

# MODELISATION DE L'ACTION

## 3.1. Introduction

Du point de vue de l'informaticien concepteur, le Modèle du Processeur Humain définit un canevas théorique utile à la compréhension et à l'évaluation quantitative des mécanismes généraux de l'interaction homme-ordinateur. GOMS et Keystroke remplissent un point précis du canevas pour le cas particulier de la modélisation du comportement de l'utilisateur expert engagé dans des tâches de routine. Le point de vue behavioriste de ces modèles doit être complété avec la perspective cognitive : GOMS et Keystroke modélisent le comportement observé ; l'objet de ce chapitre, "vers une théorie de l'action", est d'analyser les processus psychologiques qui conduisent à ce comportement.

Le point de départ de la théorie de l'action de D. Norman repose sur l'hypothèse que l'individu élabore des modèles conceptuels et que ces modèles sont les données déterminantes du comportement [Norman 86]. Nous verrons que cette hypothèse n'est pas en contradiction avec celle de H. Simon exposée au chapitre précédent, selon laquelle c'est l'environnement qui détermine le comportement. L'organisation de ce chapitre suit les deux éléments principaux de l'hypothèse de D. Norman : la première partie introduit la notion de modèle conceptuel ; la seconde décrit la structure du processus de manipulation du modèle. Ce chapitre s'achève par une évaluation.

## **3.2. Modèle conceptuel**

Un modèle conceptuel est une représentation mentale. Il dépend étroitement de la connaissance déjà acquise et de la compréhension de la situation présente. Il évolue avec l'expérience. Il est incomplet et imprécis mais il guide l'essentiel du comportement.

On distingue deux formes de modèle : le modèle de conception et le modèle de l'utilisateur.

1) Le Modèle de Conception est le modèle conceptuel de l'outil. Puisque la raison d'être d'un outil est d'aider l'utilisateur à accomplir un ensemble de tâches, le modèle de conception doit résulter d'une étude approfondie des besoins, des possibilités et des limitations de l'utilisateur type.

2) Le Modèle de l'Utilisateur est la représentation mentale que l'utilisateur élabore à propos de l'outil. Il résulte de l'interprétation que l'utilisateur fait de l'image.

L'image d'un outil est sa présentation physique, son interface d'utilisation. Le concepteur a donc pour tâche de définir une image qui conduise l'utilisateur à construire, au cours de l'interaction avec l'outil, un modèle compatible avec le modèle de conception. Si l'image est explicite, cohérente et intelligible, alors on peut espérer que l'utilisateur élabore le modèle adéquat

Ces définitions et remarques s'appliquent à tout outil, à l'ordinateur en particulier. Dans le cas de l'ordinateur, un troisième type de modèle peut intervenir : le modèle qu'un programme "intelligent" élabore pour représenter l'utilisateur. Ce modèle sert de base aux interfaces capables d'évoluer dynamiquement en fonction des caractéristiques et de l'état mental actuel de l'individu.

La figure 3.1 situe les diverses classes de modèles conceptuels et montre le rôle central de l'image, passerelle entre le monde physique du système informatique et le monde psychologique de l'utilisateur. Chacun des deux mondes s'exprime dans un langage spécifique. L'image assure le passage entre les deux langages. L'expérience montre que ce passage n'est pas

toujours facile à franchir même dans les cas les plus simples de la vie courante.

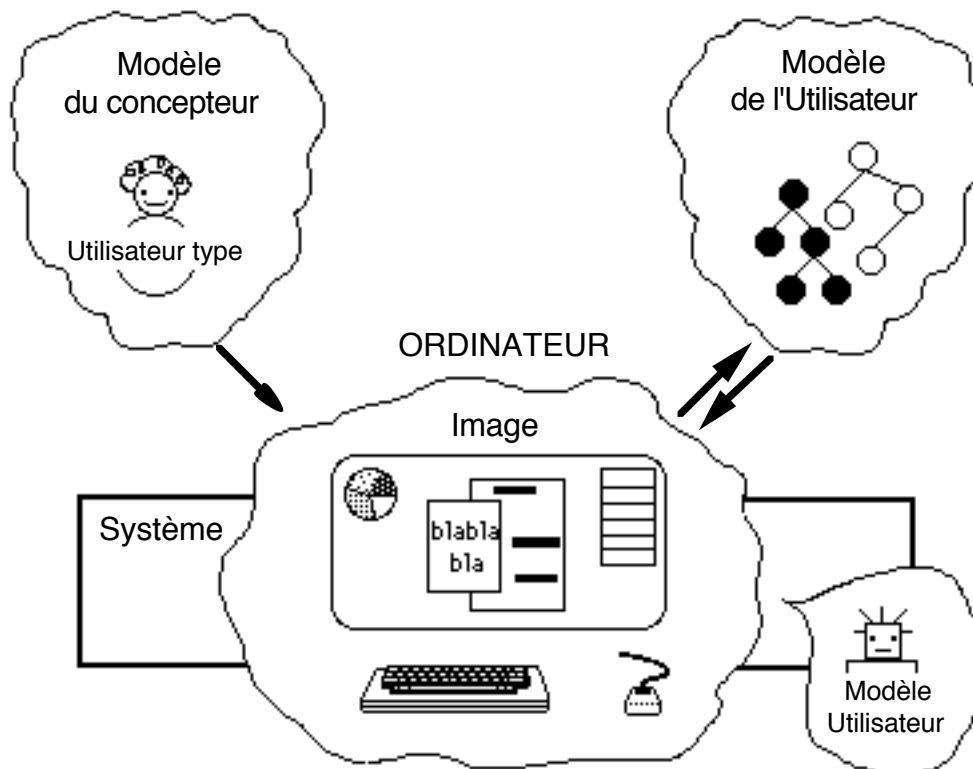


Fig. 3.1 : Modèles conceptuels et notion d'image.

Prenons comme exemple une illustration inspirée de [Norman 86]. La tâche considérée consiste à remplir une baignoire équipée de deux robinets indépendants, l'un pour l'eau chaude, l'autre pour l'eau froide. Du point de vue de l'utilisateur :

- les éléments signifiants comprennent :

- d, le débit total de l'eau et
- t, la température du bain.

d et t sont les variables psychologiques qui interviennent dans l'expression du but visé.

- l'outil "baignoire à 2 robinets indépendants" se manifeste par quatre variables physiques :

- $d_c$ , le débit du robinet d'eau chaude,

- $t_c$ , la température du robinet d'eau chaude,
- $d_f$ , le débit du robinet d'eau froide,
- $t_f$ , la température du robinet d'eau froide.

- le robinet d'eau chaude sert à contrôler  $d_c$  et le robinet d'eau froide permet d'agir sur  $d_f$ . Les robinets sont des dispositifs physiques de commande pour agir sur les variables physiques  $d_c$  et  $d_f$ .
- les équations (1) et (2) expriment les relations entre les variables psychologiques et les variables physiques :

$$d = d_f + d_c \quad (1)$$

$$t = (d_c t_c + d_f t_f) / (d_f + d_c) \quad (2)$$

La formalisation de la tâche "remplir la baignoire" met en évidence plusieurs difficultés pour l'utilisateur :

1) Un problème de correspondance entre les variables physiques et les dispositifs physiques de commande. Dans sa représentation mentale, l'utilisateur sait, en raison de son expérience passée, que  $d_c$  et  $d_f$  sont contrôlables avec les robinets. Deux questions se posent alors pour réaliser ce contrôle. Comment, dans l'image, distinguer le robinet d'eau froide du robinet d'eau chaude? Dans quel sens faut-il tourner un robinet pour agir sur le débit ?

2) L'inadéquation des relations entre les variables physiques et les variables psychologiques. Les équations (1) et (2) montrent que pour refroidir la température du bain tout en conservant le même débit, il faut simultanément augmenter  $d_f$  (pour faire plus froid) et diminuer  $d_c$  (pour conserver  $d$  constant) ou bien diminuer  $d_c$  (pour faire moins chaud) et augmenter  $d_f$  (pour conserver  $d$  constant). Quel que soit le raisonnement choisi, l'utilisateur se voit dans une situation difficile où il lui faut manipuler simultanément deux robinets dans des sens contraires. Le but est difficile à atteindre en raison de l'inadéquation des relations entre les variables physiques et psychologiques.

3) Un problème d'évaluation des résultats.

Avec deux becs verseurs, il est difficile d'évaluer le débit total. Il est encore plus difficile (voire sensoriellement douloureux) de déterminer la température obtenue! Cette analyse indique qu'un robinet mélangeur résoud le problème de l'évaluation et qu'un robinet thermostatique définit un isomorphisme entre les variables physiques et les variables psychologiques.

Nous venons de mentionner les notions de variables et les activités de contrôle et d'évaluation. Ces notions et ces activités forment un sous-ensemble des aspects d'une tâche

### **3.3. Les aspects d'une tâche**

D. Norman dans [Norman 86] indique que la réalisation d'une tâche met en jeu au moins sept activités :

- 1) L'établissement d'un but
- 2) La formation d'une intention,
- 3) La spécification d'une suite d'actions
- 4) L'exécution des actions,
- 5) La perception de l'état du système,
- 6) L'interprétation de l'état du système,
- 7) L'évaluation de l'état du système par rapport au but

Précisons la définition de but, d'état et de mécanisme de contrôle.

1) Un but est une représentation mentale de l'état à atteindre. Par exemple, "intervertir deux mots" d'un texte constitue un but. Dans le cas où le but ne peut être directement réalisé avec les commandes du système, l'utilisateur élabore un plan de résolution dans lequel le but est décomposé en une hiérarchie de sous-buts plus simples jusqu'à ce qu'ils soient réalisables avec une commande.

2) L'état d'un système se définit par un ensemble de variables physiques. Par exemple, le placement de la fenêtre (lucarne sur le contenu d'un document) et la localisation du curseur (point d'insertion de texte dans un document), sont des variables physiques qui intéressent directement l'utilisateur dans le but "intervertir deux mots".

3) Un mécanisme de contrôle est un dispositif physique qui permet d'agir sur les variables physiques. Par exemple, la souris permet de placer le curseur à l'endroit voulu ou de déplacer la fenêtre sur le contenu du document

Supposons qu'un utilisateur rédigeant une lettre avec MacWrite, ait pour but B : "Intervertir mot<sub>1</sub> et mot<sub>2</sub>". Le système se trouve dans l'état physique  $E(V1, V2)$ . La variable physique  $V1$  désigne la localisation actuelle du curseur et  $V2$  représente la suite de mots constituant la lettre soit :  $V2 = \{m_1 m_2 \dots m_i \text{ mot}_1 \text{ mot}_2 m_{i+3} \dots m_n\}$ .  $V1$  et  $V2$  sont contrôlables au moyen des dispositifs physiques souris et clavier. Pour simplifier l'exemple, nous supposerons que le contenu de la lettre est entièrement visible sur l'écran. Reprenons maintenant les étapes cognitives mises en jeu pour satisfaire B.

#### 1) Etablissement d'un but

Soit  $B$ , la représentation mentale de B.

#### 2) Formation d'une intention

L'utilisateur examine l'état du système et le compare au but recherché. Pour que la comparaison soit possible, état et but doivent être exprimés dans le même formalisme : ici, celui de la représentation mentale. L'état  $E(V1, V2)$  du système est traduit en une représentation mentale équivalente  $E(V1, V2)$  où  $V1$  et  $V2$  sont respectivement les représentations mentales de  $V1, V2$ . La différence (ou la distance) entre la représentation mentale  $E(V1, V2)$  de l'état du système et le but  $B$  donne naissance à une intention. Cette intention est la décision d'agir pour atteindre B.

#### 3) Spécification d'une suite d'actions

L'intention doit se concrétiser en une suite d'actions sur les dispositifs physiques de commande. Pour cela, les spécifications internes du but doivent être exprimées dans le formalisme de l'image :  $B$  est donc traduit en l'état physique souhaité  $E'(V1, V2)$  où  $V2 = \{m_1 m_2 \dots m_i \text{ mot}_2$

$\text{mot}_1 \text{ mot}_2 \dots \text{mot}_n$ .  $V_2$  désigne la suite de mots du document dans lequel  $\text{mot}_1$  et  $\text{mot}_2$  se trouvent maintenant inversés. Cette traduction requiert la connaissance des relations entre les variables psychologiques et les variables physiques. Elle nécessite aussi la connaissance des liens entre les variables physiques et les dispositifs de commande qui permettent de les modifier. En l'occurrence, l'utilisateur doit associer la variable psychologique "lieu d'insertion" à la variable physique "curseur" et il doit connaître le lien entre le curseur et le dispositif de commande "souris". Le résultat de la traduction est un plan de résolution, spécification mentale des actions à exécuter.

#### 4) et 5). Exécution des actions et perception de l'état du système

L'exécution des actions sont les actes moteurs qui conduisent au changement de l'état physique du système :  $E(V_1, V_2)$  devient  $E'(V_1, V_2)$ . Si l'image du système est sensible aux actions de l'utilisateur, le changement de l'état est immédiatement perceptible (on parle alors de retour immédiat d'information). La perception de l'état est une représentation mentale de  $E'$ , soit :  $E'(V_1, V_2)$ .

#### 6) Interprétation de l'état physique perçu

$E'(V_1, V_2)$  est maintenant interprété dans les termes des variables psychologiques d'intérêt, de celles notamment qui interviennent dans  $B$ .

#### 7) Evaluation

La comparaison entre l'interprétation de l'état perçu  $E'_m(V_{1m}, V_{2m})$  et  $B_m$  constitue l'évaluation. Cette évaluation peut conduire à la révision du plan et produire de nouveaux buts et intentions.

Les sept phases identifiées par D. Norman ne sont pas nécessairement présentes dans l'accomplissement d'une tâche ni forcément appliquées dans cet ordre. En particulier, les trois premières étapes sont récursives car au cours de la résolution, il est toujours possible de revenir sur le but, les intentions et la spécification du plan. L'intérêt premier de cette classification est d'identifier les points critiques, notamment la distance entre les variables psychologiques et les variables physiques, distance qui

complique les opérations de traduction et, par voie de conséquence, l'utilisation du système. D. Norman distingue la distance d'exécution de la distance d'évaluation.

### 3.4. Distance d'exécution et distance d'évaluation

La notion de distance exprime la dissimilitude entre la représentation de l'image et celle maintenue dans le modèle de l'utilisateur. La distance d'exécution traduit l'effort de mise en correspondance entre la représentation mentale interne de la tâche à effectuer et la représentation physique externe imposée par l'image. La distance d'évaluation traduit l'effort cognitif inverse (voir la figure 3.2). Puisque, de fait, ces distances sont inévitables, l'objectif du concepteur est d'en raccourcir la longueur par l'intermédiaire de l'image

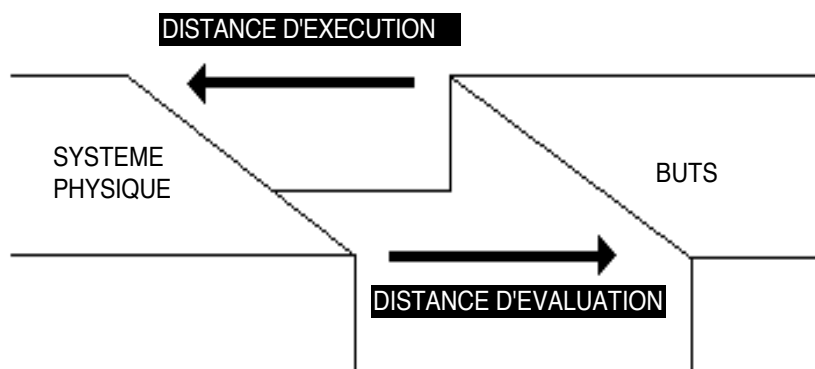


Fig. 3.2 : Les distances d'exécution et d'évaluation.

Du point de vue linguistique, l'image est un langage d'interface constitué de deux dialectes : l'un pour spécifier les expressions d'entrée, l'autre pour présenter les concepts du système sous forme d'expressions de sortie. Elle peut donc s'étudier selon les deux composantes : sémantique et forme. Hutchins et ses co-auteurs intègrent ces deux vues et parlent de distance sémantique et de distance articulatoire [Hutchins 86] :

1) La distance sémantique exprime les relations entre les buts que l'utilisateur s'est fixé et la signification des expressions du langage d'interface. Elle traduit l'effort de mise en correspondance entre la



connaissance sémantique que l'utilisateur a dans le domaine de la tâche et la connaissance sémantique qu'il possède sur le système.

2) La distance articulatoire reflète les relations entre la signification d'une expression du langage et sa forme. Elle traduit l'effort de traduction entre la connaissance sémantique que l'utilisateur a du système et sa connaissance syntaxique des éléments de présentation du système.

La figure 3.3 résume les relations entre buts, signification et forme. Elle montre comment les distances d'exécution et d'évaluation se décomposent en distance sémantique et en distance articulatoire. La distance d'exécution comprend la distance sémantique d'entrée et la distance articulatoire d'entrée. Dans l'autre sens, la distance d'évaluation inclut la distance articulatoire de sortie et la distance sémantique de sortie.

1) La distance sémantique d'entrée définit la différence entre le but recherché (par exemple, supprimer mot1) et la signification de l'expression d'entrée (par exemple, la commande couper). Cette distance est parcourue par l'activité qui aboutit à la formation d'une intention. Elle fait intervenir deux formes de connaissances sémantiques : celle sur le domaine et celle sur le système.

2) La distance articulatoire d'entrée représente la mise en correspondance entre la signification de l'expression d'entrée (commande couper) et la forme de cette expression (par exemple, sélectionner mot1 avec la souris puis taper la touche "retour-arrière"). Elle est franchie par la spécification des actions physiques. Elle fait intervenir les connaissances sémantiques et syntaxiques que l'utilisateur possède sur le système. L'exécution des actions entraîne la mise à jour (ou l'apparition) d'expressions de sortie. La perception de ces expressions engendre une forme (par exemple, mot1 vient de disparaître).

3) La distance articulatoire de sortie définit la différence entre la forme d'une expression de sortie (mot1 a disparu) et sa signification (par exemple, mot1 est détruit). Elle est franchie dans l'étape d'interprétation.

4) La distance sémantique de sortie établit la relation entre la signification de l'expression de sortie (mot1 est détruit) et l'objectif de l'utilisateur (supprimer mot1). Elle est franchie dans l'étape d'évaluation.

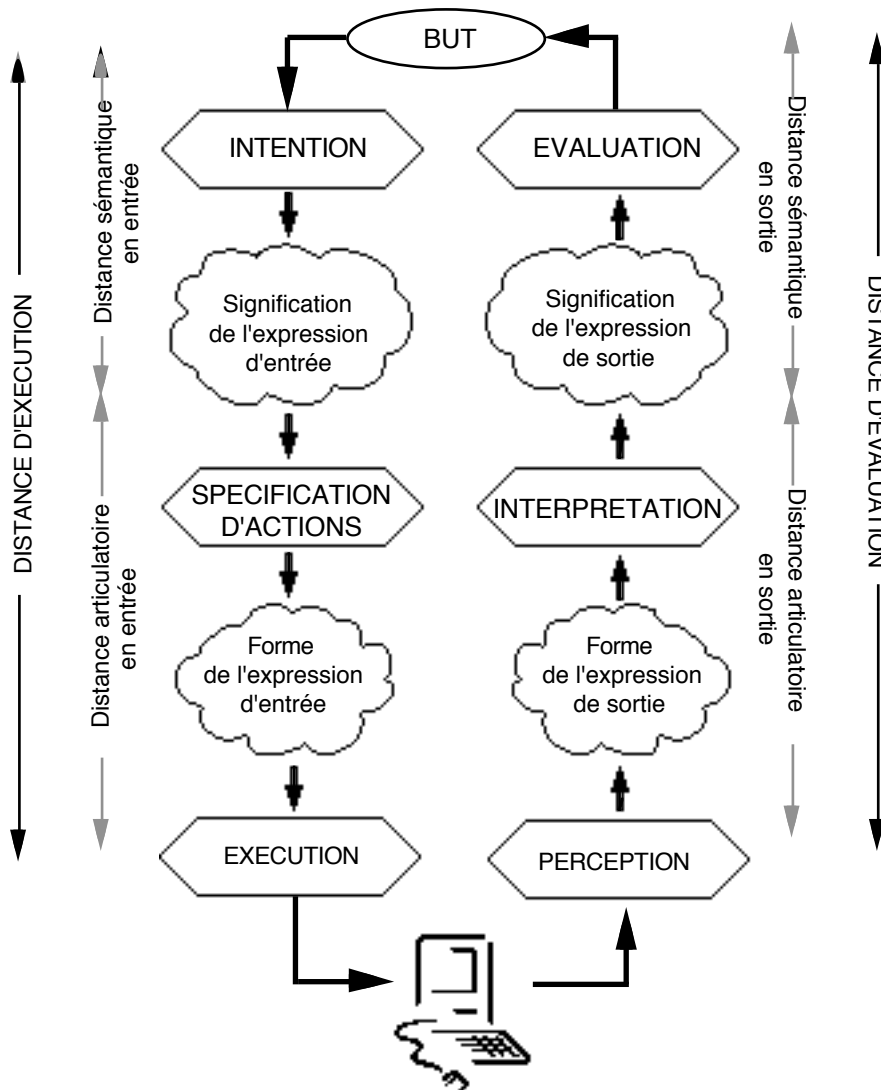


Fig. 3.3 : Distances sémantiques et distances articulaires.

### 3.5. Evaluation de la théorie de l'action

La théorie qui vient d'être présentée ne prétend pas être la théorie de l'action mais une théorie sur les processus cognitifs sous-jacents à la réalisation d'une tâche. Avec son point de vue cognitif, cette théorie complète les modèles GOMS et le Modèle du Processeur Humain en plusieurs points : elle précise la notion d'état, elle prend en compte les erreurs, elle explique les difficultés de l'utilisateur et justifie l'utilité de la notion de modèle conceptuel.

D'une théorie à l'autre, nous retrouvons la décomposition hiérarchique d'une tâche en buts et sous-buts jusqu'à ce que les sous-buts correspondent à des unités cognitives procédurales, un savoir-faire compilé. Alors que GOMS se contente de définir un but "comme une structure symbolique qui décrit un état à atteindre", la théorie de l'action précise la notion d'état. Elle distingue l'état effectif de l'état perçu. L'état effectif est une fonction portant sur des variables physiques, caractéristiques du modèle conceptuel du système, tandis que l'état perçu est la traduction de l'état effectif sous forme de variables psychologiques, caractéristiques du modèle conceptuel de l'utilisateur.

La différence de représentation entre le monde physique et le monde mental met en évidence la nécessité, pour l'utilisateur, d'effectuer des traductions. Contrairement à GOMS qui a une vue synthétique du comportement, la théorie de l'action analyse le processus de traduction qui sous-tend le comportement. Cette analyse identifie des phases auxquelles correspondent des besoins. Lorsque les besoins ne sont pas ou sont mal satisfaits, il y a risque d'erreur. Une erreur est commise soit parce que l'image n'explicité pas l'état effectif du système, soit parce que la correspondance entre les variables physiques et psychologiques est complexe, soit encore parce que les dispositifs de contrôle des variables physiques ne sont pas adaptés à la tâche. Tandis que GOMS se cantonne au cas idéal de l'utilisateur infallible, la théorie de l'action choisit le parti réaliste d'identifier les sources d'erreur et d'expliquer les difficultés rencontrées par l'utilisateur.

Erreurs, difficultés et réussites tiennent au modèle conceptuel que forge l'utilisateur avec l'expérience. Cette utilisation en lecture et en écriture du modèle conceptuel traduit bien l'aspect "système adaptatif" qui sert d'hypothèse à GOMS. En effet, l'adaptation revêt deux aspects complémentaires : l'assimilation et l'accommodation. Assimiler, c'est modifier l'environnement pour qu'il corresponde à soi. Accommoder, c'est se changer pour correspondre à l'environnement. L'utilisateur assimile lorsqu'il franchit la distance d'évaluation : dans les phases de perception et d'interprétation, il traduit une représentation de l'état de l'environnement en une représentation mentale qui peut éventuellement modifier le modèle conceptuel. L'utilisateur accommode lorsqu'il franchit la distance d'exécution : dans les phases de spécification d'actions et d'exécution, il

utilise en lecture le modèle conceptuel pour traduire l'objectif mental en une suite d'actions conformes à l'environnement.

Jusqu'ici, nous avons présenté la théorie de l'action comme un modèle explicatif des réussites, des difficultés et des erreurs de l'utilisateur. Nous verrons dans le chapitre sur les principes pratiques, l'utilité de ces explications pour classer les besoins de l'utilisateur et attirer l'attention du concepteur sur la nature de ces besoins.

La théorie de l'action n'est pas seulement un modèle explicatif. Elle peut, comme GOMS, Keystroke et le Modèle du Processeur Humain, servir de support à l'évaluation prédictive comme en témoigne ETIT. ETIT est une technique d'analyse que Moran a appliquée à l'étude des tâches de deux éditeurs de texte [Moran 83]. Moran part de la distance qui sépare la représentation mentale de la représentation physique. Il exprime sous forme de règles la mise en correspondance entre les concepts mentaux et les concepts du système. L'examen des règles fournit une métrique de la difficulté d'utilisation et d'apprentissage.

### **3.6. Résumé du chapitre**

La théorie de l'action assimile la réalisation d'une tâche au parcours d'une distance. Cette distance qui traduit la dissimilitude entre le monde mental et le monde physique fait intervenir les notions de modèles conceptuels et d'image du système ; elle comprend sept étapes.

Un modèle conceptuel est une représentation mentale. On distingue deux modèles conceptuels :

- 1) Le modèle de l'utilisateur qui est la représentation mentale que l'utilisateur a du système,
- 2) Le modèle de conception qui est la représentation mentale que le concepteur a du système pour un utilisateur type.

L'image a pour rôle d'explicitier le modèle de conception afin que l'utilisateur élabore un modèle mental en accord avec le modèle de conception.

Les sept étapes du processus cognitif qui conduit à la réalisation d'une tâche sont :

- 1) L'établissement d'un but. Un but est la représentation mentale de l'objectif en termes de variables psychologiques.
- 2) La formation d'une intention. Elle résulte de la différence entre le but et la représentation mentale de l'état actuel du système.
- 3) La spécification d'une suite d'actions. Cette spécification est une représentation mentale du plan de résolution ; elle résulte de la traduction du but en l'image physique souhaitée ; elle fait intervenir la connaissance des relations entre les variables psychologiques et physiques et entre les variables physiques et les dispositifs de commande.
- 4) L'exécution des actions. Elle consiste à manipuler les dispositifs physiques de commande pour modifier les variables physiques.
- 5) La perception de l'état du système. Elle est une représentation mentale de la nouvelle image du système.
- 6) L'interprétation de l'état du système. Elle s'effectue sur l'état perçu en fonction des variables psychologiques impliquées dans le but.
- 7) L'évaluation. Elle compare le résultat de l'interprétation au but recherché.