III.2. SAVANT 3 : enseigner par la conversation

1. Introduction

Nous décrivons ici le système SAVANT3 que nous avons développé à TELECOM-Paris. Dans son principe, ce système utilise certains des mécanismes argumentatifs que nous avons pu analyser en observant des conversations spontanées, afin de tester la cohérence des connaissances de l'étudiant. Il se caractérise par une grande économie de mise en oeuvre : une quinzaine de relations logiques suffisent à SAVANT3 pour soutenir une conversation sur un sujet donné. Le caractère naturel de l'interaction que ce système tente de reproduire permet d'espérer qu'il sera bien accepté par les étudiants.

2. L'importance pédagogique d'un contexte argumentatif

2.1. L'apprentissage au cours de la conversation

Nous avons mis en évidence dans le chapitre III.1 le fait que la plupart des systèmes dont l'objectif était l'enseignement de connaissances conceptuelles utilisaient le dialogue dit "Socratique" comme mode d'interaction avec l'élève. La justification de ce type d'approche utilise généralement comme situation modèle la relation professeur élève, comme on peut l'observer par exemple dans des cours particuliers [Collins 1976]. Ce genre de justification a été critiqué [Petrie-Brown 1989]. Pour notre part, nous cherchons à ancrer l'intérêt didactique du type de dialogue que nous essayons d'instaurer entre l'élève et la machine, non dans la relation institutionnelle enseignant-enseigné, mais dans sa contrepartie naturelle, la conversation spontanée.

Nous avons laissé entendre dans les chapitres II.3 et II.4 que la perception d'une situation incohérente pouvait jouer le rôle de déclencheur pour l'apprentissage d'un nouveau concept ou pour la modification de relations conceptuelles. Prenons un exemple conversationnel particulièrement clair de ce point de vue :

[ex_fond]
contexte : A s'étonne d'améliorer certaines de ses performances sportives alors qu'il atteint la trentaine.
A1- Moi, avec l'âge, je peux tenir des trucs en fond, mais dès que c'est le sprint... *B1- Normal.*
A2- Je tiens mieux le fond qu'avant. C'est dingue, ça! *B2- Normal, normal. Mais ça c'est normal, hein. On est plus endurant que résistant. Avec l'âge.*
C1- Quelle est la différence? *[...]*

L'incohérence est clairement exprimée en A2. Sa solution vient de l'introduction des concepts d'endurance et de résistance: l'endurance culmine plus tard que la résistance. La situation de A est ainsi présentée par B en B2 comme un cas particulier d'une loi plus générale que A et C ne semblaient pas soupçonner.

Nous pouvons considérer qu'un apprentissage conceptuel, comme ici la distinction endurance / résistance et son lien avec l'âge, intervient comme l'*explication* d'une situation perçue comme incohérente (voir chap. III.4)[1]. Mais pourquoi chercher à reproduire une situation de type conversationnel entre l'étudiant et la machine ? Après tout, on pourrait imaginer que le fait de donner de simples définitions des concepts pourrait avoir une efficacité équivalente !

La réponse à cette question est essentielle pour notre propos. Nous retenons comme critère d'efficacité pédagogique, en ce qui concerne l'apprentissage des connaissances conceptuelles, les deux aspects suivants :

⇨ ***compréhension***

⇨ ***mémorisation***

2.2. Comprendre un nouveau concept grâce à un contexte problématique

Que veut dire précisément *comprendre* ? En prolongement du chapitre II.3, nous suggérons le fait qu'un concept sera compris dès que les relations logiques qui le relient à d'autres concepts seront correctement connues. Cette définition nous permet de faire les remarques suivantes :

- un concept compris permet de s'étonner à bon escient, et réciproquement
- un concept compris permet d'expliquer les situations qui le mettent en jeu, et réciproquement
- la compréhension d'un concept n'est pas un phénomène graduel[2] (cf. chap. III.1).

En effet, si le concept *c* entre dans des relations logiques avec d'autres concepts (nous représentons ces relations par des clauses négatives, dites aussi clauses paradoxales) :

$[\, c \,\&\, c_1 \,\&\, c_2 \ldots \&\, c_m \,] \Rightarrow \mathbf{F}$

$[\, \mathbf{non}\ c \,\&\, c'_1 \,\&\, c'_2 \ldots \&\, c'_n \,] \Rightarrow \mathbf{F}$

alors ces relations sont autant d'occasions d'étonnement. De plus, ***non*** *c* et *c* sont respectivement des explications (invalidations indirectes, comme négations de prémisses omises, cf. chap. III.4) relatives aux étonnements suivants, si un interlocuteur les manifeste :

$[\, c_1 \,\&\, c_2 \ldots \&\, c_m \,] \Rightarrow \mathbf{F}$

$[\, c'_1 \,\&\, c'_2 \ldots \&\, c'_n \,] \Rightarrow \mathbf{F}$

Compte tenu de ces remarques, on comprend l'importance du contexte argumentatif pour la compréhension. Toute relation logique qui lie un nouveau concept *c* à des concepts déjà connus $\{c_i\}$ peut donner lieu à un étonnement. Supposons qu'un étonnement ait effectivement lieu, et appelons *clause de référence* (CR) la relation logique qui traduit cet étonnement. Un contexte argumentatif est l'ensemble formé par la CR et les clauses qui permettent de relier les autres concepts $\{c_i\}$ de la CR à des concepts tenus pour vrais dans la situation (comme expliqué dans le chapitre I.2).

Par exemple, le contexte argumentatif de l'extrait précédent peut être représenté ainsi :

---

[1] L'invalidation d'une situation improbable ou d'une situation (in)désirable peut donner lieu, de même, à un apprentissage conceptuel. Nous n'en parlerons pas, car le mécanisme est exactement le même que pour l'incohérence (paradoxe).

[2] Nous nous situons là dans le cas où le contexte est spécifié. Il serait certes possible de voir une gradualité dans le fait qu'entre tous les contextes imaginables, une proportion variable des relations logiques incluant le nouveau concept est connue.

$\forall$ Personne , $\forall$ Sport [plus-âgé(Personne) **&** meilleur(Personne, Sport)] $\Rightarrow$ **F**

$\exists$ *A* , $\exists$ *nage-de-fond* plus-âgé(A) **&** meilleur(*A*, *nage-de-fond*)

Le concept *endurance* vient expliquer la situation, puisqu'il intervient sous forme négative dans une version plus complète du contexte :

$\forall$ Personne , $\forall$ Sport [plus-âgé(Personne) **& non** endurance(Sport)
& meilleur(Personne, Sport)] $\Rightarrow$ **F**

La présence d'un contexte argumentatif assure que le nouveau concept sera *compris*, dans la mesure où il permet à l'apprenant d'expliquer une situation. Cette explication insère le nouveau concept dans un réseau de relations avec les concepts déjà connus. Comprendre un concept, c'est donc accepter des relations conceptuelles (que l'on peut représenter comme des relations d'incompatibilité) qui lient ce concept aux concepts connus. Mais qu'en est-il de la mémorisation ?

2.3. Mémoriser un nouveau concept grâce à un contexte problématique

Cette question de la mémorisation de connaissances conceptuelles a été abordée d'une manière tout à fait intéressante pour notre propos par John Bransford et Barry Stein [1982]. Ces auteurs ont réalisé l'expérience suivante. Ils présentèrent à des étudiants des phrases comme celles-ci :

1ère série

- The tall man bought the crackers *(l'homme grand a acheté les biscuits)*
- The bald man read the newspaper *(l'homme chauve lisait le journal)*
- The funny man liked the ring *(l'homme amusant a bien aimé la bague)*
- The hungry man purchased the tie *(l'homme affamé a acheté la cravate)*
- The short man used the broom *(le petit homme a utilisé le balai)*
- The strong man skimmed the book *(l'homme musclé feuilletait le livre)*

Ces phrases sont construites de telle manière que la relation entre la caractéristique de l'homme qui fait l'action et l'action elle-même soit arbitraire. On demande ensuite aux étudiants de ranger ces phrases dans l'ordre de leur facilité de compréhension. Mais le but du test n'est pas là. Sans qu'ils soient prévenus, les étudiants subissent un test de mémorisation : "Quel est l'homme qui acheté la cravate ?", "Quel est celui qui aimait la bague ?", etc. Le résultat obtenu par les étudiants fut médiocre : 4,2 phrases sur 10 en moyenne.

Il semble qu'un *lien* entre les éléments d'une phrase à mémoriser puisse favoriser la mémorisation. Mais quelle est la nature de ce lien ? Pour le savoir, Bransford et Stein ont présenté à un autre groupe d'étudiants les phrases suivantes :

2ème série

- The tall man bought the crackers that had been lying on the top shelf
- The bald man read the newspaper in order to look for a hat sale
- The funny man liked the ring that squirted water
- The hungry man purchased the tie so that he could get into the fancy restaurant
- The short man used the broom to operate the light switch
- The strong man skimmed the book about weight lifting

Cette fois le taux de mémorisation atteint 7,4 / 10. Pour les auteurs, l'interprétation ne faisait aucun doute :

> "[...] memory was enhanced because the elaborations helped people understand the significance or relevance of linking a particular type of man to a particular activity." [Bransford & Stein 1982]

Pourtant d'autres interprétations ont été émises, notamment le fait qu'une élaboration reliée sémantiquement aux éléments de la phrase aidait à la mémorisation. Autrement dit, le fait d'ajouter des éléments du même contexte *sémantique* va permettre aux étudiants de se souvenir des phrases. Bransford et Stein ont voulu tester cette idée en essayant des phrases comme celles-ci :

3ème série

- The tall man bought the crackers from the clerk in the store
- The bald man read the newspaper while eating breakfast
- The funny man liked the ring that he received as a present
- The hungry man purchased the tie that was on sale
- The short man used the broom to sweep the porch
- The strong man skimmed the book before going to sleep

> "These statements included elaborations that make sense semantically, but the elaborations did not help understand why it might be relevant for a particular type of man to perform a particular activity. The results indicated that students who received sentences such as these exhibited poorer memory than students who received the first list of base sentences (the list without any additional elaborations). It seems clear, therefore, that the semantic congruity of elaborations does not necessarily lead to increase retention." [Bransford & Stein 1982]

Une simple élaboration sémantique ne favorise donc pas la mémorisation. Il faut davantage. Il faut un lien "pertinent". Notre propre interprétation de cette expérience consistera simplement à préciser ce terme de pertinence (*relevance*) employé par les auteurs pour désigner les "élaborations" qui facilitent la mémorisation. Le rôle de l'information additionnelle, dans la série de phrases bien mémorisées (2ème série), est de faciliter l'insertion de l'action de chacun des personnages dans un contexte "problématique". Voici des exemples de contextes qui peuvent venir spontanément à l'esprit de celui qui lit la deuxième série de phrases :

- Nobody could see these crackers, but fortunately this man did.
- The bald man needs a hat, otherwise he may have a cold
- Some people may take offence at having their face splashed, not this funny man
- The hungry man was refused the entrance into the restaurant
- The short man could not reach the light switch
- Nobody can be interested in such a book, but a strong man

Nous pensons que la deuxième série de phrases a été bien mémorisée parce que les sujets ont imaginés ces contextes problématiques[3]. Pour la 3ème série, les sujets ont pu imaginer une scène pour chaque phrase, mais l'établissement d'un contexte problématique est, en revanche, beaucoup moins immédiat[4].

L'important est de constater, et cette expérience nous invite à le faire, qu'un *lien logique* entre les éléments d'une phrase facilite la mémorisation. Bransford et Stein suggèrent eux-mêmes les conséquences pédagogiques d'un tel constat :

> "[...] even biology novices who know the general terms *vein* and *artery* may have difficulty remembering which ones have particular properties because the relationships between each concept (vein, artery) and its properties (*e.g.* elastic, thick) seem arbitrary." [...] "biology experts who are presented with long lists of simple facts (*e.g.* arteries are elastic) versus elaborated facts (*e.g.* arteries are elastic, thick and have no valves) should produce [...] better free recall for the concept (*e.g.* artery) when it is more elaborated [...] For biological novices, however, the relationships between concepts and properties are arbitrary [...] Under these conditions, semantically congruous elaborations may hurt retention rather than help." [Bransford & Stein 1982]

---

[3] Ces exemples jouent sur la modalité désirable / indésirable, mais le même effet serait obtenu avec les modalités paradoxale ou improbable (cf. note **Erreur! Signet non défini.** p. **Erreur! Signet non défini.**).

[4] Dans les deux séries de phrases "élaborées" présentées par Bransford et Stein, la dernière phrase est sans doute la moins intéressante. Le lien entre l'homme fort et le livre d'haltérophilie est surtout associatif, alors que le fait de feuilleter le livre peut être facilement lié logiquement à l'envie de dormir.

Nous allons adopter le paradigme suivant :

Un contexte argumentatif, qui comporte une situation problématique, facilite la compréhension et la mémorisation de nouvelles connaissances.

Notre objectif, avec le développement de SAVANT3, est de nous assurer que tout nouveau concept présenté aux étudiants sera inséré dans un contexte argumentatif. Les étudiants peuvent alors plus facilement relier logiquement le nouveau concept à d'autres concepts déjà connus. Le moyen par lequel ces liens logiques seront rendus manifestes, dans notre application, consiste dans la présentation de situations potentiellement incohérentes.

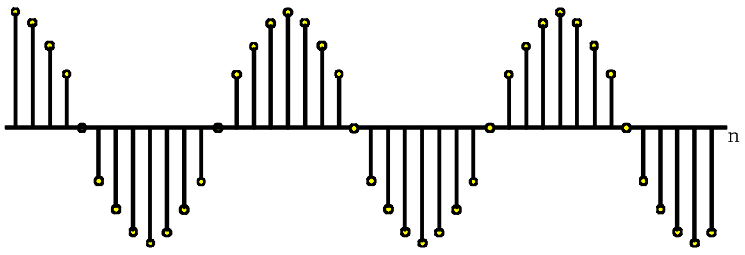
3. Principe de fonctionnement de SAVANT3

SAVANT3 repose sur un principe inspiré de l'observation des conversations : présenter à l'étudiant un contexte argumentatif organisé autour d'une contradiction. Nous venons d'indiquer l'importance de ce type de contexte pour la compréhension et la mémorisation des connaissances. Dans le chapitre III.1, nous avons rappelé qu'un principe similaire était à l'oeuvre dans la stratégie de piège du célèbre système WHY. Nous retrouvons une stratégie similaire dans le système NF-Tutor de Thomas Rätz [1992]. T. Rätz note que les explications fournies dans le cas où une contradiction est détectée sont pédagogiquement plus efficaces, et qu'elles peuvent être calculées (par opposition à une situation où les explications seraient issues de "canevas" (*pattern*) prévus à l'avance). Toutefois, son système repose sur l'existence d'une liste de conceptions erronées que l'auteur est obligé de fournir explicitement. L'explication par la mise en évidence d'une contradiction est déclenchée lorsqu'une erreur commise par l'étudiant peut être rapprochée d'une de ces conceptions erronées.

Dans notre système, les conceptions erronées ne sont pas explicites au départ. C'est l'occurrence d'une contradiction qui est détectée en premier. Cette contradiction va provoquer un étonnement de la machine, puis conduire à un dialogue. C'est ce dialogue qui va permettre à l'étudiant de corriger la conception erronée si celle-ci existe.

Au début d'une interaction avec SAVANT3, la situation n'est généralement pas incohérente. Le système va donc poser des questions ou émettre des répliques qui visent en fait à amener l'étudiant à préciser (instancier) la situation jusqu'à ce que le système puisse établir une incohérence.

Nous pouvons observer cette stratégie dans l'extrait suivant, qui porte sur la périodicité d'une sinusoïde discrète $cos[2\pi F_1 n]$. Ce signal est périodique si et seulement si F_1 est rationnelle.



Cette propriété mathématique peut donner lieu à un dialogue comme celui-ci entre la machine (M) et l'étudiant (E) (les réponses entre crochets indiquent un choix fermé) :

M1: Le signal cos[2πF$_1$ n] (pour F$_1$ donnée) est-il périodique ?
E1: oui
M2: D'après ce que j'ai compris: le signal cos[2πF$_1$ n] (pour F$_1$ donnée) est périodique.
On pourrait prendre par exemple F$_1$ = 1/2π .
E2: oui
M3: mais encore ?...
E3: ça fait cos[n]
M4: Peut-on dire que cos[n] est périodique?
E4: [non]
M5: Attendez... D'après ce que vous m'avez dit:
- le signal cos[2πF$_1$ n] (pour F$_1$ donnée) est périodique
- cos[n] n'est pas périodique
Je m'avoue quelque peu surpris !
E5: F$_1$ doit être entière
M6: D'après ce que j'ai compris: F$_1$ est un entier. cos[2πF$_1$ n] risque de ne pas beaucoup ressembler à une sinusoïde !
E6: pourquoi?
M7: Mais cela pose un problème... sauf si vous considérez que cos[2πF$_1$ n] est un signal constant ...
E7: F$_1$ doit être l'inverse d'un entier
M8: oui... mmm... D'accord...

Nous fournissons en annexe 1 d'autres extraits, dont un qui est obtenu à partir de la même connaissance, pour un élève qui fait des choix différents.

Afin de commenter le fonctionnement de SAVANT 3 pendant cet extrait, nous donnons ci-dessous les principales règles dont dispose le système sur ce sujet. Ces règles sont présentées (et utilisées par le système) sous forme d'incompatibilités. Par exemple la règle n°2 indique qu'il est impossible que F1 soit rationnelle et que la sinusoïde ne soit pas périodique. Le chiffre entre parenthèses indique si la règle doit être considérée comme évidente (règle de "bon sens", valeur 1) ou non (valeur 0).

règle 1 (0):
le signal cos[2πF1 n] (pour F1 donnée) est périodique
cos[n] n'est pas périodique
la fréquence F1 n'est pas rationnelle

règle 2 (0):
la fréquence F1 est rationnelle
le signal cos[2πF1 n] (pour F1 donnée) n'est pas périodique

règle 3 (0):
cos[2πF1 n] vient de l'échantillonnage de cos(2πfo t)
la fréquence F1 n'est pas rationnelle
la fréquence d'échantillonnage fe est multiple de fo

règle 4 (1):
cos[2πF1 n] vient de l'échantillonnage de cos(2πfo t)
F1 n'est pas un entier
la fréquence d'échantillonnage fe est diviseur de fo

règle 5 (0):
F1 est un entier
cos[2πF1 n] n'est pas un signal constant

règle 6 (0):
l'échantillonnage a lieu dans les conditions de Nyquist
la fréquence d'échantillonnage fe est diviseur de fo

règle 7 (1):
F1 est un entier
la fréquence F1 n'est pas rationnelle

règle 8 (1):
la fréquence d'échantillonnage fe est diviseur de fo
la fréquence d'échantillonnage fe est multiple de fo

règle 9 (0):
cos[2πF1 n] ne vient pas de l'échantillonnage de cos(2πfo t)

règle 10 (0):
cos[2πF1 n] vient de l'échantillonnage de cos(2πfo t)
F1 est un entier
la fréquence d'échantillonnage fe n'est pas diviseur de fo

règle 11 (1):
cos[n] est périodique
cos[0] est égal à 1
cos[n] n'est pas égal à 1 pour n non nul

règle 12 (1):
cos[2πF1 n] est un signal constant
F1 n'est pas un entier

règle 13 (0):
cos[n] est égal à 1 pour n non nul

règle 14 (1):
cos[0] n'est pas égal à 1

règle 15 (1):
la fréquence F1 est rationnelle
la fréquence F1 est quelconque

règle 16 (1):
le signal cos[2πF1 n] (pour F1 donnée) n'est pas périodique
la fréquence F1 n'est pas rationnelle
cos[n] est périodique

SAVANT3 utilise un moteur en chaînage arrière écrit en Prolog. Actuellement, ce moteur fonctionne de la manière suivante. Pour démontrer une proposition *p*, il recherche *non p* dans une règle de la base. Si (*non p*) $\Rightarrow$ ***F*** est une règle de la base, la démonstration réussit. Si *non p* est trouvée dans une règle en conjonction avec d'autres termes, le moteur cherche à démontrer ces termes sans utiliser *p* ni *non p*. Enfin si *non p* n'est pas trouvée, la démonstration échoue.

Ce moteur travaille donc en monde ouvert (pour une terme donné *p*, la démonstration de *p* et la démonstration de (*non p)* peuvent toutes deux échouer). Les faits appris en cours d'interaction avec l'étudiant sont ajoutés au fur et à mesure à la base de règles (dans la mesure où ils sont compatibles avec celle-ci). Le fonctionnement du moteur est décrit avec davantage de détails en annexe 2.

Nous décrivons maintenant le fonctionnement de SAVANT3 dans le dialogue que nous avons reproduit.

M1:	Le signal cos[2πF$_1$ n] (pour F$_1$ donnée) est-il périodique ?	⇨ *SAVANT3 émet la question associée par l'auteur à la proposition "*le signal ... est périodique*". L'auteur a imposé que ce terme soit connu dès le démarrage.*
E1:	oui	
M2:	D'après ce que j'ai compris: le signal cos[2πF$_1$ n] (pour F$_1$ donnée) est périodique.	⇨ *Tentative pour saturer la règle 1. Reprise des termes appris de la règle en cours de saturation.*
	On pourrait prendre par exemple F$_1$ = 1/2π .	⇨ *SAVANT3 émet la "question" associée par l'auteur à la proposition "*cos[n] est périodique*"*
E2:	oui	
M3:	mais encore ?...	⇨ oui *n'est pas un mot clé local prévu par l'auteur pour la proposition attendue, et n'est un mot clé global pour aucun autre prédicat. Mais il est reconnu comme une réponse "vide" (comme* certes, bien sûr*", etc.)*
E3:	ça fait cos[n]	⇨ *réponse non reconnue*
M4:	Peut-on dire que cos[n] est périodique?	⇨ *SAVANT3 pose la question directement, sans passer par la question prévue par l'auteur*
E4:	[non]	⇨ *l'élève répond dans un choix fermé oui/non*
M5:	Attendez... D'après ce que vous m'avez dit: - le signal cos[2πF$_1$ n] (pour F$_1$ donnée) est périodique - cos[n] n'est pas périodique Je m'avoue quelque peu surpris !	⇨ *La règle 1 n'est pas encore saturée, car le statut de son troisième terme est inconnu. SAVANT3 feint la surprise.*
E5:	F$_1$ doit être entière	⇨ *Cette réponse (reconnue par les mots clés* F1 *et* entier*) invalide l'étonnement grâce à la règle 7.*

M6:	D'après ce que j'ai compris: F_1 est un entier. cos[$2\pi F_1$ n] risque de ne pas beaucoup ressembler à une sinusoïde !	⇨ *SAVANT3 tente de saturer la règle 5 dont il connaît un terme sur deux. Il émet la question associée par l'auteur à la proposition "*cos... signal constant*"*
E6:	pourquoi?	⇨ *L'étudiant ne voit pas le lien*
M7:	Mais cela pose un problème... sauf si vous considérez que cos[$2\pi F_1$ n] est un signal constant ...	⇨ *SAVANT3 reconnaît une demande d'explication. Il explicite le terme suivant de la clause en cours de saturation.*
E7:	F_1 doit être l'inverse d'un entier	⇨ *SAVANT3 assimile cette réponse à "*F1 rationnelle*"*
M8:	oui... mmm... D'accord...	⇨ *Toutes les clauses candidates sont invalidées*

L'auteur a la possibilité d'imposer à SAVANT3 de démarrer l'interaction en recherchant la valeur de vérité d'une proposition donnée. Ainsi, en M1, le programme énonce la phrase associée par l'auteur à la proposition à valider : "*le signal cos[$2\pi F_1$ n] (pour F_1 donnée) est périodique*".

La stratégie employée par SAVANT3 pour gérer le dialogue consiste à essayer de valider une incompatibilité. Pour cela, il considère des règles "candidates", c.à.d. des règles qui ne sont pas marquées comme "évidentes", qui sont encore valides (une règle est valide si aucun de ses termes n'est démontrable comme étant faux) et qui ont au moins un terme vrai. Il va sélectionner, parmi ces règles candidates, la plus "prometteuse", c'est-à-dire celle qui contient le plus grand pourcentage de termes connus. Cette règle va provisoirement devenir la clause de référence.

Ainsi, en M2, c'est la règle 1 qui est sélectionnée. Le système connaît en effet le premier terme, qu'il vient d'apprendre, mais aussi le deuxième terme (*cos[n] n'est pas périodique)*, car il peut le démontrer avec les règles 11, 13 et 14.

Pourtant, bien qu'il sache le démontrer, SAVANT3 va émettre la phrase associée à ce deuxième terme pour amener l'élève à le valider (deuxième partie de M2). En effet, ce deuxième terme est connu du programme, mais sa démonstration fait intervenir des clauses qui ne sont marquées comme évidentes (règle 13). M2 a donc pour objectif d'amener l'étudiant à préciser si *cos[n]* est périodique, ou mieux, à lui permettre d'anticiper la contradiction.

SAVANT3 détecte des mots clés dans les réponses de l'étudiant, ainsi que la présence d'une négation. Il ne reconnaît pas E3, et propose le choix fermé M4. Après E4, tous les termes de la clause de référence sauf un sont validés. SAVANT3 peut donc s'étonner en M5. Mais l'*explication* E5 invalide la clause de référence 1 grâce à la règle 7.

Le programme change alors de clause de référence, et se porte sur la règle 5. Dans la deuxième partie de M6, il tente de connaître la validité de la deuxième proposition de la règle 5. Après E6, perçue comme une demande d'*explicitation*, le programme livre le terme suivant de sa clause de référence en M7. Mais l'élève change d'avis en E7. Le programme comprend que F_1 est rationnelle non entière, et il arrête la conversation car toutes les règles candidates (comme définies plus haut) se trouvent invalidées.

L'argumentation dépend des répliques de l'élève, qui a toujours plusieurs solutions pour rester cohérent. Ici par exemple l'élève aurait pu assumer le choix "F_1 entière", mais la machine aurait argumenté à propos de la fréquence d'échantillonnage. Il aurait pu aussi faire le choix "*cos[$2\pi F_1 n$]* non périodique", etc. (voir annexe 1).

4. L'acceptabilité d'une conversation avec SAVANT3

On peut être frappé en premier lieu par la "rigidité" de la manière de s'exprimer de SAVANT3. Cette rigidité est facile à corriger, par l'introduction d'une variabilité des messages utilisés. Nous avons choisi de ne pas effectuer cette modification tant que SAVANT3 est en phase de développement. En effet, notre espoir est que la pertinence de l'argumentation repousse au second plan ces problèmes d'expression.

SAVANT3 a déjà été testé dès 1988, ce qui a permis de cerner certaines limitations. Il a subi depuis des modifications assez profondes. La principale difficulté rencontrée concerne la clause de référence (voir p. 200). Le type d'interaction qu'il essaie d'instaurer repose en effet entièrement sur la *perception par l'élève de la contradiction* exprimée dans la clause de référence. C'est cette perception qui garantit la pertinence logique, pour l'étudiant, des répliques de la machine (ce qui n'est pas le cas pour M2) et de ses étonnements (ce qui est le cas pour M5, mais pas pour M6). Il s'agit là d'un problème délicat qui nécessite des expérimentations avec des étudiants. L'un des principaux moyens utilisés actuellement pour tenter de remédier au problème consiste à rappeler les termes connus de la clause de référence (comme en première partie de M2, M5, M6).

Puisque l'efficacité pédagogique de SAVANT3 repose entièrement sur la perception, par l'étudiant, de l'étonnement de la machine, nous pouvons aisément comprendre que le dialogue entre eux n'est possible que si tous deux manipulent les mêmes concepts. L'étudiant doit connaître les propositions contenues dans les règles de la base de connaissances (*i.e.* le sens des mots qui les désignent), ainsi qu'une partie de la théorie du domaine (*i.e.* certaines des règles de la base, avec éventuellement des termes manquants). De telles conditions exigent que le contexte de l'interaction pédagogique soit fortement cadré. Si elles ne sont pas remplies, l'interaction échoue immanquablement.

Le dialogue peut échouer pour d'autres raisons. Il peut arriver que l'étudiant ne partage pas les objectifs du dialogue, parce que le sujet ne l'intéresse tout simplement pas. Il n'entrera pas dans le dialogue, et refusera de prendre en compte le premier étonnement de la machine.

SAVANT3 ne se préoccupe pas de la négociation des objectifs du dialogue, contrairement à d'autres systèmes comme KANT [Baker 1991a] ou le système de M. Joab [Joab 1990]. De tels systèmes reconnaissent les questions de l'utilisateur qui portent sur l'opportunité de la tâche elle-même (ce sont des "méta-interventions"). Notre système néglige cet aspect, pour deux raisons principales. Premièrement, nous ne sommes pas ici en situation de résolution de tâche, et la notion d'objectif est donc moins évidente. Deuxièmement, le pari qui est à la base de l'approche est que l'élève sera motivé par le sentiment que ses déclarations sont contradictoires, comme le sont les interlocuteurs dans les conversations spontanées (M. Baker [1991b] identifie aussi le "conflit interne" comme générateur d'actes dialogiques dans des dialogues de résolution de problème). Toute la difficulté, comme on l'a dit, est donc de parvenir à faire ressentir cette contradiction le plus rapidement possible.

Rappelons enfin que, de par sa conception même, SAVANT3 a un domaine d'emploi bien délimité. Comme nous l'avons dit dans le chapitre III.1, il est destiné à l'acquisition de concepts (ou à l'élimination de fautes conceptuelles), ce qui exclut les situations d'acquisitions purement procédurales. Les tentatives pour faire traiter à SAVANT3 des sujets incluant des savoir-faire ne vont pas de soi. Nous envisagerons tout de même dans le chap. III.4 l'utilisation de SAVANT3 comme un système critique.

5. Conclusion

L'objet de ce chapitre était de présenter SAVANT3, qui est un système d'Enseignement Assisté par Ordinateur bâti sur des principes originaux issus de notre théorie de la pragmatique conversationnelle. SAVANT3 est conçu pour converser avec l'étudiant à propos de sujets techniques. Lorsque la conversation s'instaure, l'étudiant apprend : il modifie sa théorie du domaine, jusqu'à ce que celle-ci soit compatible avec celle du système.

Notre pari est que ce type d'interaction, basé sur la résolution de contradiction, sera agréable pour l'étudiant. Il est inspiré de l'observation des conversations naturelles, et nous pouvons ainsi espérer que l'étudiant prendra plaisir à argumenter avec la machine comme il le fait dans la vie courante.

L'un des intérêts principaux de SAVANT3 vient du fait qu'il repose sur peu de principes. Si SAVANT3 parvient à être validé comme moyen d'enseigner, alors il sera possible de faire remonter cette validation vers les principes qui gouvernent son fonctionnement. Ceux-ci sont par ailleurs suffisamment précis pour que l'auteur n'ait pas à gérer le dialogue. SAVANT3 est sans doute l'un des premiers systèmes qui évite à l'auteur de

se soucier de l'interaction. L'auteur se contente idéalement de fournir une connaissance au système, et c'est le système qui prend en charge la relation avec l'apprenant.

Bien entendu, SAVANT3 doit être perçu comme une partie d'un système d'EAO autosuffisant. Dans nos expérimentations, il était inséré dans un ensemble comportant un programme de présentation "narrative" des concepts, un système multimédia interactif (SAVANT2), une encyclopédie conceptuelle (SAVANT1), et des programmes de simulation. Les dialogues n'interviennent qu'aux moments où l'auteur veut s'assurer que l'étudiant a une théorie correcte du domaine qui vient de lui être présenté.

Nous décrivons, dans le chapitre III.3, le système auteur associé à SAVANT3. Ce système auteur fait encore l'objet de recherches. Notre objectif est de le baser lui aussi sur les principes de la conversation, de manière à faciliter le travail de l'auteur et ainsi à parvenir à un système viable sur le plan économique.

Annexe 1 : Exemples d'interaction avec SAVANT3

L'extrait suivant utilise la même connaissance que celle de l'extrait donné dans le texte du chapitre. Le dialogue qui s'ensuit est assez différent de celui qui a été analysé. Noter que grâce au dialogue, l'étudiant (fictif) de ce dialogue prend conscience de l'insuffisance de la contrainte *F1 non entière*.

le signal cos[2π F1 n] (pour F1 donnée) est-il périodique ?
----- non
D'après ce que j'ai compris: le signal cos[2π F1 n] (pour F1 donnée) n'est pas périodique.
tout cela a des conséquences sur F1 !
----- oui
mais encore ?...
----- F1 ne peut pas être entière
Ah, Ok... oui...mmm...
D'après ce que j'ai compris: le signal cos[2π F1 n] (pour F1 donnée) n'est pas périodique.
F1.n peut être entier.
----- F1 ne doit pas être rationnelle
oui... D'après ce que j'ai compris: la fréquence F1 n'est pas rationnelle.
le signal cos[2π F1 n] vient de l'échantillonnage à la fréquence fe d'une sinusoïde continue cos(2 π fo t).
----- certes...
Attendez... D'après ce que vous m'avez dit:
- cos[2π F1 n] vient de l'échantillonnage de cos(2π fo t)
- la fréquence F1 n'est pas rationnelle
Je m'avoue quelque peu surpris !
----- pourquoi?
Mais cela pose un problème... Si on considère que la fréquence d'échantillonnage fe est multiple de fo ...
----- non, fe n'est pas multiple de fo
D'accord...

L'extrait suivant provient d'un cours de communications numériques. Il s'agit de définir un type de codage pour transmettre une page dactylographiée. Dans ce dialogue (inspiré d'une interaction réelle), l'étudiant choisit de faire du TELETEXTE. Il aurait pu choisir un codage par pixels de type FAX. Dans les deux cas, il doit rester cohérent. C'est d'ailleurs l'exigence de cohérence qui permet au dialogue de progresser.

est-ce qu'un texte dactylographié peut constituer un message numérique?
----- oui
D'après ce que j'ai compris: la page de texte peut être considérée comme un message numérique.
il faut choisir des symboles pour décrire chaque détail de la page
----- les caractères
oui...mmm...D'après ce que j'ai compris: les points de l'image ne sont pas pris comme symboles.
que pourrait-on coder dans la page, en dehors des caractères typographiques?
----- s'il y avait des figures ?
D'après vous, on veut coder les figures (schémas, photos, signature).
C'est surprenant, dans la mesure où l'on suppose que les points de l'image ne sont pas pris comme symboles!

----- mais c'est une page dactylographiée !
D'après ce que j'ai compris: on réalise un codage de type TELETEXTE.
il n'y a pas que les caractères typographiques à coder...
----- non
mais encore ?...
----- la mise en page, les marges
oui...mmm... Attendez... D'après ce que vous m'avez dit:
- on veut coder les positionnements (paragraphes, etc...)
Je m'avoue quelque peu surpris !
----- il faut des codes spéciaux
Ah bon!... D'accord...

La connaissance utilisée pour cet extrait est donnée ci-dessous :

règle 1 (1):
les points de l'image sont pris comme symboles
les ordres de traitement de texte font partie des symboles
on ne réalise pas un codage mixte (texte + images)

règle 2 (1):
la page de texte ne peut pas être considérée comme un message numérique
les points de l'image sont pris comme symboles

règle 3 (1):
la page de texte ne peut pas être considérée comme un message numérique
les caractères typographiques sont pris comme symboles

règle 4 (1):
on réalise un codage de type FAX (télécopie)
les points de l'image ne sont pas pris comme symboles

règle 5 (0):
On réalise un codage de type TELETEXTE
on veut coder les attributs des caractères (police, taille, lignage,...)
les ordres de traitement de texte ne font pas partie des symboles

règle 6 (1):
on ne réalise pas un codage de type TELETEXTE
les caractères typographiques sont pris comme symboles

règle 7 (0):
la page de texte peut être considérée comme un message numérique
on ne réalise pas un codage de type FAX (télécopie)
on ne réalise pas un codage de type TELETEXTE
on ne réalise pas un codage mixte (texte + images)

règle 8 (0):
la page de texte ne peut pas être considérée comme un message numérique

règle 10 (0):
on veut coder les figures (schémas, photos, signature)
les points de l'image ne sont pas pris comme symboles

règle 11 (1):
on réalise un codage de type FAX (télécopie)
on réalise un codage de type TELETEXTE
on ne réalise pas un codage mixte (texte + images)

règle 12 (0):
on réalise un codage mixte (texte + images)

règle 13 (1):
on réalise un codage de type TELETEXTE
les caractères typographiques ne sont pas pris comme symboles

règle 14 (1):
on ne réalise pas un codage de type FAX (télécopie)
les points de l'image sont pris comme symboles

règle 19 (0):
on réalise un codage de type TELETEXTE
on veut coder les positionnements (paragraphes, etc...)
les ordres de traitement de texte ne font pas partie des symboles

L'extrait suivant fait aussi partie d'un cours de Communications Numériques. Il traite de l'interférence

entre symboles (IES). L'élève finit par réaliser que l'on peut annuler l'IES (chevauchement des impulsions à l'instant de l'échantillonnage) sans exiger que les impulsions ne se chevauchent pas du tout. Noter la surprise prématurée manifestée à la fin de l'extrait par le système (les réponses entre crochets indiquent une réponse à un choix fermé).

Les termes qui provoquent l'IES: p((n-k)T), peuvent-ils être tous nuls (pour n différent de k) ?
----- oui, je crois
D'après ce que j'ai compris: on ne modifie pas les paramètres de l'émetteur.
On peut peut-être jouer sur la durée des impulsions MIA, pour résoudre ce problème d'IES ?
----- on peut raccourcir les impulsions
D'après vous, on choisit des impulsions p(t) plus courtes.
C'est surprenant, dans la mesure où l'on suppose que on ne modifie pas les paramètres de l'émetteur !
----- pourquoi ne peut-on pas changer l'émetteur?
On ne modifie pas les paramètres de l'émetteur parce que cela fait partie des données du problème.
D'après vous, on choisit des impulsions p(t) plus courtes. C'est surprenant, dans la mesure où l'on suppose que on ne modifie pas les paramètres de l'émetteur!
----- on ne les prend pas plus courtes
D'après ce que j'ai compris: les termes d'IES: p((k-n)T) sont nuls (avec k≠ n).
On suppose au départ que les impulsions se chevauchent. Comptez-vous supprimer ce chevauchement ?
----- il faut écarter les impulsions
Oui... mmm... D'après vous, on ralentit l'émission des symboles. C'est surprenant, dans la mesure où l'on suppose que on ne modifie pas les paramètres de l'émetteur!
----- Ok
Mais cela pose un problème... D'après ce que j'ai compris:
- on ralentit l'émission des symboles
- on ne modifie pas les paramètres de l'émetteur
Que proposez-vous comme solution ?
----- [on ne ralentit pas l'émission des symboles]
Attendez... D'après ce que vous m'avez dit:
- les termes d'IES: p((k-n)T) sont nuls (avec k≠ n)
- on ne ralentit pas l'émission des symboles
Je m'avoue quelque peu surpris !
----- on peut changer la forme des impulsions
Ah bon!... Attendez... D'après ce que vous m'avez dit:
- on change la forme p(t) de l'impulsion pour annuler l'IES
- on ne modifie pas les paramètres de l'émetteur
Je m'avoue quelque peu surpris !
----- si on filtre, ça change bien leur forme
Ah bon!... D'accord...

Le dialogue suivant met en scène un étudiant qui aurait des conceptions fortement erronées sur la notion de filtrage. Noter que le système n'émet jamais de jugement de valeur. Il se contente de montrer l'absurdité de certaines réponses (*e.g.* "passe-haut").

si vous considérez la suspension de votre voiture comme un filtre, quel genre de filtre est-ce ?
----- passe-haut
D'après vous, votre suspension est un filtre passe-haut. C'est surprenant, dans la mesure où l'on suppose que le système de suspension est transparent aux très basses fréquences!
----- il n'est pas transparent aux basses fréquences
D'après vous, le système de suspension n'est pas transparent aux très basses fréquences. C'est surprenant, dans la mesure où l'on suppose que la voiture suit les variations lentes du terrain!
----- ah oui
Mais cela pose un problème... D'après ce que j'ai compris:
- la voiture suit les variations lentes du terrain
- le système de suspension n'est pas transparent aux très basses fréquences
Que proposez-vous comme solution ?
----- [le système de suspension est transparent aux très basses fréquences]
D'après vous, votre suspension est un filtre passe-haut. C'est surprenant, dans la mesure où l'on suppose que le système de suspension est transparent aux très basses fréquences!
----- c'est un passe-bande
Donc, vous considérez maintenant que votre suspension n'est pas un filtre passe-haut. Est-ce bien le cas?
----- [votre suspension n'est pas un filtre passe-haut]
D'après ce que j'ai compris: votre suspension peut entrer en résonance. votre suspension n'est pas en très bon état, n'est-ce pas ?
----- elle n'entre pas en résonance
Oui... mmm... D'après vous, votre suspension ne peut pas entrer en résonance. C'est surprenant, dans la mesure où l'on suppose que votre suspension est un filtre passe-bande!
----- Ok, ce n'est pas un passe-bande
D'après ce que j'ai compris: votre suspension est un filtre passe-bas. considérez-vous votre suspension comme "dure" ?
----- non
D'accord...

Annexe 2 : Le moteur de clauses paradoxales

1. Principe du démonstrateur

Il n'y a pas de conversation possible sans raisonnement. Dans les applications comme SAVANT3 et PARADISE, ces raisonnements sont effectuées sur des clauses négatives, que nous appelons clauses paradoxales. En effet, tout ensemble de formules logiques peut être mis sous la forme d'une conjonction de disjonctions. Nous écrivons les disjonctions comme des conjonctions fausses, (*a* ***ou*** *b*) devenant [***non*** *a* ***& non*** *b*] ⇒ ***F***. ***F*** désigne une proposition toujours fausse. La connaissance devient ainsi une conjonction de clauses négatives (ou clauses paradoxales) de ce type.

Nous devons maintenant expliquer comment sont menées les démonstrations. Le moteur que nous décrivons maintenant fonctionne en ordre 0 dans SAVANT3, et en ordre 1 dans PARADISE. Nous le décrivons avec des propositions pour alléger l'écriture.

Pour démontrer un terme *p*, la stratégie adoptée est la suivante :

- soit à prouver *p*
- rechercher ***non*** *p* dans une clause ne contenant aucun terme (ni son contraire) qui soit en cours de démonstration (hormis bien entendu ***non*** *p*)
- ***non*** *p* trouvé dans la clause *C*
- prouver les autres termes de *C* s'il y en a

Insistons sur les avantages et les inconvénients d'un tel moteur par rapport à un moteur classique comme celui de Prolog.

Dans un ensemble de clauses de Horn comme celles d'un programme Prolog, les termes démontrables sont désignés à l'avance, et sont dans la "tête" de la clause. De plus, ces termes ne peuvent pas être niés. Ceci entraîne certaines limitations. Par exemple, à partir de la clause Prolog :

a ← b, c

le moteur Prolog ne peut pas établir ***non*** *b* à partir de la connaissance de ***non*** *a* et de *c*, alors que c'est logiquement possible. En revanche, à partir de la clause paradoxale :

[**non** a **&** b **&** c] ⇒ **F**

notre moteur peut conclure à la négation de chacun des trois termes à partir de la connaissance des deux autres.

Bien entendu, ces possibilités plus étendues, dont nous ne pouvons pas nous passer dans la mesure où nous voulons analyser et reproduire les raisonnements conversationnels, entraînent un temps de réponse plus long. Le démonstrateur doit en effet essayer un nombre important de combinaisons. Nous avons donc dû consacrer nos efforts à diminuer le temps de réponse.

2. Optimisation

Nous avons implanté et optimisé ce moteur, de manière à ce qu'il fonctionne efficacement tant en ordre 0 qu'en ordre 1. L'optimisation consiste principalement à mémoriser les résultats partiels. Elle est compliquée par le fait qu'il existe plusieurs contextes de démonstration, car les raisonnements peuvent utiliser, ou ne pas utiliser :

- les clauses marquées comme évidentes
- les clauses avec la modalité "indésirable" (pour certaines applications)

- les déclarations de l'utilisateur
- des hypothèses "essayées" par le programme utilisant le moteur

3. Monde ouvert

Une caractéristique essentielle de notre moteur est qu'il fonctionne en monde ouvert. Un terme *p* peut être vrai parce que démontrable, ou faux parce que ***non*** *p* est démontrable, ou encore indéterminé parce que ni *p* ni ***non*** *p* ne sont démontrables. Cette propriété est absolument nécessaire dès qu'il s'agit de reproduire les raisonnements utilisés dans les conversations.

4. Révision

Le moteur peut prendre en compte les déclarations de l'utilisateur dans ses raisonnements, à condition que ces déclarations soient compatibles avec la connaissance antérieure au moment où elles sont affirmées. Lorsque ce n'est pas le cas, le programme mémorise une erreur que l'application doit gérer (par exemple SAVANT3 "s'étonne" lorsqu'une erreur est commise).

Lorsque l'utilisateur fait une déclaration qui est l'exacte négation d'une de ses déclarations antérieures, le programme ne constate pas une erreur. Il révise la connaissance apprise, et retire de sa mémoire tous les résultats partiels établis avec les déclarations de l'utilisateur.

5. Complétude

Ce moteur présente un inconvénient : il n'est pas complet. En d'autres termes, il peut arriver que certains termes soient vrais (et logiquement démontrables) sans que le moteur soit capable de les valider. Considérons la connaissance suivante :

[a **&** **non** b] ⇒ **F**
[**non** a **&** c] ⇒ **F**
[b **&** c] ⇒ **F**

Il est possible de déduire ***non*** *c* de cette connaissance : si *a* est vrai, alors *b* est vrai (première clause), ce qui entraîne ***non*** *c* (troisième clause) ; si ***non*** *a* est vrai, alors ***non*** *c* est encore vrai (deuxième clause). Or même dans le monde ouvert dans lequel nous nous plaçons, on peut assurer que soit *a*, soit ***non*** *a* est vrai. Donc ***non*** *c* est vrai.

Pourtant, le moteur que nous utilisons dans nos applications est incapable d'établir ***non*** *c*. Il va adopter la démarche suivante :

- soit à prouver ***non*** *c*
- rechercher *c* dans une clause
- *c* trouvé dans la deuxième clause
- prouver ***non*** *a*
- rechercher *a* dans une clause ne contenant pas *c*, ***non*** *c*
- *a* trouvé dans la première clause
- prouver ***non*** *b*
- rechercher *b* dans une clause ne contenant pas *c*, ***non*** *c*, *a*, ***non*** *a*
- échec
- *c* trouvé dans la troisième clause (backtracking)
- prouver *b*
- rechercher ***non*** *b* dans une clause ne contenant pas *c*, ***non*** *c*
- ***non*** *b* trouvé dans la première clause
- prouver *a*
- rechercher ***non*** *a* dans une clause ne contenant pas *c*, ***non*** *c*, *b*, ***non*** *b*
- échec

La solution consiste à adopter une stratégie de preuve par réfutation. Nous avons développé un nouveau moteur, dont il est facile de montrer qu'il est complet, puisqu'il implante directement une stratégie de résolution correspondant à la technique de la résolvante [Haton et al. 1992]. La stratégie peut être décrite de la manière suivante :

- soit à prouver *p*
- ajouter la règle $p \Rightarrow \boldsymbol{F}$ à la base de connaissances
- tenter de saturer une règle en prouvant tous ses termes
- en cas de "backtracking", ôter les règles qui ont été ajoutées

Cette stratégie est d'ailleurs tout à fait dans l'esprit conversationnel. Une proposition *p* est vraie parce que son contraire ***non*** *p* provoquerait une incohérence.

Nous avons implanté ce moteur, et commencé à l'optimiser (y compris en ordre 1). Mais nous ne l'avons pas encore intégré à nos applications. Il est à noter que nous n'avons jamais pu constater, dans les dialogues, un quelconque défaut qui serait dû à l'incomplétude de notre premier moteur.