

Pascal SALAMBIER

Étude empirique et modélisation d'une activité de diagnostic cognitif

Introduction

Le développement grandissant depuis quelques années de l'étude des activités cognitives mises en œuvre dans des situations complexes naturelles, et plus particulièrement dans les situations de travail, a conduit à reconsidérer certaines questions théoriques et méthodologiques abordées jusqu'alors dans le seul cadre de travaux menés selon une approche expérimentale classique.

Ainsi une recherche menée sur l'activité de correction de texte en situation réelle a-t-elle permis d'invalidier certains modèles psycholinguistiques d'accès au lexique (Pavard, 1987) basés sur des résultats obtenus en situation artificielle sur du matériel non signifiant.

Au niveau méthodologique, cette nouvelle perspective d'appréhension de l'activité cognitive comme totalité complexe organisée se caractérise par l'abandon d'une approche favorisant un découpage a priori et plus ou moins artificiel de facteurs jugés principaux. Elle met l'accent sur le caractère naturel des situations étudiées, et privilégie une analyse fine des activités, essentiellement de nature qualitative, mais revendiquant une certaine rigueur du point de vue des modalités de recueil des données, et de la reproductibilité de l'analyse des protocoles (formalisation et

interprétation). Pour que les conditions de cette reproductibilité soient assurées, il est nécessaire que les protocoles utilisés soient accessibles (Woods, sous presse¹) et que les options théoriques sous-jacentes qui ont guidé le codage et l'interprétation des données soient explicitées.

Au niveau théorique, cette approche se démarque d'une conception logiciste de la rationalité. Cette distanciation n'implique pas l'adoption d'une position inverse selon laquelle l'activité d'un opérateur travaillant dans une situation de travail complexe, serait totalement irrationnelle², mais plutôt que l'activité de cet opérateur est traversée par une rationalité propre qui répond de façon satisfaisante aux caractéristiques et aux modifications de son environnement, compte tenu de ses connaissances, de ses objectifs et de ses ressources cognitives à un moment donné. Il ne s'agit donc plus d'étudier ici une cognition idéalisée mais une cognition "pratique" telle qu'on peut empiriquement en observer les manifestations en situation naturelle.

Après avoir précisé brièvement certaines des options sur lesquelles se fondent notre approche, nous présenterons, au travers d'une étude centrée sur une activité de diagnostic cognitif en situation d'autoformation, un exemple des apports possibles de cette démarche d'analyse des activités cognitives en situation réelle.

I. Un cadre général d'analyse des situations complexes orienté vers la conception : l'ingénierie cognitive

La perspective présentée ici s'inscrit dans un champ d'investigation théorique et pratique qui a connu depuis quelques années un développement important lié à une demande croissante émanant de

¹ Malheureusement, comme le fait remarquer Woods, les pratiques habituelles des éditeurs de publications et des auteurs ne favorisent pas la satisfaction de ce prérequis pour la reprise des analyses et des interprétations effectuées par un tiers.

² Ce qui aurait pour conséquence d'hypothéquer gravement toute entreprise d'étude et de compréhension de cette activité (cf. Cherniak, 1986).

situations réelles complexes, et plus particulièrement des situations de travail. L'*ingénierie cognitive* (Woods et Roth, 1988 ; Norman, 1986 ; Rasmussen, 1986)³ se réfère à l'émergence d'un sous-champ disciplinaire (pour reprendre la formulation de Hoc à propos de l'ergonomie cognitive) dont l'objectif est de fonder la conception de dispositifs d'interaction homme-machine sur des principes, des connaissances issues des sciences cognitives et plus particulièrement de la psychologie cognitive.

L'ingénierie cognitive trouve son origine dans l'émergence massive d'outils informatiques dans les situations professionnelles, et plus particulièrement dans la disponibilité récente de systèmes intelligents

³ Soyons clairs : nous avons tout à fait conscience en employant ce terme de participer, même involontairement, à l'inflation sans doute regrettable d'appellations diverses qui toutes se réfèrent d'une façon ou d'une autre à une démarche visant à l'utilisation de connaissances sur le fonctionnement cognitif pour la conception de systèmes interactifs. Mais cette multiplication des termes va au-delà d'une quelconque volonté de se singulariser à tout prix et révèle parfois des orientations fondamentalement différentes. Le terme très connoté de "psychologie de l'ingénierie" par exemple (*engineering psychology*, Gopher et Kimchi, 1989) renvoie à une approche, d'origine essentiellement américaine, qui vise à l'application des résultats de laboratoire issus de la psychologie expérimentale, à la conception de dispositifs de présentation de l'information (cf. Montmollin, 1984). L'ergonomie cognitive telle qu'elle a pu être définie par Hoc notamment (Hoc, 1990a) nous semble bien s'inscrire dans la même perspective que l'ingénierie cognitive dans le sens où nous l'entendrons ici. Par contre l'ergonomie cognitive définie par Rasmussen (1987) s'en éloigne notablement et s'apparente plutôt à ce que les anglo-saxons appellent *human-computer interaction*. Le terme de *cognitive engineering* lui-même prendra un sens différent selon qu'on le traduit par génie cognitif ou par ingénierie cognitive ; le génie cognitif subordonne en effet la phase d'analyse et de description formalisée de l'activité cognitive à la production rapide d'un modèle informatique de cette activité, ce que ne fait pas l'ingénierie cognitive dans l'acception qui sera la sienne ici. Nous retiendrons cette dénomination, non pour des raisons de priorité chronologique mais parce que les aspects fondamentaux propres à cette démarche ont été particulièrement bien définis par Woods et Roth (1988) sous cette appellation, notamment concernant sa finalité, orientée sur la conception de systèmes coopératifs (Woods, Roth et Bennett, 1987).

d'aide au travail (prise de décision, diagnostic, planification, résolution d'incidents).

Elle s'attache à apporter des éléments de solutions à des problèmes qui ne relèvent pas directement de l'informatique (et de l'IA) vue sous son angle purement technologique ; par exemple : comment spécifier une mise en oeuvre efficace des ressources techniques disponibles, qui optimise les performances du couple opérateur/machine, ou quel type d'assistance faut-il fournir à un opérateur engagé dans une tâche complexe ?

L'ingénierie cognitive peut être caractérisée par un certain nombre de traits qui la distinguent d'autres approches de conception de systèmes interactifs⁴ :

- ILIC traite de *l'activité d'opérateurs dans des situations complexes* (typiquement des situations de travail). Pour l'IC la réduction artificielle de cette complexité, sa subordination à des décisions de nature exclusivement technique ne fait qu'éluder temporairement le problème de la prise en compte de cette complexité, qui réapparaîtra inévitablement en situation d'utilisation dans des conditions nettement plus défavorables. Pour éviter cette faillite dans le processus de conception il s'avère nécessaire d'étudier et de mieux comprendre l'activité d'opérateurs dans des situations complexes afin d'identifier quels sont les éléments qui déterminent cette complexité, comment elle est prise en compte par l'opérateur et dans quels cas cette gestion de la complexité au travers de l'activité n'est plus possible.

- L'IC met l'accent sur le critère de *validité écologique*. L'IC traite de situations du monde réel non délimitées par des frontières posées a priori

⁴ Cette partie reprend certaines des idées formulées par Woods et Roth, tout en introduisant des modulations sur des points particuliers et en introduisant de nouveaux éléments notamment en ce qui concerne le rôle de l'analyse des activités complexes en situation réelle.

pour les besoins d'une expérimentation. Ces situations constituent le lieu privilégié d'investigations scientifiques et de recherches de solutions pratiques à des problèmes concrets de conception.

- L'IC traite *la sémantique du domaine abordé*. Envisager la conception d'un dispositif d'assistance indépendamment du contenu des tâches oblitère dès avant sa réalisation l'efficacité d'un tel système (Woods et Roth, 1988 ; Hoc, 1990a) ; c'est pourquoi l'IC s'attache à identifier les caractéristiques du domaine déterminantes dans l'activité des opérateurs, et à envisager l'introduction d'un nouveau dispositif par rapport à ces caractéristiques. Corollaire obligé de cette perspective : l'utilisation d'un langage de description de l'activité cognitive indépendant d'une situation particulière, qui permette de s'abstraire des langages d'expression propres au domaine ou au dispositif technique considéré.

- L'IC s'attache à traiter des problèmes relatifs au couplage opérateur/système en privilégiant l'aspect *interface avec la tâche* (Alty & al., 1985) par rapport à l'aspect interface-utilisateur. Il apparaît en effet que la majorité des problèmes critiques constatés dans les situations d'interaction homme/ordinateur est due soit à une inadéquation fonctionnelle, soit à une impossibilité pour l'opérateur d'appréhender le champ des compétences du système ou d'évaluer le niveau de correction des réponses fournies (Goransson et al., 1988).

- L'IC implique *un mode d'appréhension empirique* des activités cognitives. Par rapport à la formulation initiale de Woods et Roth, nous serions tentés d'ajouter que l'analyse de l'activité en situation réelle ou recréée constitue la forme privilégiée de cette appréhension empirique. Cette prééminence des études de terrain n'interdit pas le recours à des expérimentations où l'on s'efforce de se rapprocher le plus possible de la réalité.

- La nature des *problèmes théoriques abordés* dans le champ de l'IC est délimitée en regard de leur pertinence, définie par les objectifs pragmatiques poursuivis (objectifs de conception notamment).

- D'un point de vue *méthodologique*, le but de l'IC n'est pas de produire des résultats quantitatifs soumis à des modes de validation classiques essentiellement statistiques, mais de générer des modèles descriptifs et explicatifs plutôt que normatifs de l'activité d'opérateurs engagés dans une tâche ou un type de tâche donné. Cette orientation conduit à favoriser l'analyse fine d'un nombre limité d'enregistrements de l'activité, par rapport à une multiplication de relevés essentiellement quantitatifs visant à apporter les matériaux nécessaires à certains types d'analyses mathématiques. Mettre l'accent sur des analyses de type qualitatif ne veut pas dire sacrifier toute rigueur dans l'appréhension des problèmes traités. Mais cette rigueur se focalise plus sur la reproductibilité des méthodes de recueil et d'analyse que sur les outils statistiques de validation.

II. Le diagnostic cognitif

1. L'activité de diagnostic dans les situations complexes

L'activité de diagnostic est généralement apparentée à la catégorie plus large des activités de résolution de problèmes d'induction de structure. Dans ce cadre, l'activité de diagnostic a été décrite comme un processus d'identification de la classe d'événements à laquelle appartient la situation traitée. Ce processus d'identification consiste à établir une correspondance entre cette situation et les descripteurs des classes connues.

La référence aux situations de diagnostic médical a grandement contribué à la détermination de cette conception "orientée classification" de la résolution des problèmes de diagnostic (association entre symptômes et pathologies). C'est à partir de travaux menés dans ce domaine que Clancey par exemple a élaboré un modèle général des situations de diagnostic basé sur la mise en oeuvre successive de trois mécanismes : abstraction des données, appariement heuristique et

raffinement (Clancey, 1985). Le domaine de validité de cette conception reste toutefois limité aux domaines où pré-existe déjà une catégorisation des problèmes. Dans les situations où cette condition n'est pas vérifiée, l'activité de diagnostic ne peut être ramenée à une simple activité de classification.

Une des spécificités du diagnostic dans les situations de travail réside dans le fait que le résultat du diagnostic en lui-même n'est que très rarement l'objectif final poursuivi par l'opérateur. L'activité de diagnostic est la plupart du temps finalisée par une prise de décision qui peut se concrétiser par une action immédiate ou différée, ou même par l'absence d'action.

Si le résultat du diagnostic détermine la prise de décision, l'inverse est également vrai : le cheminement de l'activité de diagnostic peut en effet varier notablement selon les objectifs que s'est assignés l'opérateur. Bainbridge (1981), par exemple, a montré dans le domaine du contrôle de processus que le type de diagnostic auquel se livrent les opérateurs diffère selon qu'ils se donnent comme objectif de maximiser la production ou de la freiner. Dans une situation de même type, Hoc (1989) a mis en évidence que le champ des hypothèses envisagées par les opérateurs était fortement contraint par leurs possibilités d'action.

L'activité de diagnostic peut être d'autre part être influencée par les caractéristiques de la situation ; dans certaines situations à haut risque où les contraintes temporelles sont fortes par exemple, la prise de décision doit parfois être extrêmement rapide. Le niveau de profondeur du diagnostic effectué par l'opérateur pourra donc varier en fonction de la pression temporelle et d'une estimation des risques encourus : dans un premier temps, un diagnostic grossier débouchera sur une action permettant de stabiliser une évolution critique des paramètres considérés ; un diagnostic plus poussé pourra être ensuite mené si nécessaire, pour parvenir à une compréhension plus fine de la situation (Abbott, 1988).

Une autre caractéristique fondamentale de l'activité de diagnostic tient au rôle important joué par le niveau d'expérience de l'opérateur, défini par le nombre de situations différentes traitées dans le passé. Il s'avère en effet que certains problèmes de diagnostic ne peuvent être résolus par la simple mise en oeuvre de connaissances théoriques du domaine, et requièrent une expérience antérieure du problème.

La caractérisation d'une activité de diagnostic est étroitement dépendante du type de domaine dans lequel elle est étudiée ; l'importance relative des différents aspects de cette activité variera selon que l'on s'intéresse au diagnostic médical, au diagnostic dans les situations de contrôle de processus dynamique ou au diagnostic de pannes sur des dispositifs techniques.

Des cadres généraux de description des différentes phases des activités de diagnostic ont été néanmoins proposés qui tentent de faire la synthèse des résultats obtenus dans diverses situations.

Hoc (1990b), par exemple, a proposé de distinguer cinq composantes des activités de diagnostic :

- l'acquisition d'indices par exploration plus ou moins contrôlée de l'environnement, et par recherche orientée par des hypothèses préalables ;
- la génération d'hypothèses qui fait intervenir différents niveaux et types de connaissances ; les hypothèses retenues permettent d'orienter les stratégies de prise d'information et de réduire le nombre de lignes de raisonnement possibles (Woods, Roth & Pople 1987) ;
- l'interprétation d'indices, qui tient compte d'éléments divers (valeurs de certains paramètres, contexte, etc.) afin de fournir une explication cohérente du point de vue de l'opérateur qui permette de rendre compte de l'ensemble des indices identifiés ;
- l'évaluation des hypothèses qui peut être effectuée par confrontation entre données observées et données prévues, ou par intervention directe de l'opérateur ;

- la localisation d'une composante défectueuse ; cet aspect est caractéristique des situations où l'opérateur doit localiser l'élément responsable d'une panne sur un dispositif.

La combinaison et l'ordre de mise en oeuvre de ces différentes composantes déterminent des types de stratégie de diagnostic qui varient selon la situation et le niveau d'expertise de l'opérateur. La typologie proposée par Rasmussen constitue la référence la plus fréquemment citée pour illustrer les tentatives de classification de ces stratégies. Elle s'appuie essentiellement sur la distinction entre référence à un modèle normal du fonctionnement (stratégie de recherche topographique) et référence à un modèle anormal du fonctionnement (stratégie de recherche symptomatique, cf. Rasmussen, 1986).

2. *Le diagnostic cognitif en EIAO*

Le suivi et l'assistance d'un apprenant engagé dans la résolution d'un problème nécessite la construction d'une représentation de cet apprenant (profil général, but courant, connaissances acquises, lacunes, etc.). Cette activité de recueil et d'actualisation d'informations à propos de l'apprenant, généralement référencée sous le terme de *diagnostic cognitif*⁵ (Ohlsson, 1987), requiert la mise en oeuvre de mécanismes d'inférence plus ou moins complexes.

Wenger (1987) a proposé un cadre général d'analyse du diagnostic cognitif dans lequel il distingue trois niveaux.

- *Le niveau comportemental* : à ce niveau le diagnostic porte sur les comportements ou le produit de ces comportements, sans chercher à inférer l'état des connaissances sous-jacent à ces comportements.

⁵ Van Lehn (1988) fait le parallèle avec l'activité de diagnostic médical qui consiste à inférer un état physiologique caché (une pathologie) à partir de signes observables (symptômes). Dans le cas du diagnostic cognitif il s'agit d'inférer un état cognitif interne, par définition non accessible, à partir des manifestations observables de cet état.

- *Le niveau épistémique* : le diagnostic consiste ici en une interprétation des comportements qui donne lieu à une évaluation de l'état des connaissances de l'élève.
- *Le niveau individuel* concerne les aspects ayant trait aux caractéristiques particulières de l'élève.

A partir d'une reformulation critique de cette distinction entre diagnostic comportemental et diagnostic épistémique, Balacheff (1991) propose de distinguer entre deux types de modélisation jusqu'à maintenant confondues dans l'architecture des tutoriels intelligents (TI) :

- *Le modèle comportemental* est élaboré à partir du traitement des événements observables au niveau de l'interface entre le système et l'élève. Ce traitement procède par filtrage sélectif des événements pertinents et par codage de certaines séquences d'événements par un descripteur de plus haut niveau. Cette transformation est désignée sous le terme *d'homomorphisme de comportement*.
- *Le modèle épistémique* est dérivé du modèle comportemental ; le traitement qui permet ce passage s'appuie sur les événements observables afin de reconstituer le processus de résolution de problème et l'état des connaissances de l'élève. C'est ce traitement qui constitue pour l'auteur le diagnostic cognitif proprement dit.

Les techniques de diagnostic actuellement implémentées dans les TI ont été passées en revue par Van Lehn (1988) auquel on pourra se référer pour plus de précisions sur la question. On peut de façon schématique (et donc quelque peu grossière) classer ces techniques selon trois approches :

L'évaluation purement syntaxique qui juge le comportement de l'élève en termes d'écart à un modèle de référence.

L'encodage et le stockage d'une expertise de diagnostic spécifique du domaine d'enseignement décrit des configurations de comportements typiques (souvent des erreurs).

La simulation dynamique du fonctionnement cognitif de l'élève.

On notera deux points particuliers concernant l'approche de l'activité de diagnostic cognitif en EIAO :

Les techniques de diagnostic employées mettent l'accent sur la mise en œuvre de mécanismes parfois extrêmement sophistiqués d'acquisition et de traitement des comportements de l'élève, mais recourent peu à des processus inférentiels opérant sur les informations déjà acquises et faisant appel à des expériences antérieures. Dans le domaine de la modélisation de l'utilisateur, seuls quelques systèmes, non dédiés spécifiquement à l'enseignement, utilisent des modèles qui leur permettent d'inférer des données à propos d'un utilisateur à partir d'une phase préalable d'acquisition d'informations sur cet utilisateur⁶.

L'accent est mis sur le processus de modélisation de l'élève par le système, mais sa réciproque, c'est-à-dire la construction d'une représentation du système par l'élève, n'est pas considérée ; ce qui paraît pour le moins étrange surtout si l'on rapproche ce fait de propositions récentes visant à reconsidérer les objectifs de l'EIAO dans la perspective des *systèmes de communication de la connaissance* (Wenger, 1987).

⁶ Voir par exemple les systèmes GRUNDY (Rich, 1979), KNOME (Chin, 1986) et GUM (Kass et Finin, 1988).

3. Les études empiriques de l'activité tutorielle

L'intérêt, si ce n'est la nécessité, de recourir à des études de l'interaction tutorielle préalablement à la conception d'un système d'enseignement automatisé a été souligné à de nombreuses reprises (Rosenberg, 1987 ; Ford, 1986 ; Halff, 1988 ; Ohlsson, 1987). Un certain nombre de travaux se sont engagés dans cette voie pour tenter de mettre en évidence les stratégies pédagogiques utilisées par les tuteurs et/ou d'identifier des stratégies de contrôle du dialogue tutoriel (Salembier, 1990) : on notera entre autres les études de Collins et al.(1975), de Stevens et Collins (1977), de Littman, Pinto et Soloway (1986), et de Leinhart et Greeno (1986). Mais peu d'études se sont spécifiquement intéressées au versant diagnostic de l'activité tutorielle⁷.

Si l'on élargit le champ des situations de dialogue tutoriel aux situations de dialogue expert-consultants, on trouve un certain nombre d'études dont les résultats permettent d'identifier des dimensions importantes des activités d'assistance⁸.

Aaronson et Carroll (1987) par exemple ont mis en évidence que les consultants sont souvent engagés dans une démarche de négociation avec l'expert et ne se cantonnent pas dans une attitude d'acceptation passive d'une solution. Les résultats obtenus indiquent en effet que la plupart des questions posées par les consultants s'apparentent à des requêtes de vérification où ils proposent déjà, à un niveau de finesse variable, un premier diagnostic du problème rencontré.

Un dialogue de consultation peut également dans certains cas prendre la forme d'une véritable coopération expert-consultant orientée vers la résolution du problème posé. Dans cette perspective De Medeiros (1991)

⁷ On notera par exemple le travail de Jerrams-Smith (1989) qui malheureusement subordonne la finesse de l'analyse des données recueillies à des considérations "d'implémentabilité" rapide des résultats.

⁸ Outre les travaux auxquels il est fait allusion ici on pourra se référer entre autres aux études de Coombs et Alty (1980), McKendree et Carroll (1986).

a montré dans le cadre d'une étude sur le diagnostic téléphonique, qu'une partie importante de l'activité de l'expert consistait alors à tenter de rendre plus homogènes les représentations du problème (la sienne et celle du consultant) afin de faciliter les échanges. Cette homogénéisation des représentations peut passer par une phase préalable de mise à niveau des connaissances du consultant ; si l'expert, après évaluation rapide du niveau de compétence de son interlocuteur, estime cette phase trop coûteuse en temps il changera alors de stratégie de conduite du dialogue et prendra le contrôle de l'initiative. Toujours dans le domaine du diagnostic téléphonique, l'importance de cette phase préalable d'évaluation des connaissances du consultant a été mise en évidence de façon plus spécifique dans une étude menée par Cahour et Falzon (Falzon, 1987 ; Cahour, 1988). Les résultats obtenus mettent l'accent sur l'importance de cette évaluation et sur son influence sur l'activité de diagnostic proprement dite ; le résultat de cette évaluation constitue en effet un indice important dans le processus de catégorisation par l'expert du problème qui lui est décrit par le client au téléphone, ainsi que dans la conduite du dialogue.

On notera que dans ce type de situation la partie évaluation des connaissances du consultant (que l'on pourra assimiler en partie à l'activité de diagnostic cognitif dans les dialogues tutoriels) ne constitue pas la part principale de l'activité : l'objectif primordial de l'expert est de diagnostiquer l'origine du problème qui lui est soumis et de proposer une solution. Plus généralement ces études posent le problème de la transposition des résultats obtenus dans des situations de dialogue homme/homme à des situations d'interaction homme-machine, où du fait des limitations technologiques, il n'est pas toujours possible (ni même souhaitable d'ailleurs) de simuler toutes les caractéristiques des situations d'interaction homme-homme. En ce qui concerne les systèmes d'enseignement automatisé par exemple, une des limitations technologiques les plus apparentes a trait à la restriction du champ des

données susceptibles d'être traitées : lors d'une interaction tutorielle, l'enseignant utilise un certain nombre d'indices mimo-gestuels et auditifs (intonation) afin d'inférer l'état de l'élève à un moment précis (hésitation, doute, compréhension, etc.). Ces indices sont captés par le biais de modalités sensorielles (vision, audition) dont ne disposent pas les tuteurs artificiels, qui doivent la plupart du temps se contenter d'informations telles que les entrées au clavier.

Figure 1. Limitation du spectre des indices accessibles aux systèmes d'enseignement automatisé par rapport à un tuteur humain. On qualifie habituellement cette restriction de problème de la largeur de bande (bandwidth) en référence à la notion de bande passante.

De ce point de vue les études qui se sont efforcées de reproduire ces restrictions de la communication entre deux interlocuteurs en recourant à des dispositifs tels que télétype, téléphone, terminaux connectés marquent un progrès certain quant au degré de réalisme des situations d'expérimentation. Mais cette simulation des contraintes matérielles de

prise d'information ne suffit pas pour espérer reproduire une future situation de dialogue homme-machine. On sait en effet que dans une situation identique, un individu ne se comportera pas de la même façon selon qu'il croit avoir affaire à un interlocuteur humain ou à une machine (Chin, 1984). Le modèle qu'un agent se fait d'un autre agent (et notamment la représentation de ses connaissances dans un domaine particulier) va en effet orienter son comportement en ce qui concerne par exemple la forme et le contenu de ses productions langagières (voir Falzon, 1986, pour un passage en revue d'études ayant mis en évidence ce phénomène). Cette représentation des connaissances supposées de l'autre peut amener à sous-estimer les capacités de l'interlocuteur (qu'il soit humain ou artificiel) ou à les surestimer.

III. Le cadre de l'étude : la conception d'un système d'autoformation à l'utilisation de progiciels (Claes & Salembier, 1988)

L'étude présentée partiellement ici s'inscrit dans un projet industriel visant à la conception d'une gamme d'outils spécifiques d'assistance à l'utilisateur d'applications bureautiques, depuis le stade de l'initiation jusqu'à celui de l'enrichissement/réactualisation des connaissances (Claes, 1988 ; Claes et al., 1988). Notre travail, répondant en cela à la demande de l'équipe de conception, a plus spécifiquement porté sur l'activité de suivi et d'assistance à l'élève en situation d'autoformation, lors des phases de réalisation d'un exercice.

1. Méthodologie

1.1. Description de la situation d'expérimentation ergonomique

Il était demandé à un sujet de jouer le rôle d'un apprenant et de participer à ce qui lui était présenté comme l'évaluation d'un système d'assistance à l'autoformation. Après avoir répondu à un questionnaire très court portant sur les tâches réalisées habituellement et sur les outils informatiques utilisés, le sujet devait lire un document de quelques pages présentant brièvement les fonctions principales d'un logiciel de traitement de texte qui faisait l'objet de l'autoformation. L'apprenant devait ensuite effectuer un exercice d'édition et de correction de document ; cet exercice incluait des éléments identifiés comme sources possibles de difficultés lors d'une étude antérieure en situation réelle. L'apprenant disposait pour réaliser cette tâche d'un système d'assistance lui permettant de poser des questions en langage naturel. On l'informait que le système d'assistance contrôlait ses actions et qu'il pouvait intervenir spontanément et lui envoyer des messages. Le rôle du "système d'assistance" était tenu par un comparse situé dans une pièce adjacente (méthode dite du *Magicien d'Oz*⁹), et qui avait pour consigne de suivre, et d'assister l'apprenant si nécessaire. Il disposait pour ce faire d'un écran de contrôle lui permettant de visualiser les actions effectuées par l'apprenant, et d'un terminal pour la réception et l'émission de messages. La durée des sessions était conditionnée par le temps mis par le sujet pour réaliser l'exercice, le point d'arrêt étant fixé à 50 minutes.

Le rôle du comparse était joué alternativement par deux formateurs spécialisés dans l'enseignement d'applications bureautiques et possédant une grande expérience acquise dans des situations diverses (cours traditionnel, tutorat individualisé,...). Dix sujets ont participé à cette expérimentation ; ils ont été choisis en fonction de leurs connaissances initiales en informatique et de leur familiarité avec le type de tâche. L'objectif était de disposer d'un échantillon de sujets suffisamment varié.

L'ensemble des entrées clavier a été enregistré via un capteur d'événements et l'intégralité des sessions a fait l'objet d'un enregistrement vidéo qui servait ensuite de support au recueil des verbalisations de l'élève, du formateur qui avait participé à la

⁹ La technique du Magicien d'Oz (*Wizzard of Oz*) consiste à simuler le fonctionnement d'un système qui n'existe pas encore en faisant appel à un comparse qui traite les actions du sujet situé dans une autre pièce, alors que celui-ci pense interagir avec un dispositif informatique (Salembier, 1991). Le "système" ainsi simulé envoie des messages, des suggestions, ou peut modifier directement l'environnement dans lequel travaille le sujet (caractéristiques de l'interface, mode de présentation des informations, etc.).

session, et du second formateur qui n'y avait pas pris part (méthode de verbalisation consécutive assistée par les traces). Durant cette phase il était demandé au formateur de commenter la session en explicitant notamment les indices pris en compte pour juger de l'état des connaissances de l'élève, les raisons motivant une intervention et plus généralement tout élément qui lui paraissait important. Certains points verbalisés par les formateurs (notamment le rattachement du sujet à une classe d'apprenant à partir d'une évaluation locale de ses connaissances) ont fait l'objet de demandes ponctuelles d'explicitation ou de précisions supplémentaires¹⁰.

Le statut des verbalisations visant à l'extraction de connaissances expertes ou non a fait l'objet de nombreuses discussions, que ce soit en psychologie, en ergonomie ou en IA (Leplat et Hoc, 1981 ; Evans, 1988 ; Neale, 1988 ; Caverni, 1989). Nous nous contenterons ici de mettre l'accent sur certaines caractéristiques du mode de recueil utilisé dans cette étude :

- Ces verbalisations peuvent être confrontées avec des données observables, ce qui permet dans une certaine mesure, de limiter les biais de rationalisation a posteriori.
- L'utilisation d'un enregistrement de l'activité permet un recueil des verbalisations dans un contexte proche de la situation réelle (par opposition à certaines méthodes où il est demandé à l'opérateur/expert d'explicitier une expertise indépendamment de tout contexte réel de mise en oeuvre de ces connaissances).
- Le recueil croisé à partir d'un même enregistrement permet une comparaison entre ce qui est verbalisé par le formateur et l'apprenant, et par les deux formateurs. Dans le premier cas il s'agit de vérifier le niveau de validité des interprétations effectuées par le formateur ; dans le second cas de tenter de mettre en évidence les écarts entre les stratégies de diagnostic et d'assistance mises en oeuvre par les deux formateurs.

¹⁰ Notons que, à la différence de la grande majorité des paradigmes expérimentaux utilisés dans les travaux menés autour de la typicalité, les traits qui vont permettre ici une description en intension de la catégorie (dans notre cas une classe d'apprenants) ne sont pas connus et donnés a priori mais inférés à partir des verbalisations des formateurs.

1.2. Modèle d'analyse des données

Nous avons effectué une analyse essentiellement qualitative des données recueillies en nous centrant sur la construction et l'actualisation de la représentation de l'apprenant par les deux formateurs. Les protocoles verbaux ont été analysés en termes d'attribution d'états intentionnels (croyances, buts, etc.) exprimés par des attitudes propositionnelles (AP) telles que "penser que p", "vouloir p",¹¹ etc. Nous nous sommes plus spécifiquement inspirés de la taxonomie proposée par Kobsa (1986) pour classer les différents modes de modélisation d'un agent (et plus particulièrement l'utilisateur d'un système interactif) en fonction du contenu modélisé.

Kobsa distingue ainsi une quarantaine d'états intentionnels répartis en croyances et buts de base (concernant les connaissances factuelles particulières ou générales d'un agent) et croyances et buts complexes (à propos des croyances et buts d'un autre agent).

Les *croyances simples* se réfèrent à des croyances à propos de faits spécifiques ou généraux : il peut s'agir d'événements présents, passés ou à venir, de propriétés générales d'objets considérés dans une situation donnée, de relation entre objets, de règles générales.

Les *croyance complexes* se réfèrent aux croyances et aux buts d'autres agents ; elles sont, selon la terminologie employée par Dennett l'apanage des systèmes intentionnels d'ordre supérieur à un. Une des caractéristiques de ces croyances est qu'elles peuvent être emboîtées (*l'agent a croit que l'agent b croit que l'agent c croit p*) ; les systèmes qui traitent de tels emboîtements sont qualifiés de systèmes intentionnels d'ordre trois et plus. Dans le cadre de cette étude l'analyse s'est limitée aux croyances complexes d'ordre deux

¹¹ Une mise au point s'impose sans doute quant au statut que nous donnons ici à cette description de l'activité en termes d'AP. Les AP constituent ici pour nous *un intermédiaire descriptif commode* pour appréhender une dimension de l'activité des sujets et des tuteurs. Ce qui ne veut pas dire que nous adhérons forcément à une conception des processus cognitif qui serait sous-jacente à ce mode de formalisation. Nous sommes en même temps conscients des limites de cette position : bien souvent certaines options théoriques sur lesquelles s'appuie un outil de représentation finissent par "remonter à la surface" et contraindre le traitement des données.

(le tuteur *a* croit que l'élève *x* croit que *p*), et 3 (le tuteur *b* croit que le tuteur *a* croit que l'élève *x* croit que *p*)¹².

La majorité des AP utilisées pour analyser les protocoles verbaux recueillis appartiennent aux catégories croyances et buts simples saturés et croyances complexes à propos d'un agent connu. Dans plusieurs cas néanmoins, l'emploi de croyances et buts simples non saturés et de croyances complexes à propos d'un agent inconnu peut se révéler pertinent.

Les croyances simples non saturées permettent d'exprimer une croyance relative à l'existence supposée d'un fait ; par exemple une requête formulée par le sujet telle que : "Quelle est la fonction qui permet de souligner un mot ?" sera décrite par une AP de type "*agent a BEL. \$ x p(x)*" soit dans l'exemple cité plus haut "le sujet croit qu'il existe une fonction *F* telle que *F* permet de souligner un mot".

Les croyances complexes à propos d'un agent inconnu permettent d'exprimer des références faites par les tuteurs aux caractéristiques typiques des membres d'une classe particulière d'apprenants.

Extrait du protocole verbal :

Tuteur : "Bon là elle essaie de mettre en gras et elle appuie sur la touche Surimpression ; on voit souvent *ÿa* en formation, c'est souvent des gens qui ne connaissent pas Startext mais qui ont l'habitude d'utiliser une TTX."

Tuteur A BEL. \$ un apprenant *a* tel que si *a* WANTS (*mettre mot en gras*) et Action : <Surimp> alors *a* ne connaît pas Startext et *a* connaît TTX.

Pour les protocoles verbaux des formateurs, chaque attitude propositionnelle est connectée à un indice (un événement observable ou une autre attitude propositionnelle) qui a le statut d'antécédent d'une opération d'inférence, c'est-à-dire une opération de formation et de transformation des croyances. Les événements observables sont des actions ou des séquences d'actions décrites par un descripteur de plus haut niveau (macro-action) dans les termes du langage de commande de l'application ; l'absence de tout événement observable peut également avoir le statut d'indice.

¹² Nous garderons de préférence les termes anglais (bien que le résultat ne soit pas très plaisant à l'oreille...) notamment celui de *belief* qui a l'avantage, à la différence de sa traduction approchée de "croyance", de ne pas véhiculer de connotation ambiguë.

Les événements comportementaux peuvent être combinés à une série de descripteurs dont les valeurs vont caractériser les aspects de l'environnement considérés explicitement par le formateur à un moment donné (contexte physique de la survenue de l'événement). Ils peuvent également être associés à des données sur l'historique de l'interaction et à des attitudes propositionnelles sur l'état des connaissances du sujet ou sur son but courant supposé. L'association de ces différents éléments constitue des configurations d'indices que nous avons catégorisées en trois groupes (figure 2) ¹³ :

- *État du sujet* (connaissances et but du sujet).
- *Contexte restreint* (valeur prise par certains paramètres du système, moment dans la session, historique de l'interaction).
- *Contexte étendu* (contexte restreint plus état du sujet).

Figure 2. Catégorisation des inférences explicitées par les tuteurs

¹³ Cette catégorisation s'inspire d'une définition de la notion de contexte proposée par Sperber et Wilson (1989, p. 31).

L'AP peut par ailleurs être affectée d'un modalisateur exprimant une évaluation du niveau de certitude sur son contenu : "*Tuteur BEL. qu'il est possible que le sujet S1 WANTS (Sortir de l'Aide)*".

IV. Résultats

Sur 160 inférences identifiées par l'analyse des protocoles verbaux des formateurs, 125 soit 78% sont des opérations de mise en relation entre une configuration d'indices comportant au moins un événement observable de type comportemental, et une évaluation de l'état des connaissances du sujet ou une décision d'action (tableau 1). Une part importante des explicitations de leur activité produite par les tuteurs sera donc catégorisée comme acquisition et interprétation d'indices. Sur les 35 inférences restant, 28 (soit 17,5% du nombre total) décrivent des inférences permettant le passage d'un niveau local d'évaluation des connaissances du sujet (en termes de notions connues ou maîtrisées) à un niveau plus global (appartenance à une classe de sujets).

A ce niveau on ne distingue pas de différences notables entre les deux tuteurs ; on remarquera simplement que le nombre d'inférences explicitées par le tuteur B est nettement plus important (102 contre 58 pour le tuteur A).

	Tuteur A		Tuteur B		Total	
Instruction du contexte	46	79,1%	79	77,5%	125	78,1%
Sans référence à la situation	12	20,9%	23	22,5%	35	21,9%
Total	58		102		160	

Tableau 1. Répartition générale des inférences explicitées par les tuteurs A et B.

Sur les 125 inférences catégorisées dans la rubrique "instruction du contexte", 106 n'ont été évoquées qu'une fois. Sur les 28 inférences explicitées exprimant le passage d'une évaluation locale à une évaluation globale, 13 explicitations concernent des inférences qui apparaissent deux fois ou plus ; la majorité des 15 autres explicitations sont des combinaisons différentes des mêmes indices.

1. Instruction du contexte

Les résultats indiquent que pour les inférences de type "instruction du contexte" les indices traités par les tuteurs sont majoritairement des événements contextualisés ou rapportés à une évaluation de l'état du sujet (état des connaissances, but poursuivi) : 87% pour A, et 84,8% pour B contre seulement 13% pour A et 15,2% pour B d'indices comportementaux isolés (tableau 2).

La prise en compte d'actions indépendamment de toute référence au contexte ou à l'état supposé du sujet se limite à certains cas particuliers :

- utilisation par le sujet d'une fonction sophistiquée dont la mise en oeuvre ne peut être due au hasard d'une fausse manipulation ou d'une exploration du système (combinaison ésotérique de touches, accès difficile au bout d'une arborescence de menus) ;
- recours à un terme technique du domaine dans une demande d'assistance ou référence à l'existence d'une fonction spécifique ;
- indépendance de la fonction vis-à-vis des actions antérieures ; l'utilisation de cette fonction ne nécessite pas le respect de pré-requis particuliers et constitue par elle-même un indice.

	Tuteur A		Tuteur B		Total	
Evénements comportementaux observables isolés	6	13%	12	15,2%	18	14,4%
Evénements comportementaux contextualisés	40	87%	67	84,8%	107	85,6%
Total	46	100%	79	100%	125	100%

Tableau 2. Catégorisation des inférences rattachées à l'instruction du contexte

Au niveau de la catégorisation des éléments constitutifs des configurations d'indices la répartition entre les trois groupes (contexte étendu, contexte restreint, état du sujet) est à peu près équilibrée (tableau 3) ; on notera simplement dans le cas du tuteur A une légère prédominance de la prise en compte de l'état du sujet (42,5%).

	Tuteur A		Tuteur B		Total	
Etat de l'apprenant	17	42,5%	23	34,3%	40	37,4%
C.suj	1		1		2	
But suj	15		22		37	
C.suj+But suj	1		0		1	

Contexte étendu	11 27,5%	20 29,9%	31 29%
Ctx+But suj	9	18	27
Ctx+But suj+C.suj	0	2	2
Ctx+C.suj	2	0	2
Contexte restreint	12 30%	24 35,8%	36 33,6%

Tableau 3. Catégorisation des configurations d'indices (C.suj=état des connaissances du sujet ; But suj=but courant du sujet ; Ctx=contexte restreint)

La prise en compte de l'état de l'apprenant s'opère essentiellement au travers de la prise en compte du but supposé poursuivi (dans 15 cas sur 17 pour le tuteur A et dans 22 cas sur 23 pour le tuteur B) ; le diagnostic nécessite ici la vérification de l'adéquation entre but et action réalisée pour atteindre ce but, contrairement aux cas où, pour le tuteur, l'action est signifiante par elle-même indépendamment de tout autre élément (voir plus haut). Afin de juger de la correction des actions de l'apprenant, le formateur se réfère à des modèles de performance que l'on pourra décrire dans le domaine d'enseignement qui nous occupe ici par des transitions (commandes, fonctions,...) entre des états du progiciel.

On distingue parmi ces modèles de performance (cf. figure 3) :

- des modèles optimaux qui décrivent des procédures jugées parfaitement correctes ;
- des modèles sous-optimaux qui permettent d'atteindre l'objectif fixé mais qui ne sont pas jugés optimaux par le formateur, soit qu'ils mettent en oeuvre des séquences d'actions plus coûteuses en temps (chemin détourné) ou plus risquées (fonctions destructives par exemple), soit qu'ils omettent d'utiliser une fonctionnalité qui facilite la réalisation d'une manipulation ;
- des modèles erronés qui ne permettent pas d'atteindre l'objectif tel qu'il a été défini par l'énoncé de l'exercice, ou tel que l'étudiant se l'est assigné. Mais contrairement à des techniques de diagnostic purement syntaxiques

utilisées dans certains TI, le formateur ne se contente pas d'évaluer un comportement erroné en terme d'écart à une procédure optimale. Ces erreurs, ainsi d'ailleurs que certaines procédures optimales et sous-optimales sont associées à des niveaux de connaissances relatives à des notions.

Figure 3. Représentation sous forme de graphe des transitions entre états de l'application pour réaliser une opération donnée. Les séquences de transition entre états décrivent des procédures jugées optimales, sous-optimales ou erronées par le formateur dans le contexte courant.

On notera par ailleurs le rôle quasi inexistant joué à ce niveau par l'état des connaissances de l'élève, et ce quel que soit le tuteur.

Figure 4. Exemple de prise en compte de l'historique de l'interaction et verbalisation associée. L'élève devait ici réaliser une modification consistant à positionner une chaîne de caractères à un certain endroit de la page-écran ; pour ce faire il a appliqué une procédure jugée sous-optimale par le formateur et même inapplicable dans le contexte courant. A partir de cet événement le tuteur infère un état de connaissance relatif à la notion de niveau supérieur "gestion du texte à l'écran". Lors d'une seconde modification de même type, l'apprenant applique la même procédure qui cette fois donne l'effet souhaité du fait d'une modification du contexte restreint (présence d'une tabulation). Si l'on considère uniquement cette seconde phase, on pourrait en tirer une conclusion hypothétique sur le niveau de connaissance relatif à la notion incriminée. Mais si comme le fait le formateur, on réintègre les actions de l'élève dans l'historique de l'interaction où sont organisés temporellement tous les événements pertinents concernant ce type de manipulation, le résultat du diagnostic est différent.

Les configurations de type "contexte restreint" se retrouvent dans les situations :

- où l'évaluation par le tuteur d'une action du sujet nécessite la prise en compte de la valeur de certains paramètres du système (mode d'édition actif, type de format du document, position du curseur à l'écran, etc.) ; une partie importante de ces couplages entre état du système et action

renvoie à des événements typiques répertoriés par le tuteur, et auxquels sont associés une interprétation qui n'a pas à être construite localement ;
 - où la référence à l'historique de la session se révèle nécessaire¹⁴ (répétition d'un comportement, référence à une assistance antérieure afin de préciser le contenu d'un message d'aide, etc.).

Les configurations d'indices répertoriées sous la catégorie "contexte étendu" se réfèrent :

- à des cas particuliers de traitement d'une demande d'aide qui nécessitent, outre la prise en compte de l'état de l'élève (essentiellement son but), la référence à des éléments qui définissent le contexte restreint (historique et valeurs prises par certains paramètres du système notamment) ;
- à des diagnostics pour lesquels il n'est pas possible de rattacher l'action observée à un comportement répertorié ;
- à des opérations de contrôle de l'environnement permettant de vérifier si l'élève ne dispose pas d'une information présente à l'écran pour trouver une solution lui-même à son problème.

Le comptage des occurrences des différents éléments constitutifs des configurations d'indices explicitées par les formateurs (toutes catégories confondues) confirme le rôle prédominant de l'identification du but poursuivi par le sujet (67 occurrences sur un total de 151), le poids non négligeable accordé aux indices relatifs à l'état du système et à l'historique de l'interaction tutorielle (36 et 31 occurrences), et inversement la faible importance de la prise en compte de l'état des connaissances du sujet (7 occurrences seulement).

¹⁴ On rapprochera ce résultat d'un reproche qui a été fréquemment formulé à l'encontre des systèmes d'enseignement automatisé : l'absence de prise en compte de l'historique dans le traitement des réponses de l'élève (Wenger, 1987).

Figure 5. Nombre d'occurrences des différents éléments constitutifs des configurations d'indices contextualisées.

2. Le passage de l'évaluation locale des connaissances à l'évaluation globale de l'apprenant

Parmi les 35 inférences ne relevant pas, selon notre catégorisation, d'une instruction de la situation, 28 (soit 17,5% du nombre total d'inférences) expriment le passage d'un niveau d'évaluation locale des connaissances (en termes de notions connues/non connues, maîtrisées/non maîtrisées) à un niveau global. Ces inférences présentent un intérêt particulier si on les rapproche d'un aspect particulier de l'activité des tuteurs : la génération d'attentes. Le rattachement d'un apprenant à une classe spécifique d'apprenants permet en effet aux tuteurs d'inférer des informations sur ce sujet qui ne sont pas le résultat d'un processus d'instruction de la situation, mais qui proviennent des traits typiques constitutif de cette classe d'apprenants. L'activité de diagnostic des tuteurs ne se limite donc pas à la catégorisation d'indices

et à l'interprétation de ces indices en termes d'évaluation locale. Le passage à un niveau global d'évaluation permet la mise en oeuvre d'un processus de comparaison entre attentes et événements observés.

Un comportement peut ainsi être catégorisé d'une part selon qu'il est optimal, sous-optimal ou erroné et d'autre part selon qu'il est attendu ou inattendu, selon qu'il est compatible ou non avec le modèle courant des attentes (cf. extrait de protocole verbal).

Extrait de protocole verbal :

Tuteur : "Bon là elle a utilisé la sélection pour effacer la ligne c'est pas faux hein... mais elle aurait pu utiliser directement le symbole ligne Exp. : "Tu pensais qu'elle l'aurait fait ?"

Tuteur : "Ben oui comme elle a l'air visiblement de bien connaître Startext, je pensais elle doit connaître la procédure avec le symbole caché et qu'elle allait s'en servir là... mais bon c'est pas très important c'est pas faux quoi..."

Dans une optique de modélisation de l'activité de diagnostic cognitif une façon pratique de rendre compte de ces classes d'apprenants consiste à les représenter par des structures schématiques de type stéréotype qui sont définies comme des ensembles de traits dont les valeurs vont caractériser un apprenant prototypique (Claes et Salembier, 1988). Cet apprenant prototypique est donc vu ici comme une entité construite d'attributs typiques. Ces stéréotypes s'intègrent dans un réseau hiérarchisé d'autres classes stéréotypiques ; le stéréotype "Utilisateur d'éditeur de textes" par exemple est une sous-classe d'une catégorie de niveau supérieur "Utilisateur de systèmes informatiques" dont il hérite de l'ensemble des traits. Au sommet de cette hiérarchie on trouve un stéréotype plus général (le stéréotype "Elève") qui renvoie à une conception globale des caractéristiques communes à tout apprenant du point de vue du formateur ("souhaite que l'on réponde aux questions qu'il pose" ; "ne souhaite pas être interrompu systématiquement lors d'une phase pratique" ; "utilise les connaissances dont il dispose déjà").

3. La gestion dynamique des stéréotypes

La phase d'instruction du contexte par le tuteur peut donner lieu à l'émergence d'incohérences apparentes (activation de plusieurs stéréotypes contradictoires, écart entre indices non attendus et modèle courant de l'apprenant). C'est pourquoi les informations contenues dans le modèle de l'apprenant sont susceptibles d'être modifiées par l'occurrence d'un événement nouveau qui peut avoir valeur de contre-évidence pour une valeur d'attribut.

Exemple : le stéréotype activé donne la valeur par défaut "connue" pour la notion X. L'élève effectue une opération interprétée comme une mise en oeuvre incorrecte de cette notion ; le formateur en déduit qu'il ne connaît pas cette notion, d'où incohérence avec la valeur courante du modèle élève. Dans ces conditions plusieurs cas de figure peuvent apparaître :

- Le trait incriminé (notion X : connue) n'est pas déterminant pour l'attribution du stéréotype et la valeur est corrigée ou laissée en suspens dans l'attente d'une confirmation ou d'une infirmation.
- Le trait est déterminant (sa vérification est nécessaire à l'attribution du stéréotype) et le changement de valeur implique l'activation d'un stéréotype différent ; selon le niveau de certitude globale attaché au stéréotype activé (le niveau de certitude peut être élevé si l'on se trouve à un moment avancé de l'interaction et qu'une majorité significative d'événements sont venus conforter ce choix, ou faible si l'on se trouve au début d'une session d'enseignement et que plusieurs stéréotypes candidats sont activés) on pourra avoir sélection d'un nouveau stéréotype candidat, ou production d'une interprétation qui tente de donner un sens à cette incohérence apparente afin de sauvegarder (éventuellement provisoirement) le stéréotype activé.

Dans certains cas (notamment en début de session, quand la représentation de l'apprenant n'est pas encore solidement fixée) on peut même avoir co-existence temporaire de plusieurs stéréotypes candidats contradictoires jusqu'à ce qu'une suite d'événements significatifs permettent de faire un choix.

Le formateur utilise ainsi plusieurs mécanismes et exploite plusieurs caractéristiques de ses modèles stéréotypiques afin de gérer les conflits éventuels entre stéréotypes et les incohérences entre indices et représentation courante de l'élève :

- Un mécanisme d'attribution hypothétique et non-monotone des valeurs d'attributs du modèle ; l'évaluation d'un niveau de connaissances reste souvent largement hypothétique et est susceptible de changer avec le temps, à la fois pour tenir compte de l'apprentissage et de la maîtrise par l'apprenant de nouvelles notions (ce qui reste l'objectif de l'interaction tutorielle) et pour refléter l'affinement progressif de la modélisation.
- La non-équivalence entre traits typiques ; les traits qui définissent l'appartenance à une classe stéréotypique d'apprenants n'ont en effet pas le même statut : certains sont nécessairement associés à cette classe, d'autres le sont de façon non-obligatoire.
- Une certaine inertie de la représentation comme garant de la pertinence des prises de décision (interventions notamment). Afin de minimiser les risques d'interruptions intempestives dues à des erreurs de diagnostic, le formateur adopte une stratégie conservatrice qui le conduit à différer une décision de révision du modèle qui entraînerait une intervention éventuelle, tant que le niveau de certitude attaché à son modèle de l'élève est insuffisant, ou lorsque qu'il y a incohérence manifeste entre le comportement de l'élève et les prédictions du modèle. Cette inertie permet de traiter correctement une erreur de réalisation (*slip*) et de ne pas l'interpréter comme révélatrice d'une conception erronée ou d'une lacune.

4. Evaluation globale et exactitude du diagnostic

Les résultats mettent en évidence l'efficacité des inférences permettant de passer d'un niveau d'évaluation locale des connaissances du sujet à un niveau global en termes d'appartenance à une classe prototypique d'apprenants (diagnostic correct pour 8 sujets sur 10 pour le tuteur A et diagnostic correct pour 9 sujets sur 10 pour le tuteur B).

Dans deux cas (même sujet), l'erreur de diagnostic des tuteurs ("le sujet n'a pas l'habitude de travailler sur un traitement de texte") est imputable à la mauvaise appréciation par le sujet de son niveau de compétence réel : le sujet travaille en effet quotidiennement sur un

traitement de texte (réponse au questionnaire) mais l'analyse de ses verbalisations montre qu'il n'en exploite pas toutes les possibilités. On peut donc en fait dire que le diagnostic posé par les tuteurs sur le niveau de compétence manifesté par le sujet au travers de ses actions correspond à une certaine réalité même si celle-ci ne correspond pas à celle exprimée par le sujet dans le questionnaire.

Dans le troisième cas, l'erreur de diagnostic commise par le tuteur A peut être expliquée par le faible nombre de données disponibles. Le sujet a en effet rapidement décidé de mettre fin à l'expérimentation pour cause de non conformité du comportement du système d'aide à ses attentes.

	Tuteur A	Tuteur B	Total
Cohérence générale	8	7	15
Concordance stricte	2	2	4
Modification	0	1	1

Tableau 4. Ecart entre le premier diagnostic posé et l'évaluation finale. Deux cas de concordance stricte (même sujet) concernent une situation où un indice (utilisation d'une fonction évoluée) a été associé à la maîtrise d'une notion importante, ce qui a conduit les tuteurs à catégoriser avec certitude l'apprenant dans la classe des sujets qui connaissent et maîtrisent Startext. Dans les deux autres cas (même sujet) la session a été interrompue rapidement à l'initiative du sujet ; les tuteurs n'ont donc pas pu disposer de suffisamment d'éléments pour affiner ou modifier leur diagnostic initial qui, ne reposant pas sur des indices déterminants, manquait d'assurance.

On remarquera par ailleurs la forte cohérence entre la nature du premier diagnostic posé par les tuteurs et l'évaluation finale du sujet (tableau 4). Cette cohérence ne signifie pas systématiquement correspondance stricte. Dans la majorité des cas (15 cas sur 20), si l'évaluation finale ne vient pas contredire le premier diagnostic, elle le précise notablement par affinement progressif (diagnostic initial : "elle a l'habitude de ce genre de système informatique", diagnostic final : "elle connaît le traitement de texte, mais pas Startext") ou par imputation

d'une certitude plus grande (diagnostic initial : *"c'est probable qu'elle ne connaît pas le traitement de texte"*, diagnostic final : *"elle ne connaît pas le traitement de texte, c'est certain"*).

Le seul cas où l'on peut observer une différence nette entre le diagnostic initial et l'évaluation finale peut être interprété par référence au rôle joué par la construction par le sujet d'un modèle du système d'assistance.

Figure 6. Effet de la représentation de l'apprenant à propos du système d'assistance sur l'évaluation des connaissances de cet apprenant par le tuteur (en italique : extrait du protocole verbal ; en standard : messages échangés entre le tuteur et le sujet).

On est ici en présence d'une manifestation du décalage entre le modèle que le tuteur se fait du sujet et le modèle que le sujet se fait du système d'assistance. Ceci conduit chacun des deux protagonistes à simplifier

progressivement le contenu de ses messages pour tenter de s'adapter au modèle qu'il se fait de son interlocuteur. Ce phénomène qui a été précédemment décrit dans la littérature sous le nom d'*effet de poursuite* (*hunting effect*) peut donc jouer un rôle direct sur l'évaluation des connaissances de l'apprenant et conduire le tuteur à produire un diagnostic erroné.

V. Discussion

Parmi les résultats de cette étude, certains points nous semblent devoir être soulignés.

(i) *Le caractère contextualisé des indices traités par les tuteurs.* La majorité des événements comportementaux traités sont intégrés dans des configurations d'indices comprenant des éléments de description du contexte et des références aux buts et connaissances des sujets. Ceci peut amener à s'interroger sur le domaine de validité des systèmes d'assistance qui se basent uniquement sur la détection d'événements ou de séquences limitées d'événements indépendamment de toute référence au contexte. Les résultats montrent par ailleurs que, s'il est possible, à partir des explicitations des tuteurs, d'identifier un noyau relativement restreint d'inférences permettant le passage d'une évaluation locale à une évaluation globale des connaissances de l'apprenant, le nombre d'inférences récurrentes relatives à l'instruction du contexte est proportionnellement beaucoup moins important. Ceci est dû en grande partie à la variété des configurations d'indices traitées qui se singularisent souvent par la prise en compte d'un élément contextuel particulier. On peut donc penser que si la formalisation du processus "d'évaluation globale" en vue d'une éventuelle automatisation semble envisageable, la formalisation de "l'évaluation locale par instruction du contexte" apparaît

comme plus problématique en raison du caractère essentiellement situé de cette phase de diagnostic¹⁵.

(ii) *Le rôle fonctionnel des stéréotypes.* Dans une perspective de fondation des processus cognitifs sur une base de “rationalité immanente”, l'existence dans le raisonnement de structures cognitives simplificatrices telles que les stéréotypes a pu être expliquée comme une sorte de raté explicable uniquement par des facteurs d'ordre socioculturel ou motivationnel.

Dans les situations complexes où les données d'un problème à résoudre sont rarement toutes disponibles, les stéréotypes ne peuvent être réduits à un ensemble de croyances simplificatrices ou à la manifestation d'une rationalité imparfaite : ils permettent d'une part de gérer la complexité de l'environnement de manière à le rendre plus "praticable", et ils contribuent également à augmenter le nombre des éléments pertinents susceptibles d'être pris en compte dans le processus de résolution. On peut penser que cette capacité d'évocation et de gestion des stéréotypes est directement liée à la pratique professionnelle du diagnostic des difficultés d'apprentissage et de l'assistance aux apprenants. C'est en effet la multiplication des expériences dans différentes situations réelles de formation, l'imputation de régularités causales découvertes à partir de ces expériences et le test de la validité des inférences réalisées qui seraient à l'origine de la construction empirique, puis de la stabilisation de ces structures schématiques. A ce titre elles constitueraient un mode d'organisation des connaissances, structurées par et pour une pratique pédagogique. Il serait de ce point de vue intéressant de comparer l'activité de formateurs expérimentés avec celle de sujets possédant une bonne connaissance du domaine mais sans aucune pratique de l'enseignement.

¹⁵ Si l'on considère par exemple la prise en compte de l'historique de l'interaction, comment déterminer les événements dont il faut garder une trace ?

Figure 7. Espace de prédiction des comportements de l'apprenant. Cet espace est partitionné en espaces emboîtés qui vont définir une certaine graduation dans le niveau des attentes¹⁶. L'espace des possibles se réfère à l'ensemble des événements compatibles avec le modèle de l'élève en général propre au formateur ; l'espace des plausibles définit l'ensemble des événements prédits par les traits typiques non nécessaires d'un stéréotype, et l'espace des probables les événements prédits par la représentation courante de l'apprenant (une version actualisée, contextualisée d'un ou de plusieurs stéréotypes sur la base des données de l'interaction apprenant/tuteur). Enfin la survenue de certains événements est jugée nécessaire pour confirmer l'hypothèse courante d'appartenance à une classe stéréotypique (prédits par les traits nécessaires du stéréotype). La définition de ces espaces est propre à chaque formateur, et les attentes générées sont donc susceptibles de varier.

(iii) *Le statut des traits typiques décrivant les classes stéréotypiques.*
On retrouve à ce niveau un problème classique discuté dans le cadre des

¹⁶ Les termes employés ici pour dénommer ces espaces de prédictions sont semblables à ceux utilisés par Dubois et Prade (1985) ; mais l'emploi que nous en faisons revêt une valeur essentiellement heuristique et ne s'appuie pas, comme c'est le cas pour Dubois et Prade, sur une définition mathématique.

travaux sur la typicalité et que l'on pourra sommairement résumer de la façon suivante : les traits qui décrivent une catégorie déterminent-ils un faisceau de conditions nécessaires et suffisantes (modèle des CNS), ou bien peut-on dire qu'aucun de ces traits n'est nécessairement partagé par tous les membres d'une catégorie (version standard du modèle roschien ; cf. Dubois, 1991). Certains auteurs se sont efforcés de trouver une voie médiane qui tient compte de l'existence de traits nécessaires (oiseau pour canari par exemple) mais non suffisants. Ainsi Putnam (1975) utilise-t-il la notion de marqueur sémantique et Lakoff établit-il une distinction entre critère d'appartenance et propriété typique (Lakoff, 1987). On retrouve ce type de distinction dans le cas qui nous occupe ici : certains traits sont nécessaires pour le rattachement d'un apprenant à une classe stéréotypique, mais ils ne sont pas suffisants. D'un autre côté certains traits permettent de poser une hypothèse mais ne constituent pas une preuve déterminante. Le statut de ces traits jouerait alors directement sur la prédiction des comportements de l'apprenant.

(iv) *Le problème de la généralisation des résultats.* Les résultats discutés dans cette présentation partielle de l'étude que nous avons menée sur l'activité de diagnostic cognitif présentent un certain nombre de similitudes avec ceux observés dans situations proches. Ainsi dans le cadre des situations de dialogue de diagnostic, certains éléments notés par Cahour (1990) sont très proches de ceux que nous avons pu observer. Ainsi la distinction introduite par l'auteur entre "évaluation de degré 1" et "évaluation de degré 2" pourra être rapprochée de celle que nous avons décrite entre diagnostic par instruction du contexte et évaluation globale. On notera néanmoins certaines différences notamment sur la nature des indices traités pour réaliser le premier niveau d'évaluation des connaissances du sujet : dans l'étude de Cahour les experts se fondent essentiellement sur les caractéristiques pragmatico-sémantiques du discours de l'interlocuteur alors que dans la situation que

nous avons étudié les tuteurs se basent surtout sur des événements comportementaux non langagiers. Ceci peut être expliqué par la différence de nature entre les deux situations (dialogue de consultation vs dialogue tutoriel) et par la spécificité des objectifs poursuivis par le type d'experts étudiés (diagnostic et résolution d'un problème soumis par l'interlocuteur vs diagnostic de l'état des connaissances d'un apprenant et aide à la résolution).

Plus largement, l'activité de diagnostic cognitif que nous avons étudiée présente des similitudes avec d'autres activités de diagnostic (dans les situations de contrôle de processus continu notamment). On retrouve ainsi certaines des phases (acquisition et interprétation d'indices, génération d'hypothèses) décrites par Hoc dans sa typologie. Une des différences essentielles entre l'activité de diagnostic cognitif en situation d'interaction tutorielle et d'autres activités de diagnostic tient au fait que dans le premier type de situation le “processus” qui fait l'objet du diagnostic va lui même se construire une représentation du tuteur (humain ou artificiel) avec qui il interagit, ce qui va éventuellement avoir des conséquences sur l'activité de ce tuteur (cf. exemple donné plus haut).

Cet aspect conduit à s'interroger sur la conception des systèmes interactifs qui intègrent un modèle de l'agent (interlocuteur, utilisateur, élève). En effet ces systèmes dans leur vocation à s'adapter aussi précisément que possible à un interlocuteur, s'inscrivent souvent dans une perspective de simulation du dialogue homme-homme. Mais comme le rappelle justement Rastier (1989, p. 133) cet objectif demeure illusoire car il procède d'une conception unilatérale du dialogue qui fausse les conditions d'une simulation, dans la mesure où la représentation du système par l'interlocuteur n'est pas prise en compte.

VI. Conclusion

En conclusion, nous souhaiterions aborder le problème de l'évolution du statut des études menées dans les situations de travail et de leur place dans la recherche cognitive actuelle. Si la confusion plus ou moins sciemment entretenue entre "terrain" et "champ d'application" a fini par succomber (du moins peut-on l'espérer) devant la critique (voir Leplat 1982, par exemple) il reste que les situations professionnelles demeurent simplement pour certains des lieux dont les caractéristiques intrinsèques spécifiques (de complexité notamment) permettent d'étudier des problèmes originaux impossibles à reproduire en laboratoire. De la même façon, leur rôle se réduirait à interroger des disciplines jugées plus "fondamentales" et habituées à exercer dans un cadre plus proche des canons académiques de la science.

Cette négation de la capacité des études de terrain à générer un corpus de connaissances originales sur les activités cognitives pose problème. Il est en effet permis de penser que la complexité de ces situations et les résultats qui en sont issus peuvent contribuer de manière sensible à enrichir l'analyse de phénomènes simplifiés et réduits à l'extrême (Dubois, 1992) et par là même à régénérer certains aspects des recherches menées en sciences cognitives.

Dans le même esprit, on peut rattacher la problématique d'étude des activités cognitives dans les situations complexes naturelles d'un courant plus large qui pose le problème de la remise en question du bien-fondé de l'adoption d'une certaine perspective rationaliste dans l'étude de la cognition. Récemment cette interrogation a connu dans le champ d'études qui nous concerne ici, l'amorce d'un nouveau développement plus radical sous la forme d'une proposition de glissement paradigmatique : suite à un examen critique des apports concrets à la conception dans le domaine de l'ergonomie, certains auteurs (De Greene, 1990) ont prôné une distanciation à la référence obligée au paradigme

newtonien en sciences humaines, arguant de la faillite du rationalisme à rendre compte des phénomènes observés dans les situations complexes. Mais la recherche de solutions alternatives ne s'est pas encore traduite semble-t-il par la mise en évidence de réponses satisfaisantes. Certaines voies font actuellement l'objet d'analyses exploratoires, en particulier les modèles causaux non linéaires ou chaotiques, qui ont déjà été appliqués aux systèmes physiques, mais dont l'utilisation en sciences cognitives est encore à ses débuts. Ces modèles peuvent en effet fournir un cadre global pour l'interprétation et la généralisation de résultats (en permettant notamment une certaine relativisation quant aux conclusions qu'il est possible d'en tirer), mais la validité d'une application des formalismes utilisés dans ces travaux aux domaines de la psychologie et de l'ergonomie cognitive reste à démontrer (Pavard, 1991).

Pascal SALAMBIER

G.I.C.

Laboratoire d'Ergonomie et Neurosciences appliquées au Travail

CNAM

41, rue Gay-Lussac, 75005 Paris

N.B. Je remercie Danièle Dubois, Gérard Claes et Jean-Michel Hoc pour leurs commentaires et remarques ; ils m'excuseront de ne pas en avoir intégralement tenu compte...

Bibliographie

AARONSON, A. & CARROLL, J.M. (1987) Intelligent help system in a one-shot dialog : a protocol study, *Proceedings of the CHI+GI'87 Conference*.

ABBOTT, K.H.(1988) Robust diagnosis as problem solving in hypothesis space, *Proceedings of the 7th AAAI conference*.

ALTY, J. & COOMBS, M.J. (1981) Communicating with university computer users : a case study, in M.J. Coombs & J. Alty (Eds.) : *Computing skills and the user interface*, Academic Press, Londres.

- ALTY, J., ELZER, P., HOLST, O., JOHANNSEN, G. & SAVORY, S. (1985) *Literature and user survey of issues related to man-machine interfaces for supervision and control system*, Esprit P600 Graphical Dialogue Pilot Phase Report.
- BALACHEFF, N. (1991) Contribution de la didactique et de l'épistémologie aux recherches en EIAO, *Actes des XIIIe Journées francophones sur l'informatique, Grenoble*.
- BAINBRIDGE, L. (1981) Le contrôleur de processus, *Bulletin de Psychologie*, 34, 813-832.
- CAHOUR, B. (1988) *Les dialogues de consultation : vers un modèle pour évaluer la compétence de l'interlocuteur*, Rapport Inria-Rocquencourt n°931.
- CAHOUR, B. (1990) Systèmes d'aide et modélisation de l'utilisateur : apport de l'étude des situations naturelles de consultation, *TSI*, 9, 5.
- CAVERNI, J.P. (1989) La verbalisation comme source d'observables pour l'étude du fonctionnement cognitif, in J.-P. Caverni, C. Bastien, P. Mendelsohn & G. Tiberghien (éds.) : *Psychologie cognitive : modèles et méthodes*, P.U.G., Grenoble.
- CHERNIAK, C. (1986) *Minimal Rationality*, MIT Press, Cambridge MA.
- CHIN, D. (1984) An analysis of scripts generated in writing between users and computers consultants, *AFIPS conference Proceedings*, 53, 637-642.
- CHIN, D.N. (1986) User modeling in UC, the UNIX consultant, *Proceedings of the CHI'86 Human Factors in Computing Systems Conference, Boston*.
- CLAES, G., OUNIS, O., RAZOARIVELO, Z., SALEMBIER, P. & SRIDHARAN, M.S. (1988) Starguide : a generator for self-tutorials, *RIAO'88, Boston, MIT*.
- CLAES, G. ET SALEMBIER, P. (1988) Analyse de l'activité tutorielle et représentation des connaissances dans un système d'autoformation, *Congrès ERGOIA'88, Biarritz*.
- CLANCEY, W.J. (1985) Heuristic Classification, *Artificial Intelligence*, 27, 289-350.
- COLLINS, A., WARNOCK, E.H. & PASSAFIUME J.J. (1975) Analysis and synthesis of tutorial dialogues, in G. Bower (ed.) : *The Psychology of Learning and Motivation* (vol. IX), New York : Academic Press.
- DE GREENE, K. B. (1990) Contextual aspects of Human factors : the case for paradigm shift, *Human Factors Society Bulletin*, 33, 9.
- DE MEDEIROS, E. (1991) Troubleshooting strategies - the constraints imposed by the telephon, *HCI International Conference'91, Stuttgart*.
- DENNETT, D.C. (1987) *The Intentional Stance*, MIT Press, Cambridge (Mass.).
- DUBOIS, D. (1991) *Sémantique et cognition ; catégories, prototypes et typicalité*, Paris, Editions du CNRS.
- DUBOIS, D. (1992) Sciences cognitives et ergonomie : effet de mode, enrichissement réciproque ou normalisation ?, *Le Travail Humain*.
- DUBOIS, D. ET PRADE, H. (1985) *Théorie des possibilités*, Masson, Paris.
- FALZON, P. (1986) *Langages opératifs et compréhension opérative*, Thèse de doctorat, Université Paris V.

- FALZON, P. (1987) *Les dialogues de diagnostic : l'évaluation des connaissances de l'interlocuteur*, Rapport n° 747, INRIA, Rocquencourt.
- FORD, L. (1986) *Lessons from TUTOR*, Research Report R132, Computer Science Department, University of Exeter.
- GOPHER, D. & KIMCHI, R. (1989) Engineering psychology, *Annual Review of Psychology*, 40, 431-453.
- GÖRANSSON, B., LIND, M., PETTERSSON, E., SANDBLAD, B. & SCHWALBE, P. (1987) The interface is often not the problem, *Proceedings of the CHI+GI'87 conference*.
- HALFF, H.M. (1988) Curriculum and instruction in automated tutors, in M.C. Polson & J.J. Richardson (eds) : *Foundations of Intelligent Tutoring Systems*, LEA, Hillsdale, New Jersey.
- HILL, W.C. & MILLER, J.R. (1988) Justified advice : a semi-naturalistic study of advisory strategies, *Proceedings of the CHI+GI'88 conference*.
- HOC, J.M. (1989) Strategies in controlling a continuous process with long response latencies : needs for computer support to diagnosis, *Int. J. of Man-machine Studies*, 30, 47-67.
- HOC, J.M. (1990a) Psychologie et ergonomie cognitive du raisonnement, *CIRILLE, 4eme Université d'été "Méthodes et raisonnement en Intelligence artificielle*, Lyon.
- HOC, J.M. (1990b) Les activités de diagnostic, in J.-F. Richard, C. Bonnet & R. Ghiglione (éds.) : *Traité de Psychologie cognitive*, tome 2, Dunod, Paris.
- JERRAMS-SMITH, J. (1989) An attempt to incorporate expertise about users into an intelligent interface for Unix, *Int. J. of Man-machine Studies*, 31, 269-292.
- KASS, R. & FININ, T. (1988) A general user modelling facility, *Proceedings of the CHI+GI'88 Conference*.
- KOBSA, A. (1986) A taxonomy of beliefs and goals for user models in dialog systems, *Proceedings of UM86, the first international workshop on user modeling*, Maria Laach, RFA.
- LAKOFF, G. (1987) *Women, Fire and Dangerous Things*, The University of Chicago Press, Chicago.
- LEINHARDT, G. & GREENO, J.G. (1986) The cognitive skill of teaching, *Journal of educational Psychology*.
- LEPLAT, J. (1982) Le terrain, stimulant ou obstacle au développement de la psychologie cognitive, *Cahiers de Psychologie Cognitive*, 2, 115-130.
- LEPLAT, J. & HOC, J.M. (1981) Subsequent verbalization in the study of cognitive processes, *Ergonomics*, 24, 743-755.
- LITTMAN, D.C., PINTO, J. & SOLLOWAY, E. (1986) Ana analysis of tutorial reasoning about programming bugs, *Proceedings of the AAAI86 Conference*.

- MCKENDREE, J. & CARROLL, J.M. (1986) Advising roles of a computer consultant, *Proceedings of the CHI'86 Conference*.
- NORMAN, D. (1986) Cognitive engineering, in D. Norman & S.W. Draper : *User Centered System Design, New perspectives on Human Computer Interaction*, Hillsdale, N.J. : Lawrence Erlbaum Associates.
- OHLSSON, S. (1987) Some principles of intelligent tutoring, in R.W. Lawler & M. Yazdani (éds) : *Artificial Intelligence and Education*, Vol.1, Ablex, Norwood.
- PAVARD, B. (1991) Quel paradigme utiliser pour étudier les systèmes complexes ?, *Proceedings of the XXVIIIe Congress of the International Ergonomics Association*, Paris.
- PAVARD, B. (1987) *Psycholinguistique du traitement de texte*, Thèse d'habilitation, Université de Paris VIII.
- PUTNAM, H. (1975) *Mind, Language and Reality*, Philosophical Papers, vol. 2, Cambridge University Press, Cambridge MA.
- RASMUSSEN, J., (1987) Cognitive engineering, *Proceedings of the 2nd IFIP Conference on Human Computer Interaction - Interact'87*, Stuttgart.
- RASTIER, F. (1989) Dialogue homme-machine et représentation de l'interlocuteur, in *L'interaction*, Paris, ASL.
- RICH, E. (1979) Modeling via stereotypes, *Cognitive Science*, 3, 329-354.
- ROSENBERG, R. (1987) A critical analysis of research on Intelligent Tutoring Systems, *Educational Technology*, November, 7-13.
- SALEMBIER, P. (1990) Les études préalables à la conception de systèmes d'enseignement automatisé, *TSI*, 9, 5.
- SALEMBIER, P. (1991) Le Magicien d'Oz : une technique d'analyse et d'investigation des situations d'interaction avec un système "intelligent", *Proceedings of the XXVIIIe congress of the International Ergonomics Association*, Paris.
- SPERBER, D. & WILSON, J. (1986) *Relevance*, Basil Blackwell, Oxford. [Traduction française, 1989, Editions de Minuit, Paris].
- STEVENS, A.L. & COLLINS, A. (1977) The goal structure of a Socratic tutor, *Proceedings of ACM national conference*, Seattle, Washington.
- VAN LEHN, K. (1988) Student Modelling, in M.C. Polson & J.J. Richardson (éds) : *Foundations of Intelligent Tutoring Systems*, LEA, Hillsdale, New Jersey.
- WENGER, E. (1987) *Artificial intelligence and tutoring systems*, Morgan Kaufmann Publishers, Los Altos.
- WOODS, D.D. & ROTH, E.M. (1988) Cognitive Engineering, in M. Helander (ed.) : *Handbook of Human Computer Interaction*, Elsevier Science Publishers B.V. (North Holland).
- WOODS, D.D. (sous presse) Process tracing methods for the study of cognition outside of the experimental psychology laboratory, in G. Klein, R. Calderwood & J. Orasanu (Eds.) : *Decision making in action : models and methods*, Ablex, Norwood NJ.

- WOODS, D.D., ROTH, E.M. & BENNETT, K. (1987) Explorations in joint human-machine cognitive systems, in W. Zachary & S. Robertson (éds.) : *Cognition, Computing and Cooperation*, Ablex, Norwood, N.J.
- WOODS, D.D., ROTH, E.M. & POPE, H. (1987) *Cognitive environment simulation : an artificial intelligence system for human performance assessment*, NUREG-CR-4862 report.