne exerçant des activités de:
 - conception
 - direction de projets
 - réalisation
 - mise en œuvre de produits, de systèmes ou de services impliquant des problèmes techniques complexes;
- => Suppose un ensemble de connaissances techniques et économiques, sociales et humaines, reposant sur une solide culture scientifique.

Qu'est ce que l'IC ? Connaissance ?

- Connaissance / knowledge ?
- Information / Information?
- Donnée / data ?

Différences?

• Donnée est :

- le résultat direct d'une mesure.
- basée sur des symboles (codage) : représentation abstraite
- un élément brut, qui n'a pas encore été interprétée, ni mis en contexte.
- Un couple formée par un concept et une mesure

Elle peut être :

- Quantitative : poids, taille, montant, âge
- Qualitative : nom, prénom, ville, pays

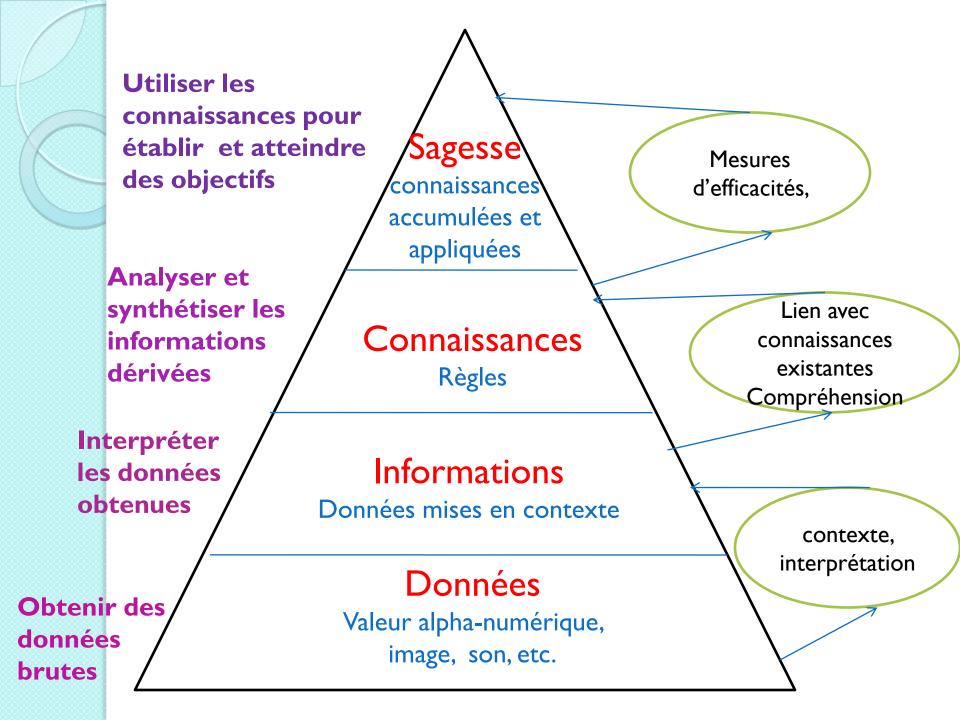
=> Une donnée seule ne permet pas de prendre une décision sur une action à lancer.

Information est:

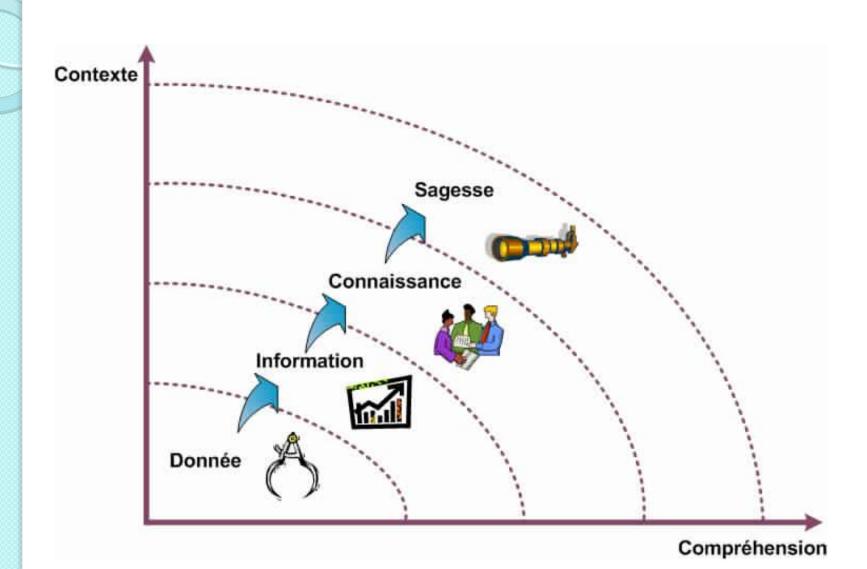
- une donnée interprétée
- Information = Donnée + contexte
- => Une information permet à un responsable opérationnel de prendre une décision locale

Connaissance est:

- Connaissance = information + mode d'emploi dans un contexte donné (sémantique)
- le résultat d'une réflexion sur les informations analysées en se basant sur l'expertise du domaine
- De l'information interprétée et mise en lien avec les connaissances existantes
- Une information validée et retenue comme une règle
- Règle utilisant des données pour en produire d'autres



- Sagesse: dernière étape qui permet de :
 - aboutir à un état d'esprit général d'évaluation final sur le contenu (informations, connaissances) et de jugement
 - s'adapter à de nouvelles situations
 - anticiper les grands changements à venir et l'évolution d'un domaine



Données/informations/Connaissances Exemple :

- Donnée: Il fait 15° dans cette pièce
 - ⇒ valeur brute indiquée par un thermomètre
- Information: Il fait froid dans cette pièce.
 - ⇒ donnée contextualisée
- Connaissance: Pour avoir chaud, il suffit de monter le chauffage
 - ⇒ Interprétation des informations

Ingénierie des connaissances

• L'IC évoquerait les techniques pour manipuler des connaissances sur ordinateur

• L'IC est synonyme de Génie cognifif

- L'Intégration des techniques d'IA et du GL en vue de concevoir et de construire des systèmes experts.
- Discipline étudiant l'extraction et la formalisation de connaissances provenant d'un expert humain en vue de leur intégration dans des systèmes experts.

Ingénierie des connaissances

• Finalité applicative

- o aider un opérateur dans sa tâche (ou un groupe)
- transmettre, enregistrer des savoir-faire, gérer des connaissances
- améliorer les performances, l'adaptabilité ou les résultats de systèmes artificiels

Interdisciplinarité

- informatique IA
- psychologie cognitive et ergonomie
- Linguistique

Ingénierie des connaissances

- L'Ingénierie des connaissances intervient pour
 - Définir une aide à l'utilisateur (méthodes, outils logiciels ou non, organisation du travail)
 - Modéliser des connaissances, individuelles ou collectives, explicites ou implicites, stabilisées ou évolutives, expertes ou techniques ...
 - Rendre ces connaissances accessibles sous une forme définie en fonction du contexte, opérationnelle ou non
- La recherche en Ingénierie des connaissances produit
 - Des méthodes et des techniques de recueil, d'analyse et de structuration des connaissances
 - Des plates-formes de modélisation
 - Des représentations des connaissances opérationnelles ou non
 - ⇒ L'Ingénierie des connaissances se veut une ingénierie

Les principaux langages d'IA

- Lisp (1960, J. MacCarthy)
- Prolog (1973, A. Colmerauer), Prolog avec contraintes
- SmallTalk (1972, A. Kay)
- JAVA (1994), C++, Scheme, ...

Les grands domaines de l'IA

- Reconnaissance et synthèse de la parole
- Reconnaissance et synthèse d'images (Reconnaissance des visages)
- Reconnaissance de l'écriture (ex: recon. cheques, codes postaux)
- Systèmes experts : MYCIN (diagnostic médical)
- Representation des connaissances

Les grands domaines de l'IA

- Traitement du langage naturel
- Résolution de problèmes combinatoires (ordonnancement, planification)
- Robotique
- Apprentissage
- Jeux (Echecs(DeeperBlue à 2600), Checkers, Othello BackGammon, GO.

Les grands domaines de l'IA

Simulation du raisonnement humain: logique floue

 Médecine: Aide à la décision (SE), prédiction de patients à risques, analyse automatique d'images médicales

Retombées de l'IA

- objets, agents, méthodologies, représentation des connaissances
- fouille de données, fouille de texte
- programmation par contraintes
- nouvelles méthodes d'optimisation

Plan du Cours

- Introduction à l'IC
- Les SBC
- Représentation des connaissances
- Les ontologies
- Méthodologies d'IC : CommonKADS

Chapitre II:

Les Systèmes à Base de Connaissances (SBC)

Plan

- Définition & Objectifs
- Architecture des SBC
- Evolution des SBC
- Moteurs d'inférences : Principes & types

Plan

- Définition & Objectifs
- Architecture des SBC
- Evolution des SBC
- Moteurs d'inférence : Principes & types

Définition

- Systèmes informatiques fonctionnant avec une base de connaissances sur un sujet donné.
- Des systèmes qui utilisent des techniques de l'intelligence artificielle (IA) pour effectuer des tâches réservées généralement aux humains

• Ces systèmes sont appliqués à divers domaines (informatique, médecine, transport, etc.).

Objectifs des SBC

- Inscrire (ou représenter) les connaissances en tant que connaissance dans un système :
 - ⇒ Pour « conserver » des savoirs, des savoir-faire et leur sémantique associée
- Disposer d'un « moteur » permettant d'enchaîner des inférences sur ces inscriptions de connaissances :
 - ⇒ Pour « exploiter » les savoirs et savoir-faire ainsi « conservés »

Rôles des SBC

 Représenter les connaissances issues de l'expertise ou/et de la pratique dans un système informatique

 Fonde le « raisonnement » sur des mécanismes d'inférence logique ou analogique

Intègre une représentation symbolique

Rôles des SBC

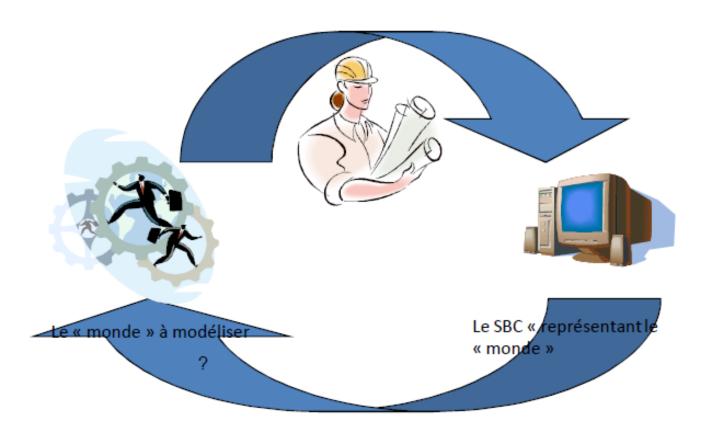
 Autorise parfois une certaine prise en compte de l'incertitude

 Se fonde sur des heuristiques (connaissances spécifiques au domaine qui guident la recherche de solutions)

 Est orienté décision, résolution de problème et peut fournir des explications

Rôle de l'Ingénierie des connaissances

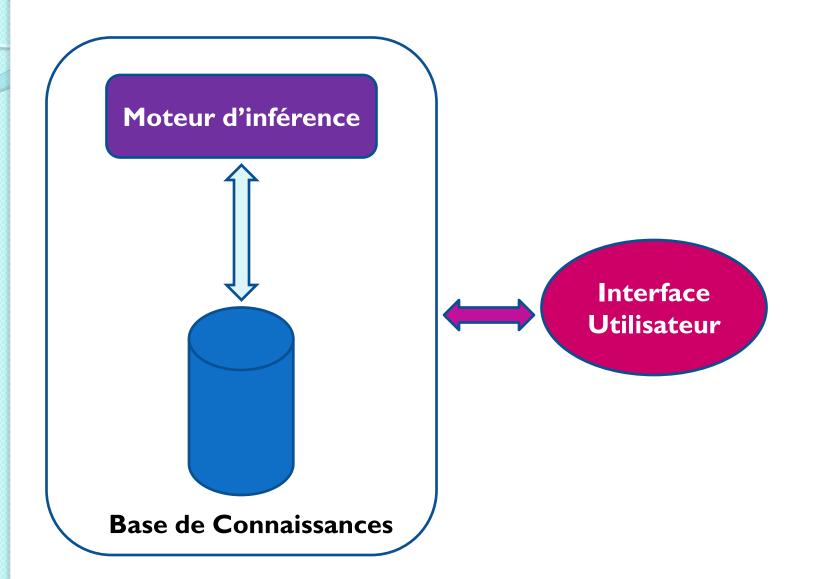
Objectif historique : faciliter la construction des Systèmes à Base de Connaissances



Plan

- Définition & Objectifs
- Architecture des SBC
- Evolution des SBC
- Moteurs d'inférence : Principes & types

Architecture d'un SBC



Architecture d'un SBC

 BC : contenant les connaissances nécessaires aux tâches affectées au système

- Moteur d'inférence : un programme qui exploite ces connaissances pour résoudre la tâche:
 - Il détermine les règles utilisables,
 - o les applique pour produire de nouveaux faits, et
 - répète ce cycle jusqu'à l'obtention de faits intéressants

Architecture d'un SBC

- Interfaces utilisateurs du système, dédié :
 - d'une part au concepteur, qui doit :
 - pouvoir modifier la base de connaissances et
 - paramétrer le moteur d'inférences de façon à améliorer les performances du système,
 - d'autre part la personne qui :
 - · soumet un problème à résoudre et
 - attend, non seulement la réponse, mais aussi des explications sur la manière dont cette réponse a été obtenue.

Remarques:

 Les connaissances formalisées dans la base sont sujettes à des critiques et des révisions;

 Les réponses obtenues doivent donc pouvoir être évaluées en fonction de leur pertinence, telle que jugée par la personne qui a soumis le problème à résoudre.

Plan

- Définition & Objectifs
- Architecture des SBC
- Evolution des SBC
- Moteurs d'inférence : Principes & types

Evolution des SBC

2 générations de SBC

- Les SBC Ière Génération : parues à partir des années 70
- Les SBC 2^{ème} Génération : parues les dix dernières années

Evolution des SBC

SBC lère Génération: (ou SE classiques)

 Ce sont des logiciels capables de reproduire le raisonnement d'un expert dans un domaine de compétence spécialisé

 Les premiers exemples de SE conçus pour l'aide au diagnostic, qu'il soit médical ou technique.

Justification de ces projets :
 les raisonnements de spécialistes seraient plus faciles à formaliser et à reproduire que des raisonnements de sens commun (la compréhension d'un récit ou d'un énoncé quelconque).

- La conception d'un tel système passe par:
 - 1. le recueil des connaissances utilisées par l'expert et
 - 2. Par la compréhension de ses schémas de raisonnement

 les connaissances étaient représentées le plus souvent de façon uniforme par des règles de la forme :

« Si Condition(s) alors Action(s) »

ce qui permettait d'exprimer les connaissances sous forme déclarative.

© Cette approche accélérait le développement des SBC.

Exemple:

- Lors d'un diagnostic médical, il s'agit de:
 - identifier et de formaliser les connaissances, issues de l'expérience et de la formation de l'expert,
 - Exploiter correctement ces connaissances pour passer des symptômes observés et des données sur le patient à un ou plusieurs diagnostics argumentés.
 - ⇒ MYCIN système expert de diagnostic médical (maladies bactériennes du sang)

- DENDRAL : ler système expert permettant d'identifier les constituants chimiques d'un matériau.
- SACHEM : Système d'Aide à la Conduite des Hauts fourneaux En Marche
 - conçu pour piloter des hauts- fourneaux en analysant les données fournies en temps réel par un millier de capteurs.

- Ces connaissances sont difficiles à expliciter (à codifier)
- Le développement d'un système expert est alors hautement itératif

- Les connaissances qu'il utilise doivent faire l'objet de fréquentes mises à jour.
- Pour faciliter leurs modifications, il est jugé préférable de ne pas les « figer » dans un programme, mais de les décrire séparément, dans ce qui est appelé une « base de connaissances ».

 Le système expert doit être alors réalisé sous la forme d'un « système à base de connaissances ».

• Elle a conduit à des systèmes ayant un faible niveau d'abstraction des connaissances.

- les SE lère Génération étaient difficiles à modifier, peu réutilisables et transférables d'un environnement à l'autre.

En pratique, le développement de ce genre d'application est très lourd car en dépassant la centaine de règles, il devient difficile de :

- comprendre comment le système expert
 « raisonne » : manipule faits et règles en temps
 réel, et donc
- d'en assurer la mise au point finale puis la maintenance.

Evolution des SBC

SBC 2ème Génération:

 Résultats des travaux menés pour pallier aux difficultés des SBCIG

 Ces travaux ont permis de poser les bases d'une ingénierie des connaissances basée sur le paradigme de modélisation.

• repose sur le principe d'une séparation des connaissances et de leur mise en œuvre

SBC 2ème Génération:

- Avec les SBC2G, la distinction est faite entre :
 - le type des connaissances, et
 - les modes de représentation utilisés pour les manipuler, ce qui offre un meilleur niveau d'abstraction

 les connaissances du système sont explicitement représentées (avec précision), de façon à être facilement compréhensibles et modifiables.

SBC 2ème Génération:

- Plusieurs représentations de connaissances ont été proposées:
 - les règles, plus ou moins complexes (avec ou sans variables),
 - les formalismes dits « à objets » ont rencontré un certain succès.

 Cette base de connaissances peut être modifiée sans que le programme qui les exploite n'en soit affecté.

SBC 2ème Génération:

- séparation censée apporter des avantages décisifs quant à :
- l'évolutivité des connaissances et,
- l'explicabilité des démarches de résolution

Plan

- Définition & Objectifs
- Architecture des SBC
- Evolution des SBC
- Moteurs d'inférence : Principes & types

Moteurs d'inférence

 chargé d'exploiter les connaissances acquises dans la base pour résoudre un problème donné.

 Les connaissances manipulées sont généralement sous forme de règles de production

« Si Condition(s) alors Action(s) »

Moteurs d'inférence

• Capacité d'explication d'un raisonnement grâce à l'examen de l'enchaînement des règles utilisées: traçabilité

Principe de fonctionnement des MI

 Le MI enchaîne des cycles de travail comportant chacun 3 phases :

- Phase d'évaluation
- 2. Phase de résolution de conflits
- 3. Phase d'exécution

Phase d'évaluation

- Sélection
 - d'un sous ensemble FI de la bases des faits (BF)
 - d'un sous ensembles R1 de la bases des règles (BR).

⇒ Les Méta-règles

Phase d'évaluation

• Mise en correspondance ou pattern Matching

 Le MI compare la partie Condition de chacune des règles de RI par rapport à FI.

 Sélectionne un sous ensemble R2 de règles de R1 jugées compatibles avec F1.

⇒ R2 est appelé ensemble de conflit

Phase d'évaluation



- Filtrage : MI d'ordre 0
- Unification: MI d'ordre I

Phase de résolution de conflits

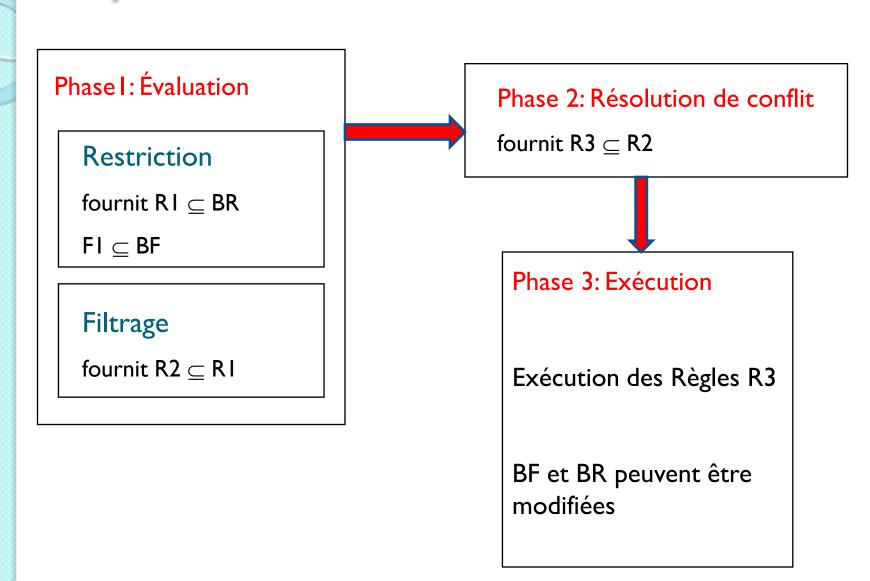
- Le MI choisit les règles qui doivent être effectivement déclenchées.
- ⇒ un sous ensemble R3 de règles de R2

Phase d'exécution

 Le MI commande la mise en œuvre des actions définies par les règles de R3 (si R3 non vide).

 Quand R3 compte plus d'une règle, l'exécution peut être différente d'un moteur à un autre

Cycle de base d'un MI



Types des Moteurs d'Inférence

• Les types des MI diffèrent selon :

- Le régime de contrôle
- Le mode de chaînage

Régime de contrôle

- Régime irrévocable
- Régime par tentatives

Régime de contrôle

Différence?

Quand R3 = \emptyset

- certains moteurs s'arrêtent
 - ⇒ Ces moteurs ont un régime de contrôle irrévocable.
- D'autres moteurs reconsidèrent l'ensemble de conflit R2 d'un cycle antérieur en examinant la possibilité de déclencher d'autres règles de R2.
 - ⇒ Ces moteurs ont un régime de contrôle par tentatives :il y a remplacement de déclenchement de règles par d'autres.

Régime de contrôle

Remarque

Quant un moteur revient sur une résolution de conflits antérieure on dit qu'il opère un retour arrière ou backtracking

Mode de chaînage

Chaînage avant

- Partir des faits établis pour en déduire d'autres
- utilisé quand on cherche les conséquences de l'ajout de nouveaux faits.

Chaînage arrière

- Partir des buts à prouver pour atteindre les faits établis
- utilisé quand on cherche à prouver un but.

Chaînage mixte

chaînage bidirectionnel

Mode de chaînage

 Chaînage mixte : nécessite la mise en place d'heuristiques particulières capables de décider, face à une situation donnée, quel est la stratégie la plus adapté pour poursuivre la résolution.

Chaînage Avant

- Schéma Avant I: Intégration immédiate des conclusions de chaque règle
- Schéma Avant 2: production et intégration des faits en largeur d'abord

Schéma Avant I:

Intégration immédiate des conclusions de chaque règle

Principe

- I. Parcourir la BR,
- 2. Trouver une règle applicable (dont les faits de la partie prémisse appartiennent à la BF (faits établis)) alors la conclusion est un fait établi,
- 3. Ajouter le nouveau fait à la BF immédiatement.
- 4. Recommence le parcours depuis le début de la BR

Exemple de BC (modélisé avec la logique de prédicats d'ordre 0)

BR:

 $RI: K, L, M \rightarrow I$

R2: I,L,J \rightarrow Q

R3: C,D,E \rightarrow B

 $R4:A,B \rightarrow Q$

R5: L,N,O,P \rightarrow Q

R6: $C,H \rightarrow R$

R7: R,J,M \rightarrow S

R8: $F,H \rightarrow B$

R9: $G \rightarrow F$

BF: A,C,D,E,G,H,K

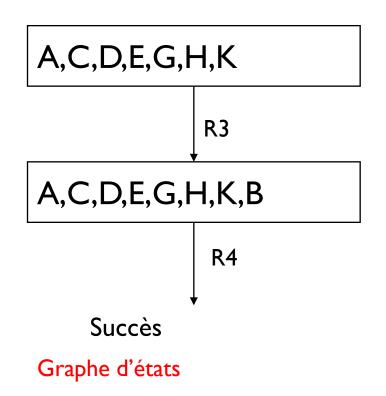
Prouver Q?

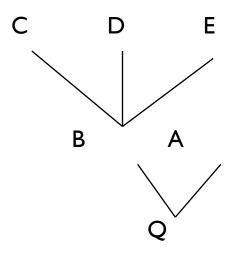
Remarque On peut utiliser:

 Une représentation par espace d'états: les règles sont considérées comme des opérateurs de changement d'états de la base des faits

 Une représentation par graphe de sousproblèmes: les règles sont considérées comme des opérateurs de décomposition

Schéma Avant I Intégration immédiate des conclusions de chaque règle





Graphe ET/OU

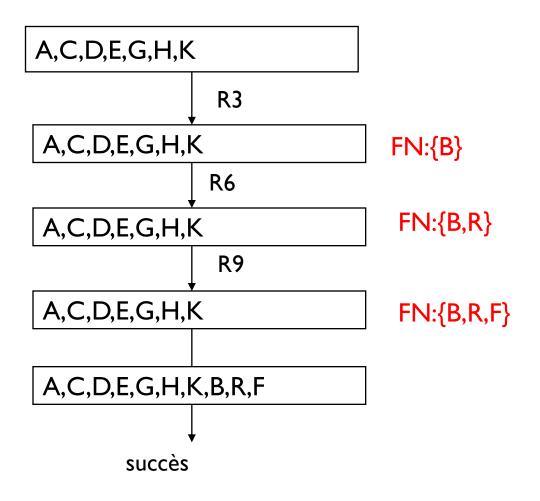
Schéma Avant 2

production et intégration des faits en largeur d'abord

Principe

- 1. Déclencher l'une après l'autre toutes les règles compatibles avec la BF avant d'ajouter leurs conclusions à la BF et de les utiliser pour déclencher de nouvelles règles.
- 2. Les nouvelles conclusions de toutes les règles sont rassemblées dans Faits Neufs. Lorsque toutes les règles ont été exploitées, FN est vidé dans la BF.

Schéma Avant 2 production et intégration des faits en largeur d'abord



Conclusion

- ❖ Schéma Avant 1 et Schéma Avant 2 fonctionnent selon un régime irrévocable puisque le déclenchement d'une règle n'est jamais remplacé par un autre.
- Ils ne permettent que des ajouts de faits établis : aucun fait ne peut être retiré: on dit que ces moteurs sont monotones

Chaînage Arrière

- Schéma Arrière I: production de sous-problèmes en profondeur d'abord
- Schéma Arrière 2: profondeur d'abord sauf si une règle est immédiatement concluante

Schéma Arrière I production de sous-problèmes en profondeur d'abord

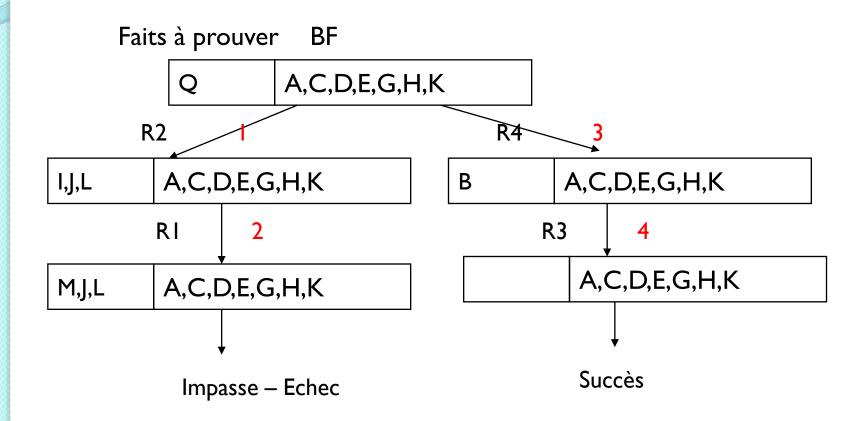
Principe

- invoque les règles en chaînage arrière càd qu'il compare la partie conclusion des règles aux faits à établir.
- maintient à chaque instant un ensemble de fait à établir.

Exemple

A l'état initial l'ensemble des faits à établir ne contient que Q. si la règle R2 est invoquée, le fait Q à établir sera remplacé par les faits à établir I,J,L.

Schéma Arrière I production de sous-problèmes en profondeur d'abord



Remarque

- Schéma Arrière I s'arrête soit parce que :
 - tous les problèmes initiaux ont été reconnus établis (liste des faits à prouver devient vide)
 - → succès
 - I'un des faits à prouver ne peut être établi
 - → Echec

Schéma Arrière 2 profondeur d'abord sauf si une règle est immédiatement concluante

Principe

- Disposant de l'ensemble des règles qui concluent sur le fait à prouver on préfère ici examiner d'abord si l'une de ces règles est immédiatement concluante, càd si les faits de sa prémisse sont déjà dans la BF.
- Cette stratégie peut être vue comme un premier pas depuis les stratégies en profondeur vers les stratégies en largeur

Remarque

Schéma arrière 1 et arrière 2 repose sur:

 une Invocation des règles en chainage arrière avec régime par tentative et monotonie

Chaînage Mixte

 Invocation des règles en chaînage mixte avec régime par tentative et non monotonie

 Dans ce type de moteur il s'agit de déterminer une séquence ou plan d'actions à exécuter à partir d'un état de départ pour résoudre un problème appelé objectif