

Comprendre les médias interactifs

Louis-Claude Paquin

5 Les environnements immersifs

5	Les environnements immersifs	1
5.1	L'interaction immersive	3
5.1.1	Définitions	3
5.1.2	les précurseurs.....	4
5.1.3	Invention du dispositif	5
5.1.4	L'espace comme interface	7
5.1.5	Le point de vue	9
5.1.6	Boucle de rétroaction	9
5.1.7	Nature de la représentation.....	10
5.1.8	Des sensations plutôt que de l'information	12
5.2	Épistémologie cybernétique.....	13
5.2.1	La première cybernétique	14
5.2.2	Le concept de système	16
5.2.3	L'auto-régulation	17
5.2.4	L'échange d'information.....	20
5.2.5	L'auto-organisation, cybernétique de second ordre	22
5.2.6	Le connexionnisme	23
5.2.7	Les neurosciences.....	27
5.2.8	La vie artificielle	29
5.2.9	Les cyborgs	32
5.3	Composants d'un dispositif d'immersion.....	34
5.3.1	La poursuite de la position et de l'orientation.....	34
5.3.2	Casque de visualisation	35
5.3.3	Gant de données	36
5.3.4	Environnement sonore	37
5.3.5	Stations graphiques	38
5.3.6	Chambre d'immersion	40
5.3.7	Autres capteurs	41
5.3.8	Les dispositifs semi-immersifs	42
5.3.9	Le partage des environnements immersifs.....	43



Comprendre les médias interactifs de Louis-Claude Paquin est mis à disposition selon les termes de la licence Creative Commons 4.0 : Attribution - Pas d'utilisation commerciale – Pas de modification.

5.4	L'environnement.....	45
5.4.1	L'ontologie.....	45
5.4.2	La fermeture.....	46
5.4.3	La modélisation 3D	48
5.4.4	Modélisation de la nature.....	49
5.4.5	La lumière	50
5.4.6	La caméra et le rendu	51
5.4.7	L'animation.....	52
5.4.8	Méthodes d'animation	54
5.5	Les personnages.....	55
5.5.1	Typologie	55
5.5.2	Modélisation.....	56
5.5.3	Animation.....	58
5.5.4	Capture de mouvement.....	59
5.5.5	Comportements.....	60
5.5.6	La parole	62
5.5.7	L'expression non verbale	66
5.5.8	L'avatar.....	68
5.6	Effet de présence.....	69
5.6.1	Définition.....	69
5.6.2	Suspension de l'incrédulité.....	70
5.6.3	Facteurs	72
5.7	Bibliographie	75



5.1 L'interaction immersive

On vient de décortiquer un premier type de médias interactifs, les « machines à contenu ». L'interface de ces médias interactifs permet de manipuler directement les représentations symboliques, organisées la plupart du temps autour d'une métaphore, tirée de préférence de la vie quotidienne. Cette interface permet d'accéder au contenu des fichiers informatiques via des menus textuels ou iconiques, ou encore de formulaires. Maintenant, on va étudier un second type de médias interactifs. Radicalement différents des premiers, les « environnements immersifs » remplacent les manipulations *indirectes* ou *symboliques* par des manipulations *directes* et *concrètes*, dans l'environnement même. Littéralement aspiré par la représentation, l'utilisateur se trouve désormais impliqué dans l'action. La différence entre ces deux types de médias interactifs est fondamentale: les premiers se fondent sur une épistémologie de type mécanique, alors que les seconds sont basés sur une épistémologie cybernétique.

5.1.1 Définitions

Pour être immersif, l'environnement doit *entourer* l'utilisateur. Avec un casque de visualisation ou dans une pièce dont les murs sont constitués d'écrans, il l'entoure directement. Lorsque l'utilisateur est représenté à l'intérieur d'un environnement par un personnage « virtuel » appelé *avatar* (cf. 5.5.8), il est entouré indirectement. Dans les deux cas, l'utilisateur ou son avatar doivent pouvoir se déplacer librement dans cet environnement, interagir avec lui, avec les objets qu'il contient, ainsi qu'avec d'autres participants, présents ou en ligne, via Internet.

L'expression « environnements immersifs » réfère à la fois au *contexte* et à la *nature* de l'interactivité. Elle est encore mouvante puisqu'elle décrit une pratique émergente, en pleine évolution et donc non consolidée. Cela ouvre la voie aux transformations, aux pratiques extrêmes ou limites, aux transferts d'un domaine à l'autre, d'un usage à l'autre. Plusieurs autres termes ont été utilisés pour décrire ce mode de communication. Chacun met l'accent sur l'une ou l'autre caractéristique :



le terme « réalité virtuelle ou artificielle » suggère que le monde est représenté de façon réaliste ou imaginaire ; dans ces expressions, le terme « monde » est souvent substitué au terme « réalité ». Le terme « téléprésence » veut rendre l'idée que la représentation est captée à un endroit différent de sa réception, et que les gestes de l'utilisateur modifieront à distance le monde concret objet de la captation.

L'expression « télévirtualité » renvoie au fait que les mondes immersifs peuvent faire l'objet d'une expérience partagée par plusieurs utilisateurs. Le terme « réalité augmentée » est utilisé lorsque la représentation de l'environnement immersif est affiché en surimposition sur le monde réel, par le biais d'écrans semi-transparents.

Comme le fait remarquer avec justesse Lev Manovich (1998), la production de représentations illusionnistes est devenue le domaine des technologies mass-médiatiques, comme la photographie, le film et la vidéo, qui sont remplacées par les technologies digitales (cf. 1.5). Avant le 20^e siècle, c'est l'art qui occupait cette fonction, pour finalement l'abandonner au profit de l'abstraction, et plus tard de l'engagement de la personne. La notion de *réalisme* même change avec les environnements immersifs: elle n'implique plus uniquement sur une illusion visuelle, mais aussi une implication multisensorielle du corps de l'utilisateur.

Les environnements immersifs s'évaluent donc non seulement par leur esthétique visuelle, mais aussi par la qualité de l'expérience que vit l'utilisateur. La résolution spatiale, la coloration, la sonorisation, la souplesse des mouvements ou encore la complexité des personnages: tous ces points constituent des critères d'appréciation importants, autant sur le plan géométrique que psychologique.

5.1.2 les précurseurs

Dans les années 1950, le cinéaste Morton Heilig cherche à développer une *machine à réalité* qui pourrait simuler la richesse sensorielle de la vie. Il avance l'idée qu'une compréhension scientifique des sens et de la perception améliorerait la puissance créatrice de l'artiste. Sa prémisse était simple mais visionnaire en son temps : si l'artiste contrôle la stimulation multisensorielle des spectateurs, il peut leur donner l'illusion et la sensation d'une expérience personnelle, " comme s'ils y étaient ".



Heilig est inspiré par d'autres tentatives en ce sens: le *Cinerama*, dont l'écran dépasse le champ de vision, et le *Cinéma 3D*, qui donne une illusion de profondeur une fois qu'on le regarde avec des lunettes de carton, verte d'un côté et rouge de l'autre. Pour Heilig, l'extension logique du cinéma, c'est d'immerger les spectateurs dans un univers fabriqué qui stimulerait tous leurs sens. Il était convaincu qu'en élargissant les stimulations sensorielles du cinéma (la vision et l'audition) aux autres sens — goût, toucher et odorat —, il parviendrait à décontextualiser la salle de cinéma pour transporter les spectateurs dans un monde « virtuel ». Il nomme ce cinéma du futur le *théâtre expérientiel*.

En 1962, il expose une étrange machine qui ressemblait aux visionneuses à péage que l'on retrouve dans les parcs d'amusements et les arcades. Appelée *Sensorama*, cette machine propulsait les spectateurs dans une excursion multisensorielle sur une motocyclette zigzaguant au travers des rues de Brooklyn. Le spectateur pouvait sentir l'odeur des pizzas, le gaz d'échappement des voitures, etc.

5.1.3 Invention du dispositif

Diplômé du Massachusetts Institute of Technology (MIT), Ivan Sutherland est un pionnier. Il défriche non seulement le champ de l'infographie avec l'application *Sketchpad* (1963), mais aussi celui de la technologie qu'on baptisera plus tard « réalité virtuelle ». Par le biais de sa participation au projet ARPA (durant quatre ans), Sutherland a pu explorer différentes hypothèses qui devaient permettre d'intégrer plus et plus les agents humains et l'ordinateur. Devenu professeur à Harvard, il propose une idée pionnière dans un article intitulé « Un dispositif d'affichage en 3D monté sur un casque » (1968) (*A Head-mounted Three-Dimensional Display*). L'idée était de placer l'utilisateur à l'intérieur de la représentation graphique d'un environnement généré par un ordinateur. Pour ce faire, il propose de créer des lunettes spéciales montées sur un casque, et qui comporteraient, en lieu et place des verres correcteurs ou grossissants, des petits écrans sur lesquels serait affichée une image de l'environnement, qui changerait en fonction des mouvements de tête de l'utilisateur.



Cette idée était déjà dans l'air à l'époque : dans les années 1950, la Philco Corporation avaient entamé des recherches dans l'idée de réaliser un dispositif qui permettrait d'opérer efficacement à distance dans des conditions de danger extrême — sous l'eau, dans l'espace ou encore dans les usines nucléaires. Ce domaine de recherche est devenu ce qu'on a appelé par la suite la « téléprésence ». Un dispositif élaboré en 1958 permettait à un opérateur de voir sur un écran les images provenant d'une caméra placée à distance. Cet écran était monté sur un casque et placé en face de ses yeux. L'utilisateur pouvait contrôler le point de vue de la caméra en bougeant la tête.

Si le projet de Sutherland était achevé dès 1968, sa réalisation (dans les années 1970) s'est avérée toutefois beaucoup plus difficile que prévu. Le poids du casque excédait le supportable, à cause du poids combiné des écrans et du dispositif de repérage, qui devait donner la position de l'utilisateur de manière à assurer un feedback approprié. Une première version du dispositif était reliée à l'ordinateur via le plafond, par l'intermédiaire d'un bras qui servait à détecter mécaniquement les mouvements de l'utilisateur. Par ailleurs, l'ordinateur devait alimenter les écrans en temps réel, de manière à produire une perspective tridimensionnelle réaliste. Pour ce faire, il a fallu concevoir certains modules : un multiplicateur de matrices, un générateur de vecteurs, ainsi qu'un module qui permettait de dédoubler l'image obtenue, avec un léger décalage qui induisait un effet de profondeur. Connectés au casque, ces modules constituent un système intégré qui immergent l'utilisateur dans un monde virtuel composé de formes géométriques.

Plusieurs expérimentations ont été réalisées dans le même sens au cours des années 1970. Notamment, l'*Architecture Machine Group* du MIT — qui deviendra en 1980 le *Media Laboratory* — a réalisé la fin des années 1970 un projet de cartographie de la ville d'Aspen au Colorado. Chaque rue et chacun de ses recoins furent filmés à l'aide d'une caméra spéciale, arrimée au toit d'une voiture. Ces séquences furent ensuite combinées à des vues aériennes de la ville captées à partir de grues, d'hélicoptères et d'avions, ainsi qu'à des séquences prises à l'intérieur de bâtiments. Assis devant un écran tactile, l'utilisateur y indiquait directement quelles



rues il voulait emprunter et dans quels édifices il voulait entrer. Il était entouré par une vue de devant, d'arrière et de côté, bref, immergé.

Cette nouvelle technologie s'est développée dans deux directions. Elle a débouché d'une part sur les *systèmes totalement immersifs*, hauts de gamme, chers et très performants, et d'autre part sur les *systèmes semi-immersifs*, abordables, voire à bas prix, qui tournent sur des ordinateurs personnels, sur des consoles de jeu affichant sur des moniteurs, ou, plus simplement, directement sur le poste de télévision. Les consoles de jeu se fondent sur les mêmes principes que les ordinateurs personnels, mais la prothèse qu'est la manette de jeu remplace le tandem souris clavier. Cette technologie s'utilise aussi bien dans la sphère des jeux que dans celle des applications utilitaires: les simulateurs pour les avions et autres véhicules destinés aux voyages extraterrestres, les modélisateurs architecturaux, qui permettent de voir l'intérieur d'un bâtiment avant qu'il ne soit construit à partir des plans, etc.

Explorons maintenant les différentes facettes des environnements immersifs : dans le rôle de la représentation, l'espace remplace la surface ; par la boucle de rétroaction et la nature de l'environnement représenté, on passe de l'information à la sensation.

5.1.4 L'espace comme interface

Dans les machines à contenu, la souris permet à l'utilisateur d'interagir avec l'ordinateur de manière directe sur la surface bidimensionnelle de l'écran. Le pointeur nommé « curseur » permet de lancer un processus, ou encore de sélectionner, de déplacer un objet typographique ou iconographique. Le type de communication entre les personnes et l'ordinateur sous-jacente aux machines à contenu véhicule une certaine représentation des interacteurs: les personnes et les ordinateurs sont conçus comme des entités contiguës, mais radicalement différentes, au point qu'il faut recourir à un langage typographique ou iconique pour que les premiers puissent accéder au contenu que livrent les seconds.



Les environnements immersifs procèdent à l'inverse des machines à contenu. Les utilisateurs sont intégrés dans un espace « virtuel », qui n'a d'existence que par la médiation de l'ordinateur. Le lieu de la représentation n'est plus la surface bidimensionnelle, qui est, en fait, une espèce de livre transposé et adapté à l'écran. Emerge un espace tridimensionnel, comme on en trouve au cinéma. Une caméra se substitue à notre regard, et détermine le point de vue adopté sur l'espace affiché à l'écran, pour capter et rendre le monde.

La communication entre les personnes et les ordinateurs se conçoit de manière radicalement différente avec les environnements immersifs : ces deux entités sont prises sur le même pied et s'inscrivent toutes deux dans la dynamique d'un même système. La personne est pleinement incluse dans l'environnement généré par l'ordinateur, ce qui veut dire que tous ses gestes sont échantillonnés et rapportés dans l'environnement, modifié en fonction en temps réel.

Le composant principal des dispositifs d'environnement immersifs, c'est donc le module qui inscrit la personne dans l'espace de l'environnement représenté. Ce module est composé d'instruments qui communiquent à l'ordinateur des informations de nature tridimensionnelle, appliquées ensuite aux coordonnées et aux autres paramètres de l'espace « virtuel ». De cette manière, on détermine à la fois la position et l'orientation dans l'espace de la caméra, qui constitue le point de vue sur le monde qu'adopte l'utilisateur. Vu le rôle prépondérant qu'il joue dans la représentation, l'utilisateur devient pleinement interacteur.

Les utilisateurs ne sont plus *en interface* avec le contenu. Ils se trouvent littéralement *au cœur* du contenu. Celui-ci doit être organisé en conséquence, dans un espace plus ou moins fermé. Cet espace est dit « virtuel » parce qu'il n'existe matériellement que par l'intermédiaire de l'ordinateur. L'écran quitte son rôle de surface en charge d'organiser ou d'afficher les contenus multimédias, que ce soit sous la forme d'assemblages de textes, d'images fixes ou animées, de vidéos. *L'écran devient un espace*, un environnement dynamique qui rétroagit sur les interactions en temps réel, en émettant les stimuli sensoriels adéquats.



5.1.5 Le point de vue

Au cinéma comme en peinture, un point de vue est construit sur la scène par le cadrage. On distingue les points de vue exocentrés et endocentrés. Il y a point de vue *exocentré* (ou *objectif*) sur le monde lorsque l'existence même de la caméra s'efface : le point de vue est celui qu'adopte l'utilisateur sur le contenu au travers d'une interface. Le point de vue est *endocentré* (ou *subjectif*) lorsque le monde semble vu au travers du regard d'un personnage auquel le spectateur peut s'identifier. L'impression caractéristique du point de vue subjectif est renforcée lorsque l'angle de vision est vraisemblable, et qu'il y a des indices de la présence d'un regard dans la scène même: déformation et mouvement de l'image, présence d'une partie du corps en avant-plan, etc. Le point de vue subjectif favorise la centration, ce qui incite l'utilisateur à participer activement à la représentation; le point de vue objectif, à l'inverse, favorise la décentration, et incite ainsi à la réflexion, au recul, ménage une distance entre le spectateur et l'objet perçu.

5.1.6 Boucle de rétroaction

Revenons un instant aux machines à contenu. On l'a déjà dit, le cycle de base du dialogue entre la personne et l'ordinateur était calqué sur celui de l'ordinateur : requête — traitement — réponse. L'utilisateur est toujours à l'extérieur du contenu. Pour y accéder librement, il doit recourir à des dispositifs de contrôle de type mécanique : icônes, boutons, menus, etc.

Les environnements immersifs sont très différents. Cette fois, l'interactivité se fonde sur les principes de la cybernétique, auxquels est consacrée la section suivante. Schématiquement, l'humain et l'ordinateur se retrouvent tous deux inclus à l'intérieur d'un même système, l'environnement. Le cycle de base repose sur la *rétroaction négative* : chaque action de l'utilisateur indiquant que la direction de son regard sur l'environnement a changé — il s'agit surtout des mouvements de la tête — entraîne un déséquilibre dans le système. Ce déséquilibre provoque sans délai une action qui tente de rétablir l'état d'équilibre initial. S'instaure une boucle sans fin entre l'utilisateur et le système. Récursivement, les machines émettent des



informations sonores et visuelles, reçues et traitées par les humains qui, en retour, transmettent leur réaction aux machines par le biais de prothèses adaptées; les machines réagissent alors en adaptant les informations sonores et visuelles envoyées aux humains en fonction, etc.

Les systèmes que constitue le tandem utilisateur — environnement immersif ne devraient pas se limiter à la boucle de rétroaction. En effet, celle-ci limite le rendu de l'environnement, restreint à suivre les gestes de l'utilisateur. Pour augmenter le degré de vraisemblance et d'immersion, les objets et les personnages qui composent l'environnement devraient se voir appliqués les principes de vie artificielle, pour devenir des entités capables de traiter l'information obtenue de l'environnement par elles-mêmes, et l'utiliser pour effectuer des actions en fonction de certains buts.

Pour simuler les capacités de reproduction et d'évolution dans des entités artificielles, on préconise généralement l'approche ascendante, qui consiste à développer quelques comportements simples et indépendants les uns des autres, qui, une fois assemblés, génèrent des comportements plus complexes. La complexité résulte de l'application itérative de règles simples lors de boucles de rétroaction qui modifient à la fois les entités et l'environnement. Ces événements sont déclenchés d'eux-mêmes suivant des règles internes au monde représenté. Par exemple, si le monde représenté est un jardin, il sera affecté par le cycle des pluies et des sécheresses, que celui-ci soit déterminé par un service météo en ligne, ou bien régulé par un algorithme probabiliste, totalement aléatoire ou contextualisé (exemple: plus la période de pluie a été longue, plus il y a de chances qu'une période de sécheresse s'ensuive).

5.1.7 *Nature de la représentation*

Contrairement aux machines à contenu (i.e. " permettant d'accéder au contenu par une interface "), les environnements immersifs font converger dans une même représentation le contenu et les caractéristiques formelles du média. Les environnements immersifs constituent des représentations audiovisuelles basées sur



des modélisations, c'est-à-dire des construits, qui entretiennent des *relations analogiques* avec leurs référents. Ces construits sont stylisés: leurs concepteurs doivent adopter une démarche artistique, et même esthétique. Le référent est conçu comme un ensemble de traits perceptifs conventionnalisés, qui constitue un modèle perceptif de l'environnement et des objets qui le composent.

Ces environnements peuvent être *réalistes*, et donc imiter des environnements qui existent ou pourraient potentiellement exister, ou bien *imaginés*, intégralement, comme dans la science-fiction. Le référent de l'environnement immersif peut ainsi être altéré de façon significative, ou bien encore ne pas exister du tout. Le référent transforme alors la réalité de manière totale ou partielle, tout en préservant certaines ressemblances qui permettent aux utilisateurs de s'y adapter.

Enfin, les environnements peuvent aussi être complètement abstraits, représenter l'organisation de bases de données ou de connaissances, ou encore le résultat d'une requête dans ces bases. Outil très performant, la cartographie tridimensionnelle d'un espace informationnel permet ainsi de repérer et de comprendre les connexions existant entre différents nœuds ou différentes régions de connaissances. L'idée consiste à traduire les catégories qui organisent un domaine en une configuration spatiale d'items ou de concepts, ce qui permet de mettre l'accent sur les relations qui existent entre eux. Dans le cas des espaces documentaires, le rendu s'opère à partir d'un concept pris comme point focal; tous les autres concepts qui y sont reliés sont disposés tout autour de ce point focal, la proximité exprimant la force de la relation statistique, et l'éloignement l'absence de ce rapport. Le degré de proximité peut se voir déterminer en fonction de plusieurs critères différents, comme la fréquence, la cooccurrence, etc. Lorsque l'utilisateur sélectionne un autre concept pour point focal, l'espace est réorganisé dynamiquement autour de ce nouvel attracteur. Il existe encore d'autres stratégies d'organisation de l'espace documentaire, présentées en détail au chapitre suivant.

Remarquons encore que puisque l'espace objet de la représentation peut être simulé et même imaginé, l'expression « environnement virtuel » est préférable à



celle de « réalité virtuelle », souvent employée pour qualifier ce dispositif, mais en ne rendant qu'un type d'environnement représentable.

5.1.8 Des sensations plutôt que de l'information

La notion d'*engagement direct* (dans l'espace représenté) fait pendant à la notion de *manipulation directe*. L'ordinateur devient un média qui permet à l'agent de participer à l'action. La communication entre les personnes et l'ordinateur n'est plus seulement de nature informationnelle, mais aussi de nature expérientielle: un contexte d'action est partagé, et chacun y agit. Il s'agit là d'une différence appréciable. Lorsque l'information est communiquée par les médias traditionnels, elle se présente comme un enchaînement stylisé de faits. La réalité perd alors son contexte et les interrelations voire les correspondances qui la caractérisent.

L'information s'obtient par traitements, extractions, abstractions de la réalité. Celle-ci se retrouve ainsi condensée et généralisée, mais ne nous atteint qu'intellectuellement: il manque la complétude de l'expérience. Par contre, une fois que nous sommes directement impliqués dans les situations, la totalité des processus qui nous caractérisent en tant que personne est mobilisée : processus conscients, inconscients, viscéraux et mentaux. L'expérience maintient et nourrit les liens contextuels de l'environnement, ce qui engage l'utilisateur dans la représentation, aussi bien au plan sensoriel qu'au plan intellectuel.

Pour que les médias interactifs soient immersifs, au même titre que les autres médias audiovisuels tels que le cinéma, la vidéo, la télévision et même la radio, ils doivent devenir bien plus qu'un simple moyen d'accéder à l'information au moyen d'une interface. Ils doivent provoquer une réelle implication de la part du spectateur, non seulement cognitive (au niveau de l'information), mais surtout perceptive, en activant certaines représentations. L'utilisateur n'a plus simplement affaire à une interface dont il doit activer les menus et les boutons, pour accéder à un contenu segmenté en fiches (ou bien à un autre type d'unité structuré selon un arbre logique, à un réseau associatif ou encore à un récit). Maintenant, il se trouve littéralement plongé dans un environnement simulé; le rapport au média n'est plus



de nature instrumentale, mais bien de nature expérientielle. Pour que cette illusion fonctionne, il faut fournir aux systèmes sensoriel, cognitif et affectif de l'interacteur un stimulus audiovisuel rendant les événements qui surviennent dans l'environnement simulé. Le terme *immersion*, on l'a déjà dit, est employé ici de façon métaphorique. Le monde réel, lui, est rendu invisible par le casque, étouffé par les écouteurs.

En tant que médias, les environnements immersifs constituent l'ultime tentative d'abolissement de la distance entre le spectateur et le spectacle. Le spectateur peut désormais faire l'expérience de la représentation directement, avec tout son corps et tous ses sens, plutôt que par son imaginaire. Les environnements immersifs sont le produit d'une activité intentionnelle dont la finalité est la production de sens. Les messages prennent la forme de stimuli visuels et sonores, élaborant un univers de représentation qui n'a d'existence que le temps de son exploration.

5.2 **Épistémologie cybernétique**

À la fin des années 1940, la cybernétique est née du contact de chercheurs venus de plusieurs horizons. C'était une première manifestation tangible de ce que la pensée machinique héritière de la révolution industrielle qui avait régné jusque là se renouvelait scientifiquement et techniquement. Suite, entre autres, aux gains de savoir générés par les développements de la thermodynamique, le contrôle des systèmes a évolué vers des dispositifs que l'on peut maintenir en un état de contrôle continu, même si le débit d'information ou d'énergie n'est pas totalement prédictible. Cette évolution est importante parce qu'elle a permis de développer des moteurs à haute intensité d'énergie. Lorsque les systèmes sont complexes, il est préférable d'étudier leurs comportements à partir de l'approche statistique plutôt que par l'approche causale, qui essaye de déterminer des causes spécifiques et précises pour chaque effet.

Après le foisonnement de la période fondatrice, la cybernétique se déclinera en plusieurs disciplines ou plusieurs champs d'application de ses principes de base : le projet d'intelligence artificielle, et plus spécifiquement le *connexionnisme* ; les



théories de l'auto-organisation, qui ont débouché sur un projet de vie artificielle dans les années 1960; et, finalement, la *théorie des systèmes* qui a traversé, durant les années 1970, des disciplines scientifiques très différentes, appartenant notamment aux sciences sociales et humaines.

L'épistémologie cybernétique ne se concentre pas sur les différences radicales qui séparent les personnes et les informations numériques contenues dans les bases de données des ordinateurs. En fait, elle met les échanges entre les constituants humains et numériques sur le même plan, à l'intérieur d'un même ensemble. Dans cette section, on va mettre à jour les fondements épistémologiques des environnements immersifs, second paradigme des médias interactifs. La cybernétique fournit des concept-clés qui favorisent et expliquent en même temps l'effet de présence qui caractérise ces médias. Ces concepts-clés sont la *rétroaction*, la *régulation* et l'*écologie*.

5.2.1 La première cybernétique

Le terme « cybernétique » est forgé à partir du grec *kubernêsis*, qui signifie littéralement « action de manœuvrer un vaisseau ». D'où le terme " gouvernail " qui désigne le dispositif qui assure cette manœuvre. Par extension, le terme désigne toute action qui consiste à diriger et à gouverner. En 1942, lorsque N. Wiener, J. Bigelow et A. Rosenblueth ont voulu donner un nom à leur méthode d'étude comportementale, ils ont donc utilisé le terme « cybernétique ». Ce même mot avait déjà été utilisé par le physicien français André Ampère (1775-1836) dans un traité sur la science du gouvernement. Dans l'étude de n'importe quel phénomène, naturel ou artificiel, la cybernétique met l'accent sur le comportement, c'est-à-dire sur les modifications que subit le phénomène en vertu de son environnement. Le terme " phénomène " est vague à souhait, et désigne potentiellement tout objet qui peut se manifester à la conscience : minéraux, végétaux, animaux, personnes, objets, etc., que ce soit par l'intermédiaire des sens ou du langage.

Au départ, l'intention sous-tendant la cybernétique était d'unifier les concepts, au-delà des barrières qui séparent la physique, la chimie et la biologie. L'accent sur la



compréhension par le comportement s'inscrivait en rupture totale avec la méthode classique des sciences, réductionniste, et fondée sur des schémas mécanistes. Cette méthode nouvelle consistait à essayer de comprendre les phénomènes de l'intérieur, en les ouvrant, en les disséquant, en isolant leurs parties constitutives pour tenter de percevoir la logique causale linéaire qui les sous-tend. Il s'agit d'une approche *holiste* (par la totalité) et sans découpage, qui veut expliquer les interactions nombreuses et complexes caractérisant les parties des systèmes, en mettant l'accent sur le mécanisme de *contrôle* et sur la *communication*.

La cybernétique privilégie l'étude des comportements des entités et des dispositifs, indépendamment de la nature physique des organes qui les composent. Elle laisse de côté les différences physiques et spirituelles entre ces entités. Par cette position, elle peut tenir compte en même temps des organismes vivants et des machines.

Le projet de la cybernétique est double. D'une part, elle veut expliquer le comportement des systèmes aussi bien artificiels que naturels, ainsi que les échanges d'information entre ces systèmes. Pour ce faire, elle part principalement de concepts mathématiques. D'autre part, elle veut réaliser des machines capables de reproduire certains comportements des systèmes naturels. Ces « machines » prendront la forme de *réseaux neuronaux*. Le cerveau constitue une source importante d'inspiration pour ces deux projets. La cybernétique fait le lien entre plusieurs domaines du savoir jusque là plus ou moins étanches, tels que les mathématiques, l'électronique, la physiologie, la psychologie, la sociologie, l'éthologie, etc.

Le livre de Norbert Wiener intitulé *Cybernetics or Control and Communication in the Animal and the Machine* (1948) reste le texte fondateur de la cybernétique.

L'objectif de cette discipline est de comprendre la nature de l'organisation et du contrôle de n'importe quel processus dont les finalités peuvent être identifiées, et ce en étudiant la manière dont l'information est communiquée. Pour ce faire, elle recourt à certaines notions-clés, telles que la *théorie des systèmes*, la notion de *feedback* et d'*échange d'information*. Une des caractéristiques épistémologiques les



plus importantes de la cybernétique, c'est l'absence de toute métaphysique, de tout déterminisme ou de tout principe agissant extérieurs aux systèmes eux-mêmes.

5.2.2 Le concept de système

La *Théorie générale des systèmes* fut élaborée par Ludwig von Bertalanffy, biologiste autrichien qui publia en 1950 *The general System Theory*. Tout comme cette théorie, la cybernétique définit les systèmes comme des *dispositifs intentionnels*, c'est-à-dire des systèmes orientés vers un but à atteindre, et non pas guidés par le hasard. La théorie des systèmes cherche à établir un cadre le plus général possible à l'intérieur duquel on peut étudier le comportement d'une entité complexe analysable, plus précisément son évolution au cours du temps. Cette théorie s'occupe d'entités abstraites, à savoir les systèmes. Pour que la théorie puisse être effective et s'appliquer à des entités concrètes, il faut établir une certaine correspondance entre un système et le type d'objet étudié : corps matériel, ensemble de corps liés d'une manière plus ou moins rigide, être vivant, société d'êtres vivants, etc.

Un système est un objet complexe. Il est formé de composants distincts reliés entre eux par un certain nombre de relations. Les composants sont considérés comme des sous-systèmes. Donc, ils appartiennent à la même catégorie d'entités que les ensembles dans lesquels ils s'intègrent. Un sous-système peut soit être décomposé à son tour en sous-systèmes d'ordre inférieur, soit être traité (au moins provisoirement) comme un système indécomposable, c'est-à-dire comme un système réduit à un seul élément.

Le système est plus complexe que ses parties. Autrement dit, il *possède des propriétés irréductibles à celles de ses composants*. Si ces propriétés qui caractérisent le système sont irréductibles à celles des sous-systèmes, c'est à cause des relations qui unissent les composants. L'évolution d'un système est conditionnée à la fois par les modifications internes des composants, et par les interactions qui peuvent s'établir entre le système et son environnement. Au cours de son évolution,



un système peut conserver une certaine stabilité; il peut aussi se transformer dans deux directions : soit il se désagrège, soit il s'intègre davantage.

Les systèmes sont biologiques ou machiniques. Les systèmes biologiques sont les systèmes vivants. Ils sont constitués de cellules plus ou moins structurées, guidés par leur instinct de survie suivant le postulat darwinien, ce qui en fait des *systèmes adaptatifs*. Parmi ces systèmes, les humains font exception : leur univers symbolique et culturel peut inverser la polarité; ils sont alors animés par leur désir de destruction, la pulsion de mort. Par extension, le concept de système vivant peut être appliqué aux sociétés. Les systèmes non biologiques, eux, sont des machines. Ces artefacts se regroupent en deux classes; s'ils sont *matériels*, on parlera d'automates, de robots, et s'ils n'ont de forme et d'existence que dans un univers numérique, ils sont *immatériels*.

Appliquée aux sciences de l'homme, la notion de système sera étendue. Elle désigne alors *l'interaction* entre un ensemble d' « entités », et la *formation*, dans cet ensemble d'entités, de *configurations* relativement stables, déterminant l'évolution conjointe des entités en tant qu'elles forment ensemble un système. L'étude des systèmes désignera, alors, l'analyse d'une entité, d'un discours, d'une théorie, d'une pratique, d'une institution, etc. Ce genre d'analyse consiste, plus précisément, en l'étude des principes internes à l'objet étudié qui détermine sa cohésion et son insertion dans divers types d'ordres : réseaux, milieux, structures, modèles, ensembles, etc. Ce genre d'étude se place à un autre niveau que celle des genres ou des disciplines prédéterminés. La notion de système permet d'articuler les savoirs à leur environnement historique, social et culturel.

Pour la cybernétique, les systèmes que forment les machines, les organismes biologiques et les sociétés se caractérisent tous par le flot de la *communication inter-système*, ainsi que par la *régulation des transformations* du système, suivant les contraintes de l'environnement. Ces deux éléments déterminent le comportement du système.

5.2.3 L'auto-régulation



Comprendre les médias interactifs de Louis-Claude Paquin est mis à disposition selon les termes de la licence Creative Commons 4.0 : Attribution - Pas d'utilisation commerciale – Pas de modification.

La cybernétique, c'est l'étude des systèmes considérés sous l'angle du *contrôle* et de la *communication de l'information*. Le contrôle (ou la régulation) d'un système, c'est le principe qui en assure la stabilité ou l'équilibre, ainsi que le principe agissant (c'est-à-dire celui qui fait fonctionner le système et qui en assure l'intégrité). Les systèmes sont *autorégulés*, en ce sens qu'ils ne sont pas soumis à une chaîne causale linéaire de type action/réaction, cause/effet, mais bien à un processus continu de rétroaction, suite à des échanges d'information avec l'environnement. Un système reçoit une information de l'environnement; celle-ci circule au travers des sous-systèmes et des composants, ce qui induit une transformation du système lui-même. La régulation détermine à la fois la redistribution de l'information reçue à l'intérieur du système, et le type d'actions déclenchées pour adapter le système aux transformations de l'environnement. La cybernétique vise à décrire formellement les mécanismes d'autorégulation d'une entité donnée, voire à les reproduire. Dans la mesure où un système a des mécanismes de régulation qui lui permettent de conserver son équilibre malgré les transformations de l'environnement, il peut contrôler son propre fonctionnement et donc se gouverner lui-même.

En tant que dispositifs intentionnels, les systèmes se subdivisent en deux catégories: les dispositifs déterministes et les dispositifs adaptatifs. Les *systèmes déterministes* sont dits *fermés*, parce que leur comportement dans l'environnement et avec les autres systèmes est fixé et prédéterminé. Ces systèmes déterministes constituent des *automates à états finis* (cf. 2.1.6), dotés d'un ensemble fermé de règles prédéterminées et immuables. En revanche, le comportement des *systèmes adaptatifs* peut se voir modifié au cours de l'action par rétroaction, de manière à adapter le système aux transformations de l'environnement. C'est pour cela que les systèmes biologiques sont dits *ouverts*: ils peuvent ajouter ou modifier certaines règles de manière à s'adapter à des contraintes imprévues. C'est pourquoi les systèmes adaptatifs sont aussi dits *auto-organisés*. Un des objectifs de la cybernétique est de déterminer la raison pour laquelle le contrôle des systèmes naturels est si versatile et si adaptatif, alors que le contrôle des systèmes conçus et fabriqués par les humains est si déterministe.



Le terme " rétroaction " équivaut au mot anglais « feedback ». Il renvoie à l'effet que les actions du système ont sur la poursuite de ces actions, via l'information qui revient au système sur les effets de ces actions. Une *rétroaction positive* fait s'accroître les réactions du système qui s'alimentent les unes des autres. Il peut alors y avoir emballement, ce qui peut du coup rendre le système instable. En revanche, une *rétroaction négative* est plus naturelle, car son but est de conserver le système équilibré et stable. Lorsque le système est dans un état extrême, le mécanisme de rétroaction combat cet état pour restaurer l'équilibre. Dans ce genre de contexte qui fait suite à un événement ayant transformé l'environnement, le mécanisme de contrôle mesure d'abord l'écart entre l'état actuel et l'état normal, puis établit le correctif nécessaire. Ensuite, il remesure l'écart résiduel et ainsi de suite.

Formellement, une boucle de rétroaction négative prend place lorsqu'un signal représentant le degré d'écart avec une certaine norme est réinjecté dans le système, de façon à rétablir l'équilibre initial, en s'opposant au stimulus original. Le système n'est au repos que lorsque le degré d'écart est égal à zéro. Une caractéristique importante des systèmes à rétroaction négative est le délai dans la réponse aux changements. Si le délai est trop long, le système peut se mettre à osciller de plus en plus, au lieu de réduire l'écart par rapport à la norme de départ. L'*homéostasie* désigne la capacité d'un système à maintenir ses différents états à l'intérieur d'une marge de variation, définie par les conditions limites dans lesquelles son organisation cesserait d'être efficace, voire d'exister.

Cette distinction entre systèmes déterministes et systèmes adaptatifs peut être traduite dans le domaine des environnements immersifs. Alors, les ordinateurs et les utilisateurs sont considérés comme des systèmes interdépendants faisant partie d'un même environnement. Mais cet environnement n'est pas celui où l'interaction des humains et des machines prend place. Cet environnement est représenté au moyen de l'audiovisuel numérique. Lorsque l'humain est seul avec l'ordinateur, leurs interrelations sont encadrées par un programme — un ensemble fermé de règles — qui détermine en même temps les interactions et les événements possibles dans le



monde de référence; bref, le registre des comportements est prédéterminé. Lorsque plusieurs agents humains sont reliés entre eux par des ordinateurs, le contrôle est distribué entre l'ordinateur et chaque agent humain. L'ordinateur spécifie et administre les contraintes du monde de référence où se tiendront les interactions entre les humains, et informe chacun d'entre eux des transformations que provoquent leurs interactions dans le monde de référence, par le biais d'une représentation 2D ou 3D.

5.2.4 L'échange d'information

Un système est en relation constante avec son environnement. Il en va de même des sous-parties du système, elles aussi constamment en relation. Ces relations réciproques s'actualisent à travers des échanges d'information sur les propriétés concrètes et abstraites qui constituent l'intrant et l'extrant de la régulation. La cybernétique étudie la *communication des signaux* à la lumière de la *théorie de l'information* développée par Claude Shannon. Celle-ci s'est vue formulée dans un article intitulé *The Mathematical Theory of Communication*, paru en 1948. Dans cet article, Shannon se pose la question de savoir à quel degré de précision peut être transmise l'information, en dépit des effets la perturbant, appelés *bruits*. Dans sa théorie, il élabore une définition quantitative, purement objective, de la notion d'information, perçue comme une quantité reliée à la nature statistique du signal. Sur cette base, il devient possible d'analyser les problèmes relatifs à la manipulation effective de l'information par les systèmes physiques. La cybernétique s'intéresse plutôt au flot de la communication intersystème. La communication d'un système implique de cinq éléments de base : 1) la source du signal ; 2) un processus d'encodage ; 3) un média de transmission ; 4) un récepteur et 5) un décodeur.

La cybernétique s'intéresse aux transformations auxquelles les informations sont soumises, mais en faisant complètement abstraction de la nature de ces signaux. Elle cherche à modéliser les opérations logiques qui président aux échanges. Un échange compte normalement trois phases: 1) une *phase d'émission* pendant laquelle le message est encodé ; 2) une *phase de transmission* à un autre système, directement, ou par propagation, au travers d'un réseau de systèmes reliés de



proche en proche ; 3) une *phase de réception* pendant laquelle le message est décodé. Dans certaines circonstances, d'autres phases peuvent s'intercaler : stockage, filtrage, etc.

Comme la cybernétique prédit les opérations de systèmes complexes (au moyen de leur distribution statistique) sans identifier complètement les facteurs opératoires en jeu, un facteur d'incertitude doit toujours être pris en compte. La théorie de l'information permet de mesurer cette incertitude à partir d'une propriété appelée « entropie » (cf. section ...). Le concept d'entropie est d'abord et avant tout défini par une expression mathématique de la *Deuxième loi de la thermodynamique*, une des lois fondamentales de la science. Dans son sens physique, ce concept s'applique aux systèmes dont les parties infinitésimalement petites n'ont pas à la même température. A chaque fraction d'intervalle de temps, une très petite quantité d'énergie change de place, jusqu'à ce que la température de tous les individus soit égalisée. L'entropie, c'est la mesure de ce changement. L'augmentation de l'entropie correspond à une décroissance de l'énergie disponible.

Dans la sphère de l'information, l'augmentation de l'entropie implique une décroissance de l'ordre du système, et cet ordre est associé à la connaissance. Plus la connaissance d'un système est détaillée, moins il reste d'incertitude sur son fonctionnement, et donc moins l'entropie sera grande. L'information est donc en relation directe avec *l'entropie négative*. L'entropie négative constitue non seulement une mesure de l'organisation des systèmes physiques inorganiques (pour lesquels la théorie de la thermodynamique a été développée), mais s'applique aussi aux systèmes biologiques et sociaux, et plus que par une simple analogie.

C'est pourquoi Norbert Wiener affirme que l'homme vit dans un monde où subsistent quelques îlots organisés au milieu d'un océan de désordre, et que s'il survit, c'est grâce à sa capacité à organiser et à faire circuler l'information dont il dispose. Cet échange d'information doit être *récuratif* et *redondant*. Cela permet non seulement d'éviter la désagrégation du système dans des conditions d'interférence ou de bruit, mais aussi d'augmenter le savoir.



Pour finir, dressons une liste partielle des facteurs générateurs de bruit dans les échanges humains : 1) la nature symbolique du contenu ; 2) le contexte culturel et social ; 3) les ellipses des échanges interpersonnels, nécessaires puisque toutes les références d'un énoncé ne peuvent être nommées ou spécifiées ; 4) la créativité et l'ouverture de la langue naturelle pour désigner des réalités ou des idées différentes ou nouvelles et 5) les affects émotifs individuels.

5.2.5 L'auto-organisation, cybernétique de second ordre

Un groupe de cybernéticiens regroupés autour du physicien H. Von Foerster au *Biological Computer Laboratory* se sont concentrés, durant les années 1960 et 1970, sur l'étude des systèmes cognitifs, c'est-à-dire des systèmes capables d'auto-organisation. Le titre de leur publication *Cybernetics of Cybernetics* (1976) instille l'idée d'une cybernétique de second ordre ou d'une méta-cybernétique. L'idée était d'observer une fonction cognitive, de la modéliser et de réaliser un mécanisme artificiel qui l'implémente et la simule. Pour ce faire, il fallait non seulement penser le système observé, mais aussi le système que forment ensemble l'observateur et le système observé. Etait ainsi introduit le thème central de la cybernétique de second ordre : le bouclage d'un système sur lui-même. C'est ainsi que la cybernétique du second ordre intègre les théories de l'« autopoïèse », qui signifie l'auto-organisation d'un système clos et bouclé sur lui-même, proposées par le biologiste H. Maturana, et poursuivies par F. Varela.

L'auto-organisation se définit comme la réalisation et le maintien d'une structure complexe dans un environnement riche en désordre. L'auto-organisation, c'est le renforcement des relations définitoires d'un système, soit par la transformation de relations existantes ou, éventuellement, par l'apport de nouvelles relations (ce qui a pour effet de doter le système de propriétés nouvelles). Ainsi, un système peut acquérir de l'information par ses interactions avec l'environnement, et élaborer cette information de façon à se doter de nouveaux mécanismes de rétroaction qui lui assure une plus grande stabilité.



Un système qui a la capacité de s'auto-organiser peut répondre à une gamme plus étendue de perturbations internes ou externes, de façon plus optimale, en diminuant les écarts par rapport à la position d'équilibre ou le temps nécessaire pour assurer sa stabilisation, etc. L'étude des êtres vivants exige que l'on puisse définir des systèmes capables d'une part de se constituer à partir de systèmes plus élémentaires, et d'autre part de se désintégrer, soit sous l'action de certaines perturbations qui dépassent le seuil d'admissibilité, soit sous l'effet de l'évolution normale du système, qui tend normalement à vieillir.

La notion de *récurtivité* est employée pour modéliser la *réflexivité* et la *complexité*. Alors que l'*itérativité*, à la base des automates à états finis, désigne la possibilité de réappliquer une même opération n fois à son propre résultat, la *récurtivité* est une réapplication à soi-même de l'opération jusqu'à l'atteinte d'une condition d'échappement. Un système *itératif* est un système générique qui applique un programme un certain nombre de fois ou sur une suite de données ; il est *toujours égal à lui-même*. Un système *récurtif* se comporte comme une machine non seulement capable d'exécuter un programme, mais aussi de *modifier* celui-ci jusqu'à le changer en cours d'exécution. Pour doter la machine de la capacité à s'auto-organiser, il faut lui apprendre, en fonction d'un but donné, à identifier des objectifs, les moyens pour y parvenir, et à comparer les objectifs atteints aux objectifs poursuivis pour ajuster les moyens déployés en conséquence. Le problème est que pour réaliser une telle machine, il faut savoir calculer la fonction de fonction, qui sert en quelque sorte de méta-programme contenant les règles d'une telle auto-adaptation. Là comme dans d'autres cas, la réalisation d'un tel système doté de caractéristiques similaires à la cognition humaine n'est possible qu'en fonction d'une fermeture du monde de référence à modéliser.

5.2.6 Le connexionnisme

En plein effervescence durant les années 1980, le connexionnisme s'inspire de la cybernétique et de certains de ses prolongements (dont le *Perceptron* de Frank Rosenblatt), ainsi que des développements des neurosciences. Avant de constituer une position épistémologique, le connexionnisme est un domaine de recherche et de



développement autour d'un outil bien précis, le *réseau de neurones formels*. Ce genre de réseau permet de simuler certaines fonctions cognitives, telles que la *classification de formes* et la *reconnaissance de la parole* et de *l'écriture*. Le connexionnisme repose sur une approche de la cognition fondée non plus sur l'inférence, le calcul logique de symboles opéré à partir de règles, mais sur la structure du cerveau humain, de mieux en mieux mise à jour par les neurosciences. Ce type de modélisation privilégie l'aspect structurel au mode de représentation.

Le réseau de neurones formels est constitué d'un ensemble d'automates très simples, inspirés de McCulloch et Pitts (1940), interconnectés. Les connexions permettent à un automate de transmettre à d'autres automates une stimulation positive — et donc excitatrice — ou négative — et donc inhibitrice. L'état d'activité d'un automate, excitation ou inhibition, se propage aux autres automates avec lesquels il est relié, et ce proportionnellement aux forces de liaison entre neurones. Ces forces sont appelées *poids synaptique*. La sommation des forces de liaison qui **constitue** ce poids peut être réalisée suivant plusieurs lois. La plus simple est la loi de Hebb : lorsque deux unités sont excitées simultanément, la force de liaison entre ces unités se trouve augmentée.

Les automates (en général identiques) sont des *automates à seuil*, c'est-à-dire capables de comparer la somme pondérée des stimulations par rapport à un seuil, et de se mettre en état d'activité si ce seuil est dépassé, ou sinon à s'éteindre. Le système est donc caractérisé, à chaque étape de son évolution, par un *vecteur d'activation* ; la transition d'une étape à la suivante résulte d'une mise à jour des valeurs d'activité, soit par tous les automates simultanément, soit par un seul.

Le processus commence par l'imposition d'un certain vecteur d'activation, qui peut se définir comme la donnée initiale. Il se poursuit par l'application itérative de la règle de transition, et se termine lorsque le système atteint un équilibre caractérisé par un vecteur particulier. L'identité du réseau, au cours de ce calcul, est préservée sous la forme du *vecteur des poids synaptiques*, qui constitue en quelque sorte la « connaissance » emmagasinée dans les connexions du réseau, et qui, en présence de données, en influence le comportement.



Ces réseaux ne sont pas programmés, mais sont dotés d'une capacité d'apprentissage. Ils présentent des propriétés remarquables : un réseau est capable, dans des conditions favorables, de généraliser, d'extraire la tendance moyenne de multiples échantillons différents afin de retrouver les types caractéristiques de l'ensemble des échantillons, ou de restaurer des données fortement altérées ou tronquées.

Chaque unité est autonome et ne propage que son influence localement, ce qui explique le caractère « massivement » parallèle des opérations. L'influence qu'exerce une unité peut certes être interprétée comme une sorte d'inférence, mais si elle en est une, il s'agit seulement d'une inférence plausible, et non certaine : cette influence est modulée et peut être soit renforcée, soit bloquée par d'autres influences. L'état stable correspond à un équilibre, et non à un arrêt.

Quant au mode de représentation des réseaux neuronaux artificiels, il n'est pas moins original que leur mode de calcul. Un système d'intelligence artificielle classique stocke en mémoire des expressions quasi linguistiques, interprétables comme des descriptions d'objets, de faits ou de règles. Selon les besoins du programme, il les consulte, les reproduit ou les transforme conformément aux règles formelles. Un réseau connexionniste, en revanche, ajuste ses poids synaptiques et capte ainsi les régularités et les aspects saillants de l'environnement. Celui-ci n'est donc pas *décrit* en termes langagiers (ce qui le rendrait lisible), mais *reflété* par le système à travers des vecteurs de poids synaptiques, qui sont proprement illisibles. Les représentations sont dites *distribuées*, parce que chaque concept correspond à l'activation d'un grand nombre d'unités, et non pas d'une seule ; inversement, chaque unité contribue à la représentation de plusieurs concepts.

Dans la plupart des versions de réseaux massivement distribués, c'est-à-dire de type PDP (*Parallel Distributed Processing*), l'architecture des réseaux comporte, selon D. Rumelhart et F. Fogelman, des couches à plusieurs niveaux d'interconnexion : les neurones formels reçoivent des inputs d'une couche inférieure et les renvoient vers une couche supérieure, sans bouclage. Ce type de réseau



permet l'*apprentissage supervisé*: en modifiant les poids des connexions, on peut faire en sorte que le réseau réponde à un vecteur input, constitué de la série de tous les inputs reçus de la couche inférieure, par un vecteur output donné et fixé d'avance. Le réseau sert ainsi de classificateur. Comme les connexions sont représentables par une matrice, les « catégories » du réseau seront les « valeurs propres » liées aux vecteurs inputs.

Des réseaux de neurones d'architecture différente ont été développés à partir des travaux de J. Hopfield, biologiste de la BELL Telephone. Parus en 1982, ces travaux exposent les lois fondamentales de la biologie théorique : la recherche de l'énergie minimum, la recherche d'états stables avec autocorrection d'erreurs et reconstitution de données manquantes. En fait, ce que fait Hopfield dans ce modèle, c'est mettre en œuvre les lois de l'auto-organisation des systèmes de la cybernétique du second ordre. Dans les versions de réseaux de neurones à la Hopfield, les neurones sont tous bouclés les uns sur les autres. Partant de plusieurs vecteurs inputs, le réseau se stabilise dans un même état, appelé « attracteur », ce qui permet d'associer plusieurs inputs différents. Considéré dans son ensemble, le réseau remplit certaines fonctions cognitives, ou plutôt les simule : il reconnaît, classifie, associe sans qu'aucun de ses éléments ne soit programmé à cet effet. Ces réseaux sont particulièrement résistants à la dégradation; on peut en détruire certains éléments sans endommager les fonctions du réseau. De plus, ils sont très tolérants au bruit : on peut déformer l'input, et même si la déformation est considérable, il sera quand même reconnu.

Dans tous les cas, l'approche connexionniste veut combler l'hiatus entre le *fonctionnement des neurones* et la *production de représentations symboliques*, problème laissé de côté par le cognitivisme classique (cf. 4.4.2). Mais il y a un problème de taille: les réseaux connexionnistes ne sont pas capables, a priori, de conserver la structure de l'information qu'ils traitent, ou du moins d'en tirer parti pour des transformations inférentielles autres que l'apprentissage. Les transformations inférentielles consistent, dans certaines conditions, à associer explicitement un individu A, nœud ou neurone, à un individu B.



Dans les environnements immersifs, les réseaux neuronaux sont surtout utilisés pour les fonctions de perception: ils tirent parti de leur capacité de filtrage et de catégorisation dans des situations réelles, où les stimuli sont déformés, où il y a du bruit et des interférences. On retrouve notamment des réseaux neuronaux dans les modules de reconnaissance du visage dans une image à partir des yeux, dans les modules de reconnaissance des gestes et d'autres mouvements corporels, à partir d'images vidéo, ainsi que dans les modules de reconnaissance de la voix.

5.2.7 Les neurosciences

Les neurosciences s'intéressent à la *réalisation biologique de la cognition*, et en particulier aux processus perceptifs et moteurs. Leur objectif est de caractériser le système nerveux en tant que système physico-chimique. Au cours des années 1980 et 1990, les neurosciences ont donné lieu à des développements importants parce que les techniques d'analyse du fonctionnement cérébral s'amélioraient constamment. Notamment, le *scanner à positron* et la *résonance magnétique nucléaire*, techniques électroencéphalographiques de plus en plus précises, permettaient d'enregistrer le fonctionnement du cerveau. Multidisciplinaires par nature, les neurosciences tentent d'intégrer, aux plans méthodologique, conceptuel et finalement théorique, d'une part les disciplines traitant des différents aspects de la structure, et du fonctionnement du système nerveux, et d'autre part les sciences de la cognition, qui étudient les principes et les règles de fonctionnement des systèmes de traitement de l'information, aussi bien humains qu'artificiels. L'objectif des neurosciences, c'est de comprendre les processus grâce auxquels le cerveau humain prend connaissance du monde, mémorise et transforme ces connaissances sous forme de représentations, et les utilise pour planifier et mener les actions qui permettent à un individu de modifier son environnement.

Les neurones sont des cellules excitables, capables de convertir des signaux chimiques en activité électrique et réciproquement. Comme la plupart des cellules, ils possèdent un corps cellulaire, qui regroupe le noyau et toute la machinerie nécessaire à la synthèse des protéines et à leur métabolisme. Les neurones se distinguent des autres cellules par la présence de prolongements plus ou moins



ramifiés, appelés *neurites*. Ceux-ci se répartissent en deux classes fonctionnelles : un *axone* et des *dendrites*. Les dendrites reçoivent les signaux électriques et chimiques provenant de neurones qui y sont connectés. À l'inverse, l'axone transporte des signaux électriques vers les autres neurones, sur des distances qui peuvent atteindre des dizaines de milliers de fois la taille du corps cellulaire. Ce sont les axones qui contactent les dendrites d'autres neurones pour constituer des structures spécialisées, appelées *synapses*. Un neurone peut compter plusieurs milliers de synapses sur ses dendrites, et son axone peut en former tout autant. Le cerveau humain est constitué d'environ 10^{11} neurones, de plusieurs centaines ou milliers de types différents, connectés par approximativement 10^{14} contacts synaptiques. Schématiquement, le cerveau est composé d'un grand nombre d'éléments interconnectés qui communiquent entre eux par des messages très simples, de nature excitatoire ou inhibitoire. L'approche neuroscientifique explore la dynamique de la répartition topologique des changements de l'activité cérébrale.

L'organisation du système visuel a fait l'objet d'une attention particulière de la part des neurosciences, et plus particulièrement les processus d'encodage sensoriel et d'élaboration des représentations perceptives. Ainsi, on a étudié le temps de réponse à différents stimuli visuels, la capacité à identifier un objet dans son contexte, la perception des couleurs et ses déficiences, la vision périphérique, la sensibilité aux contrastes, aux mouvements, la perception de la profondeur, l'appréciation des distances et des angles, etc. Les neurosciences dégagent deux principes de base de ces études. Premièrement, les différents attributs sont traités de façon *modulaire* : forme, couleur, mouvement, etc. des stimuli visuels sont pris en charge par *différents sous-ensembles* spécialisés de neurones. Deuxièmement, la synthèse perceptive résulte de la convergence et de l'intégration de processus *menés en parallèle*. De multiples aires corticales, aux compétences différenciées, collaborent au même traitement de manière différente, et se combinent en une architecture fonctionnelle. Les associations multisensorielles, le rappel à la mémoire, la projection d'un schème interprétatif, etc. sont des exemples de ce genre d'intégration convergente de tâches menées en parallèle.



Au niveau sensoriel précoce, le système visuel est spécialisé dans le traitement de la lumière, c'est-à-dire de signaux électromagnétiques. Deux ensembles de récepteurs assurent la transduction. D'une part, il y a les bâtonnets, sensibles à de faibles éclairages, et d'autre part, il y a les cônes, activés par des éclairages plus importants. On distingue trois classes de cônes suivant la longueur d'onde maximale qu'ils peuvent absorber. Au niveau cortical, certains neurones répondent sélectivement à l'orientation spatiale des contours, d'autres sélectionnent la couleur, d'autres encore la direction et le sens du mouvement.

5.2.8 La vie artificielle

La *vie artificielle* est une discipline qui s'est vue accorder ce titre lors d'une conférence (la première consacrée à ce thème) tenue à Los Alamos en 1987. Elle s'intitulait *Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*. Le programme de recherche fut défini par C. Langton comme *l'étude de systèmes construits par l'homme, et qui se comportent à la manière des systèmes vivants*. Le programme est ambitieux, mais bien défini, et dispose de bases théoriques et conceptuelles des plus variées. Le postulat de la vie artificielle, c'est que le contenu informationnel de la vie peut être séparé de son substrat matériel. Il part du constat que les *techniques computationnelles* sont structurellement similaires aux *structures profondes de la vie biologique*. En proposant de séparer la matière et l'information dans les systèmes biologiques ou dans les organismes, la vie artificielle reproduit la dichotomie augustinienne entre le corps et l'âme. Cette dichotomie sera renforcée par la pensée des Lumières, Descartes en tête, à travers la dichotomie corps — esprit, ce qui légitime et renforce d'autres structures dualistes qui sous-tendent notre appréhension du monde : forme — fond, signifiant — signifié, etc.

La vie artificielle hérite ses outils, entre autres, du domaine de la dynamique non-linéaire et de la théorie de la complexité, ainsi que de la cybernétique du second ordre : l'organisme se définit comme un système qui échange de l'information, à l'instar d'un être vivant. La vie artificielle se fonde sur les principes de l'auto-organisation exposés plus hauts. Elle se base entre autres sur la récursivité: le



processus n'est plus *linéaire* — input -> traitement -> output — mais *circulaire* : action -> réaction. Autrement dit, le comportement du système n'est plus conditionné par son input, comme dans le cas du schéma linéaire, mais bien par son organisation interne. La vie artificielle se fonde ainsi sur les principes de l'autopoïèse, du grec *autopoiein*, qui signifie littéralement la " (re)production de soi-même ". Le processus caractéristique de la vie artificielle est *néguentropique*, c'est-à-dire qu'il se définit comme un processus inverse à la production de désordre ou de désorganisation: il engendre et spécifie continuellement sa propre organisation. L'autopoïèse, c'est la capacité dont est doté un système capable de se re-produire, et de re-produire son identité en dépit des changements de l'environnement.

La vie artificielle intègre aussi des idées originaires de l'étude du chaos, et principalement la notion d'*émergence*. Dans le domaine des sciences physiques, le terme " chaos " s'est appliqué à tout phénomène qui semble n'obéir à aucune loi, et donc impossible à prévoir. Dans son sens populaire, « chaotique » est synonyme de « aléatoire ». En physique, le terme se rapporte à une classe de phénomènes bien définie, où l'imprédictibilité est certes présente, mais sans empêcher l'existence d'un ordre sous-jacent, déterminé par des « attracteurs ». Les interactions locales entre des éléments simples autour des attracteurs produisent, à l'échelle globale, un comportement plus complexe que les interactions qui le produisent. C'est pourquoi l'on parle parfois de *subsomption* pour qualifier la relation des éléments à l'architecture globale. Ainsi, la complexité apparaît spontanément comme le résultat de l'application de règles simples. À chaque niveau de l'évolution prébiotique, biotique et sociale, des propriétés nouvelles émergent à cause de la diversité des éléments ; l'accroissement du nombre de liaisons entre ces éléments et le jeu des interactions non linéaires conduisent à des comportements difficilement prédictibles.

La vie artificielle s'attache à modéliser certains comportements ou certains mécanismes primordiaux des systèmes vivants, tels que l'*auto-régulation*, la *prédation* ou la *reproduction*. La vie artificielle consacre une grande partie de ses efforts à réaliser des organismes virtuels, c'est-à-dire à simuler des créatures vivantes à l'aide des langages formels de l'informatique. On peut en gros distinguer



deux grandes catégories de créatures artificielles: les *organismes bio-morphiques* et les *organismes anthropomorphiques*. Dans le premier cas, ce sont les principes fondamentaux de la vie qui sont modélisés et simulés, par exemple à travers les automates cellulaires et les algorithmes génétiques. Dans le second cas, les organismes modélisés et implantés sont dotés de fonctions supérieures qui s'apparentent à celles de l'humain, au plan comportemental ou cognitif ; les réseaux neuronaux (cf. 5.2.6) sont principalement utilisés à cet effet.

Les *automates cellulaires* remontent aux années 1940. Stanislas Ulam, mathématicien, est l'un des premiers à en avoir mis au point. Ulam s'intéressait à l'évolution de constructions graphiques générées à partir de règles simples. La base de ses constructions était formée d'un espace à deux dimensions divisé en « cellules », formant ensemble une sorte de tableau. Chaque cellule pouvait être dans deux états différents, allumé ou éteint. Partant d'une configuration donnée, la génération de cellules suivante était déterminée en fonction de règles de voisinage. Exemple de ces règles: si une cellule est en contact avec deux cellules allumées, elle s'allume, sinon elle s'éteint. Ainsi, le nouvel état d'une cellule ne dépend que de son état actuel et de l'état de son voisinage. Ces règles sont universelles, communes à l'ensemble de l'espace de l'automate cellulaire. Les automates cellulaires fonctionnent de manière discrète, c'est-à-dire que le temps s'écoule par à-coups. À la génération t , chaque cellule examine son environnement et détermine son état futur en fonction de ce genre de règles simples. Lorsque l'ensemble des cellules a été traité, et seulement à ce moment-là, toutes les cellules adoptent l'état calculé. De cette manière, un traitement simultané en parallèle est simulé.

Les *algorithmes génétiques* sont des mécanismes mathématiques simples. Ils génèrent une séquence d'événements qui fait évoluer le système par un principe de sélection naturelle. Un algorithme génétique agit comme un mécanisme non-linéaire de rétroaction, qui change graduellement les paramètres d'un système vers un état optimum. Ce mécanisme de rétroaction est régi par les règles d'adéquation (*fitness*) est non linéaire, parce que le principe qui contrôle la boucle de rétroaction n'est pas basé sur une moyenne ou sur un autre principe mathématique progressif ou



dégressif, mais généré aléatoirement. Une autre branche de la vie artificielle est consacrée à la conception de robots autonomes dont l'aspect et le comportement sont analogues à ceux des animaux ou des humains. D'autres réalisations matérielles utilisent les matériaux mêmes des êtres vivants. S'inspirant de la biologie moléculaire, la vie artificielle tente d'instrumentaliser la sélection naturelle, procède à des opérations cellulaires telles que l'hybridation, à la construction par la réplication de sa propre structure, ou encore procède au regroupement de protéines, d'enzymes et d'acides nucléiques afin de constituer des organismes cellulaires. La génétique, les technologies de la reproduction et nanotechnologies participent ainsi au projet de la vie artificielle.

5.2.9 Les cyborgs

Cyborg est le nom de l'approche la plus radicale de l'interface entre le corps et les machines. Comme les cybernéticiens approchent les *systèmes mécaniques* et les *systèmes vivants* à partir des mêmes concepts, à savoir les notions liées à la machine à traitement de l'information, la conception d'organismes hybrides devient tout à fait pertinente. La technologie du cyborg relie de manière très intime la machine et l'individu, de sorte que finalement, elle peut être utilisée pour altérer le corps. Loin de constituer une discipline organisée et systématique, cette technologie se nourrit 1) de *prothétique*, c'est-à-dire du domaine de la construction d'organes artificiels, nécessitant ou non un système de commande pour l'interface vivant — artificiel ; 2) de la *microchirurgie des implants* ; 3) des *neurosciences*, rebaptisées *wetware* dans les milieux branchés (tiré du titre d'un roman de Rudy Rucker, *wetware* désigne le support physique de la cognition, comme le « hardware » s'oppose au « software »); 4) de la *chirurgie esthétique*, reconstructive et même constructive et 5) de la *robotique*, c'est-à-dire du domaine qui traite de la construction d'automates physiques évoluant dans l'environnement réel. Le concept de cyborg se nourrit aussi, et peut être avant tout, de l'imaginaire futuriste de la science-fiction.



« Cyborg » est une contraction de *cybernetic organism* due à Manfred Clynes et Nathan S. Kline. Ceux-ci précisent qu'en tant que systèmes autorégulés, les cyborgs doivent fonctionner sans recours à la conscience de façon à coopérer avec les contrôles homéostatiques autonomes du corps humain. En effet, tant que les systèmes artificiels sont contrôlés par la conscience des systèmes humains, la distance entre les deux demeure entière. En d'autres mots, il y a cyborg, c'est-à-dire extension artificielle d'un système vivant, lorsque cette extension est sous le contrôle homéostatique autonome inconscient du système vivant.

Selon les thèses évolutionnistes de Charles Darwin, la nature trouve différents moyens biologiques d'adaptation pour faire face aux problèmes que présente l'environnement. Les organismes qui survivent sont ceux qui ont su faire preuve du plus grand degré d'adaptation, et qui ont ainsi évolué en surmontant les obstacles. L'idée est que par le cyborg, les humains prennent activement part à leur propre évolution biologique, sans attendre les altérations naturelles de l'hérédité, en induisant artificiellement certaines altérations en modifiant biochimiquement, physiologiquement et électroniquement tout ou partie du corps humain.

Parmi les réalisations les plus probantes de cette proto-discipline, il faut citer la création d'un réseau neuronal bio-computationnel par Babatunde Ogunnaike et Jim Schwaber. Ce réseau a été créé à partir d'une mince tranche de la cervelle d'un rat, couplée aux processeurs d'un ordinateur qui simulait adéquatement certains contrôles homéostatiques d'un sous-système lent (la contraction artérielle) et d'un sous-système rapide (l'effet de pompage du cœur).

Le cyborg fascine la sphère des arts. Stelarc, alias Stelios Arcadiou, est le chef de file et le principal acteur de *l'art corporel cybernétique*. Dans ses performances, il utilise des technologies médicales et prothétiques sur son propre corps, de manière à attirer l'attention sur la nécessité d'une évolution participative et auto-gérée de son corps. De même, par ses performances de chirurgie plastique, Orlan cherche à transgresser les frontières du corps et d'avancer vers l'idée d'une forme construite par soi-même. Orlan emprunte le menton de la Vénus de Boticelli, les lèvres de l'Europa de Boucher, le nez de la Psyche de Gérôme et le front de la Mona Lisa. Elle



écrit un discours avec et sur son propre corps, par la mutilation sélective et par la vivisection, quête qui la mène à devenir physiquement ce qu'elle est intérieurement.

5.3 Composants d'un dispositif d'immersion

C'est à la fin des années 1980 que sont apparus les premiers prototypes fonctionnels de dispositifs immersifs. Dans cette section, on a choisi de présenter, à titre d'exemple, la station graphique d'environnement virtuel VIEW (*Virtual Environment Workstation*), développée au Centre de recherche Ames, rattaché à la division des facteurs humains de l'aérospatial de la NASA, l'agence spatiale américaine. Cette station est constituée d'un casque qui contient un système de visualisation stéréoscopique à grand angle. Ce système est contrôlé par la position de la tête, par la voix et par les gestes de l'opérateur. Ce dispositif est conçu comme une interface multimodale avec l'environnement, mieux adaptée aux capacités sensorielles et cognitives des humains que les interfaces habituelles en 2D. Avec ce dispositif, l'utilisateur est non seulement en mesure d'explorer les 360 degrés d'un environnement de synthèse ou réel (recensé à distance), mais aussi d'intervenir dans cet environnement.

5.3.1 La poursuite de la position et de l'orientation

La *technologie électromagnétique* est le dispositif qui permet de déterminer de manière spécifique la position et l'orientation dans l'espace d'un objet, que ce dispositif s'incarne dans un casque de visualisation ou dans un gant de données. A la base du système Polhemus, cette technologie est composée de trois éléments : 1) une *source magnétique* ; 2) un *capteur*, qui génère un courant électrique variant suivant la position et l'orientation de l'utilisateur par rapport à la source du champ magnétique ; 3) un *boîtier de contrôle*, qui contient le circuit électronique permettant de numériser les différentes valeurs captées et de les transmettre à l'ordinateur. Comme le champ magnétique est de faible intensité, l'utilisation de ce système est problématique dans des espaces dépassant le mètre cube. Ce système enregistre des informations sur la tête de l'utilisateur : son élévation (en termes d'azimut), son mouvement de rotation, sa position suivant les axes de la hauteur,



de la largeur et de la profondeur. L'ordinateur utilise ces informations pour générer les images de l'environnement en fonction des changements de point de vue détectés.

5.3.2 Casque de visualisation

Le *casque de visualisation* (ou *visiocasque*) sert de module d'affichage. Il est placé directement devant les yeux de l'utilisateur, et a le même volume qu'une paire de jumelles. Comme pour les jumelles et les lunettes, le module est divisé en deux écrans couleurs à cristaux liquides, à la résolution élevée. Dans le cas du visiocasque, la division en deux écrans ne sert pas à ajuster la lentille à chaque oeil, mais à créer un effet stéréoscopique de perception tridimensionnelle par un léger décalage entre les deux images projetées. De cette manière, l'effet de présence se voit augmenté (cf. 5.6). Une lentille s'intercale entre les yeux et les écrans. Sa focale est un grand angulaire, ce qui permet de s'approcher des capacités visuelles humaines et de remplir entièrement le champ de vision.

Il est nécessaire d'appliquer un modèle optique pour créer des images dont l'orthostéréoscopie soit correcte. L'orthostéréoscopie, c'est l'absence de déformation dans une image lorsqu'il y a vision stéréoscopique (qui rend la troisième dimension visible). Ce modèle de correction consiste à appliquer une distorsion aux pixels sur l'écran. De plus, l'utilisateur doit adapter sa vision normale pour pouvoir visualiser l'écran. En effet, l'affichage stéréoscopique fournit la convergence des deux yeux, mais pas l'accommodation, alors que ces fonctions sont habituellement solidaires. L'utilisateur doit du coup apprendre à contrôler séparément son cristallin, qui forme une lentille biconvexe. Sa plasticité permet d'en modifier la courbure, ce qui rend possible l'accommodation, c'est-à-dire le déplacement du point focal, qui donne la vision la plus nette, en avant-plan du champ visuel.

Le casque est doté d'un capteur Polhemus. Celui-ci fournit à la station graphique la position et l'orientation du casque. L'ordinateur fournit les images qui correspondent au point de vue qu'adopte l'utilisateur sur le monde représenté. Il existe une stratégie encore plus précise de localisation du point de vue de l'utilisateur, qui



consiste à coupler ce système à un dispositif de poursuite de la position de la pupille de l'utilisateur. Cette stratégie permettrait en outre d'insérer un curseur à l'endroit où se pose son regard dans l'environnement 3D.

5.3.3 Gant de données

VPL Research Inc a développé en 1985 un gant de donnée qui permettait d'interagir avec les objets inclus dans l'espace 3D. Ce gant est doté d'un capteur Polhemus, donc du même type que le capteur du casque de visualisation. Ce capteur permet de déterminer *et* la position absolue de la main, *et* son orientation par rapport aux trois dimensions de l'espace de représentation. Des fibres optiques sont insérées sur le dessus de chaque doigt du gant. Une diode est montée au bout de chaque fibre optique, et un phototransistor convertit le niveau de la lumière en courant électrique, dont le voltage diffère selon les flexions de chaque jointure des doigts, ce qui permet de déterminer approximativement l'angle de flexion. Enfin, des senseurs permettent de déterminer le niveau de la pression qu'exerce l'utilisateur qui manipule un objet virtuel inscrit dans l'environnement immersif.

Cela dit, la main a beau essayer de saisir un objet dans l'espace virtuel, elle va passer au travers. On a tenté de contourner ce problème de plusieurs manières; par exemple, un signal sonore est émis lorsqu'il y a contact entre la main et l'objet. En dotant les gants d'un dispositif de feedback tactile qui fournit un retour d'effort, notamment par des suites de très petits actionneurs en selenoid, on pourrait retourner à l'utilisateur un feedback sur la solidité des surfaces qui délimitent les objets.

Les gestes de la main dont on a parlé jusqu'ici (retransmis par le gant) servent tous à *manipuler* des objets, qu'il s'agisse d'objets virtuels, ou d'objets réels manipulés à distance à l'aide d'une main robotisée. Mais il est également possible d'assigner d'autres interprétations aux gestes de l'utilisateur. Par exemple, pointer un espace sur un mur avec l'index peut revenir à le transpercer d'un rayon destructeur, alors qu'un mouvement du plat de la main, au contraire, restaure la surface. Les marionnettistes établissent déjà depuis longtemps de telles correspondances entre



les mouvements des doigts de la main et un certain éventail d'actions. Pour que l'ordinateur puisse reconnaître des gestes, il faut préalablement qu'il en ait fait l'apprentissage, le plus souvent par le biais de réseaux de neurones (cf. 5.2.6). On utilise les mêmes techniques pour reconnaître des gestes à partir d'informations cartésiennes sur les différentes parties du corps extraites à partir d'une image vidéo.

Un module de reconnaissance des gestes doit permettre de classer les mouvements et les configurations du corps en différentes catégories, sur base d'exemples auxquels le réseau de neurones a été exposé. Une fois le geste catégorisé, on peut extraire des informations sur la façon particulière dont le geste a été exécuté. La reconnaissance fonctionne à deux niveaux. En poursuivant continuellement la position et l'orientation du gant, on peut déterminer une série de configurations particulières des doigts, qui constituent ensemble une posture. Les postures captées s'accumulent en série dans un historique. A partir de cet historique, on peut isoler des trajectoires de postures, au moyen, notamment, des modèles de Markov. Les trajectoires ainsi repérées sont soumises à un sous-système sous forme de vecteurs, qui les classifie en gestes.

De tels systèmes permettent non seulement de manipuler des objets, mais ajoutent aussi une *dimension haptique* à la communication sensorielle entre la personne et l'ordinateur, jusque là strictement audiovisuelle. Le terme " haptique " (du verbe grec *haptain*) désigne la sensation du toucher. Dans les environnements immersifs, c'est tout le corps de l'utilisateur qui est répertorié, ce qui accroît d'autant l'effet de présence.

5.3.4 Environnement sonore

Au plan sensoriel, l'aspect sonore de l'environnement est tout aussi important que l'aspect visuel, privilégié jusqu'à maintenant. L'aspect sonore contribue aussi à susciter l'effet de présence, et même de manière prépondérante (cf. 5.6.3). Le casque est donc aussi doté d'écouteurs alimentés par un synthétiseur de sons en temps réel, spatialisés par la stéréophonie. Le module sonore met en jeu la faculté



qu'ont les agents humains de localiser, le plus souvent avec justesse, la source des sons qui les entourent. Ce module fournit aux utilisateurs des indications ponctuelles et dynamiques, qui viennent renforcer ou compléter les informations fournies par la voie visuelle. Parmi les indications possibles, mentionnons le signal sonore (déjà évoqué plus haut) qui indique que la main est en contact avec un objet « virtuel », le signal qui indique qu'un geste a bien été reconnu, ou encore le signal qui vient confirmer qu'une manipulation a bien été accomplie, etc.

Les paramètres auditifs sont traités de la même manière que les paramètres visuels. Il sont dynamiquement modulés et coordonnés avec les paramètres visuels. De cette manière, on représente les positions relatives de sources sonores fixes ou mobiles, et ce en fonction des déplacements de l'utilisateur dans l'environnement. Cela n'est possible que si l'acoustique de la pièce peut être modélisée. Pour chaque source sonore, l'ensemble des paramètres du champ sonore est pris en charge. Parmi ceux-ci, on compte la direction du son, sa distance virtuelle et son rapport à l'environnement, rendu par le niveau de réverbération et d'éclat par rapport à la grandeur de la pièce.

Parfois, le casque comporte également un microphone, qui, une fois connecté à un module de reconnaissance de la parole et à un analyseur morpho-syntaxique, pourrait permettre un dialogue personne-machine de type dialogique. Mais naturellement, les choses seraient bien différentes de ce qu'il se passe par l'intermédiaire du clavier dans les machines à contenu.

5.3.5 Stations graphiques

Une *station graphique* est un ordinateur doté d'un processeur dont la puissance de calcul est suffisante pour que puissent s'effectuer en tant réel des calculs géométriques et vectoriels complexes et volumineux. Les environnements immersifs sont générés par un *groupe de procédures-objets*, qui communiquent entre elles par un mode de messagerie inter-procédure. Chaque procédure, exécutée de manière constante, produit et consomme les messages asynchrones nécessaires à l'accomplissement d'une certaine tâche. La procédure centrale de cette application



gère le modèle de l'environnement, et simule son évolution en réponse aux événements générés par la modification de la posture de l'interacteur. Cette procédure est à la fois responsable de la lecture des senseurs à des fréquences prédéterminées et du déclenchement du feedback audiovisuel approprié à la modification recensée. Toutefois, l'action qui engendre la réaction est aussi source d'informations qui se propagent du senseur vers le modèle manipulé. Ces informations peuvent être modifiées par de multiples médiateurs qui s'interposent entre les senseurs et le modèle pour tenir compte des lois de la mécanique ou une logique des émotions et ainsi de suite.

Les environnements immersifs font circuler un flot complexe d'informations. Un rapport doit être établi entre les données qui sont captées par les senseurs, et leur effet dans l'environnement immersif. Ce rapport est déterminé par la nature et par les spécifications des équipements utilisés. Certaines contraintes relient directement les paramètres variables des senseurs et les paramètres de l'environnement. Plus est complexe le rapport qui s'établit entre les données que fournissent les senseurs et les actions opérées dans l'environnement immersif, plus le pouvoir d'expression de l'environnement en tant que média augmente.

Plusieurs innovations technologiques ont permis de construire et de fréquenter des environnements immersifs à partir d'un ordinateur personnel. Quelques exemples: la modélisation d'objets ou de décors par des agrégats appropriés de formes géométriques complexes qui sont ensuite recouvertes de textures, les algorithmes de codage et de compression des données de la modélisation, les algorithmes de rendu en temps réel de ces modèles, qui tiennent compte de la manière dont le point de vue est modifié par l'interaction, et enfin les cartes graphiques qui implantent les routines de base du rendu en trois dimensions sur des processeurs dédiés uniquement à l'affichage graphique, indépendants du processeur central.

Les capacités de calcul des différents ordinateurs personnels sont très disparates; ils diffèrent aussi par la plus ou moins grande rapidité de l'horloge interne, par la quantité de mémoire vive disponible et par la puissance de la carte graphique. Du coup, le rendu des environnements tridimensionnels est décomposé en niveaux de



détails: ainsi, la granularité de la géométrie, des textures, de la réflexion de la lumière et du mouvement peut varier. L'utilisateur peut alors choisir le niveau de rendu en fonction de la puissance de son ordinateur, afin de maintenir de manière constante la fréquence d'affichage des images au niveau acceptable de 30 images à la seconde.

Le *Reality Engine* est un ordinateur haut de gamme proposé par Silicon Graphic. Il a été conçu pour générer en temps réel des environnements 3D interactifs de qualité photoréaliste. Par ailleurs, le langage de modélisation de réalité virtuelle VRML (*Virtual Reality Modeling Language*) permet de manipuler graphiquement des environnements tridimensionnels sur le Web. Il existe plusieurs bibliothèques logicielles de modélisation et de rendu 3D basés sur la notion de scène et d'objets. Par exemple Java3D, proposé par SunMicrosystem, repose en gros sur deux API (*Application Program Interfaces*) DirectX pour les systèmes d'opération Microsoft Windows, et OpenGL pour tous les autres, c'est-à-dire UNIX et MacOS. Les API sont des protocoles de communication qui permettent aux développeurs d'accéder aux microprocesseurs spécialisés d'une part dans le rendu graphique 2D et 3D, et d'autre part dans le rendu sonore spatialisé. L'intérêt est qu'ils peuvent le faire sans devoir écrire du code spécifique pour une plateforme donnée. Le serveur OpenGL offre une interface pour à peu près 120 commandes distinctes à partir de primitifs géométriques : points, lignes et polygones ; il a été conçu pour fonctionner dans l'environnement UNIX, essentiellement distribué, et prend la forme client-serveur.

5.3.6 *Chambre d'immersion*

Il existe une alternative au casque de visualisation pour explorer des environnements immersifs, à savoir la *chambre de visualisation*. Le système CAVE développé au Electronic Visualization Laboratory de l'Université de l'Illinois en est un exemple. Cette chambre est composée de trois écrans de trois mètres sur trois. Des tissus tendus en constituent les murs, et derrière ceux-ci se trouvent trois projecteurs à une résolution de 1280 pixels sur 1024. Le plancher sert aussi d'écran de projection: le faisceau d'un projecteur au plafond est dirigé vers le sol à l'aide d'un miroir incliné. Ces dimensions permettent de représenter l'environnement à un



échelle 1:1 et plus, c'est-à-dire en grandeur nature, et même en très gros plan. Plusieurs utilisateurs munis de lunettes stéréoscopiques peuvent pénétrer à l'intérieur de l'environnement en même temps. Ces lunettes sont dotées d'un dispositif dynamique qui bloque alternativement l'œil droit et l'œil gauche. Elles permettent aussi de détecter la position et l'orientation de la personne. Des gants et des souris gyroscopiques référencées dans les trois dimensions peuvent aussi être utilisés dans l'interaction. Quoique très chère, cette chambre d'immersion est intéressante parce qu'elle ouvre la voie aux expériences immersives en groupe.

5.3.7 Autres capteurs

Toutes sortes de capteurs peuvent être utilisés pour recenser l'environnement ou l'utilisateur. Les capteurs se distinguent en fonction de la nature des données qu'ils peuvent saisir et transmettre : les *capteurs binaires* permettent de fermer, d'ouvrir et donc d'éteindre un circuit électrique, alors que les *capteurs analogiques* peuvent capter la variation d'un paramètre sur une échelle. Exemples de capteurs analogiques: une cellule photosensible fabrique un faible courant électrique proportionnellement à son exposition à la lumière, ou encore un senseur de proximité utilise l'effet Doppler (qu'on retrouve dans les radars, pour calculer la position des objets par rapport à la vitesse de déplacement des ultrasons émis).

Les capteurs peuvent servir à effectuer une *analyse de localisation*, c'est-à-dire la détection de présence, de proximité, de passage, de mouvement, d'orientation ou encore la mesure d'une distance, d'une vitesse, d'une accélération. Système formé de 24 satellites et de leur station terrestre, le *Global Positioning System* (GPS) permet de déterminer le positionnement d'une entité dans l'espace précis au mètre près, à n'importe quel point du globe.

Une autre analyse de localisation consiste à détecter des sources sonores avec un microphone, ou encore à détecter et poursuivre une couleur ou une forme avec une caméra vidéo. Les capteurs permettent d'analyser la variation de certains paramètres physiques mécaniques clés — le contact, la force, la pression — ou bien



la variation de paramètres physiques environnementaux — la température, l'humidité, la pression atmosphérique, le champ magnétique terrestre.

Parmi les *capteurs de l'environnement*, on trouve notamment le thermomètre et le baromètre qui mesurent la température et la pression atmosphérique, ainsi que le magnétomètre, dispositif de mesure de champ magnétique de faible intensité, etc.

Parmi les *capteurs de la personne*, il y a les capteurs de position et d'orientation corporelle dans l'espace, le plus souvent associés au casque d'immersion.

L'exosquelette, lui, est utilisé en animation. Chaque articulation qui le compose étant dotée d'un senseur, l'exosquelette permet de capter les mouvements qu'exécute une personne et d'enregistrer ceux-ci sous la forme de données structurées. Par la suite, les données cinétiques captées peuvent être appliquées à un modèle de synthèse qui acquiert ainsi une grande crédibilité perceptuelle. C'est ainsi que l'on a pu voir des dinosaures de synthèse courir et voler de façon réaliste dans *Jurassic Park* (1993), le film déjà classique de Steven Spielberg. Cet exosquelette est notamment utilisé en danse pour capter des mouvements chorégraphiques, et ensuite les faire exécuter par des danseurs virtuels.

Enfin, il y a d'autres *capteurs de gestes* que les gants de données et l'exosquelette dont il a déjà été question, à savoir le signal de caméras vidéos, utilisé pour détecter les mouvements et les expressions des personnes. D'autres senseurs permettent de capter le mouvement des yeux. Des capteurs d'ondes cérébrales permettent d'identifier certains états psychophysiologiques. Et enfin, d'autres capteurs encore mesurent les signes vitaux de l'organisme : le pouls, la pression sanguine, la tension musculaire, l'effort respiratoire, la réponse galvanique de la peau, etc.

5.3.8 Les dispositifs semi-immersifs

Les environnements immersifs sont tridimensionnels. Lorsque ces environnements sont ramenés à la surface bidimensionnelle d'un écran, ils deviennent semi-immersifs. Les objets, les personnages, et l'environnement lui-même sont rendus suivant les lois de la perspective, et à l'aide de gradients de texture. De plus, les



objets qui se trouvent plus loin de l'utilisateur sont plus petits et moins détaillés, et sont mus moins vite que ceux qui se trouvent à l'avant-plan. La lumière qui toucherait la rétine d'un individu dans la réalité est recréée dans le respect des lois de la physique. Au plan auditif, l'espace sonore est rendu par la panoramisation, réalisée à l'aide des deux canaux de la stéréophonie.

Les environnements semi-immersifs permettent l'immersion sans devoir recourir au matériel spécialisé qui vient d'être décrit. Seul un micro-ordinateur domestique ou encore une « console » de jeux est nécessaire. Dans ce genre d'environnement, il est possible d'utiliser la souris, le clavier ou encore une manette de jeux, bien mieux adaptée à ce type d'environnement. Cette manette de jeux est appelée communément *Joystick*, nom tiré du bras de commande d'un avion, et qui signifie littéralement « manche à balai ». Le *Joystick*, périphérique fétiche du joueur, permet d'augmenter la " jouabilité ". Il en existe différentes catégories: le très classique *Paddle*, genre console de jeux, le *Joystick* avec retour de force utile pour les jeux d'avion, et enfin le volant servant à la conduite des voitures.

Le *SideWinder Force feedback Pro* est actuellement la manette de jeu qui procure le maximum de sensations. Le positionnement repose sur un système optique numérique ultra précis. Les deux moteurs rendant le rendu des effets et le recentrage sont gérés par un microprocesseur inclus dans la base du *Joystick*. Il agit sur un système de 6 engrenages, qui amplifie l'effet jusqu'à 18 fois. La transmission se fait directement à la base via 4 tiges d'entraînement. La réduction au minimum des frottements, les produits lubrifiants et la légèreté du manche permettent un transfert maximal de l'effort depuis les moteurs jusqu'à la main. Un logiciel de développement permet de visualiser facilement la forme, la sensation, la fréquence, la direction, etc.

5.3.9 Le partage des environnements immersifs

Il est devenu possible d'accéder à une version partagée des environnements immersifs grâce au développement d'Internet et à sa pénétration sur les lieux de travail, d'enseignement et dans les maisons. Dans ces versions partagées, un



utilisateur détermine toujours l'environnement de départ ; les autres utilisateurs sont « invités » et n'ont besoin de fixer aucun paramètre. Toutes les données pertinentes pour l'environnement sont acheminées lors d'une première connexion sur l'ordinateur serveur de l'environnement partagé. La communication entre les joueurs est asynchrone. La modification des données, c'est-à-dire les changements de position ou d'orientation d'un personnage que provoque l'action d'un joueur sur la partie client de l'environnement partagé, est aussitôt diffusée aux autres joueurs.

EverQuest est le premier jeu partagé qui a connu un succès commercial d'envergure. Sorti en 1999, il est soutenu par la multinationale Sony. Ce jeu pour lequel des joueurs sont prêts à payer 15 dollars par mois compte 450 000 abonnés. Les mordus y passent plus de vingt heures par semaine. Le jeu est divisé en serveurs, et chaque serveur peut accueillir jusqu'à 2 000 joueurs. La pointe d'utilisation se situe à 118 000 joueurs connectés simultanément. Il s'agit d'un jeu de rôle (cf. 7.?.?) qui propose un environnement semi-immersif en 3D. Il définit ses propres espèces qui se subdivisent elles-mêmes en plusieurs races, classes, systèmes économiques, alliances et politiques. Le joueur est invité à se fabriquer un personnage, noble chevalier, elfe voleur, ou encore marchand nain rapace, et à entreprendre sa quête de savoir, d'expérience, d'alliés, dans les villes et villages à travers les cinq continents. Dans des décors remplis de donjons, de tours et de cryptes, les personnages qui peuvent acquérir jusqu'à 40 habiletés sont confrontés à des monstres puissants. Les joueurs peuvent découvrir et accumuler des artefacts pratiques ou magiques. Des douzaines de divinités influencent le déroulement de l'histoire du monde.

Dans les environnements immersifs partagés, le concept de *persistance* est très important: le monde doit continuer à évoluer et à se développer même si personne ne joue. Le temps ne passe pas sans modifier le monde, ce qui incite les joueurs à revenir jour après jour suivre l'action.

Cette technologie donne lieu à des applications inédites, sans personnage ni jeux de rôle, destinées à la communication, à la formation, à la collaboration professionnelle ou tout simplement au divertissement entre amis.



5.4 L'environnement

Tridimensionnel, l'environnement immersif dépasse la surface bidimensionnelle de l'écran. Loin de se limiter à un décor, cet environnement peut contenir des objets, des personnages, dont un *avatar* représentant l'utilisateur si l'immersion est indirecte. En infographie graphique, les environnements simulés sont appelés *scènes*. Ils sont faits de formes géométriques constituées de polygones, assemblées les unes aux autres et recouvertes de textures. Le modèle de l'environnement simulé prend la forme d'un ensemble de descriptions pour chaque objet. Sont ainsi décrits son positionnement (ses coordonnées sur les axes du x pour la largeur, du y pour la hauteur et du z pour la profondeur), ses propriétés et les instructions associées d'assemblage et de mouvement.

Pour que ces environnements existent, il faut rendre ces modèles dans un format de matrices de pixels. C'est à ce moment, entre autres, que les automates simulent les lois de l'optique en calculant la diffraction, la réfraction et la réflexion de la lumière sur les objets du modèle à partir des propriétés de leur texture. Il en va de même pour les personnages, qui sont, en plus, dotés de comportements qui dépassent les autres animations en complexité, surtout si l'on recherche l'effet de réalisme. Ces environnements ne contiennent donc ni menus, ni boutons ; l'utilisateur reçoit les stimulations sensorielles conformes à l'environnement représenté, et doit développer pour ensuite les appliquer des stratégies interactionnelles avec son corps (ou sa représentation) en respectant les contraintes caractérisant l'environnement.

5.4.1 L'ontologie

Le terme « environnement » est synonyme du terme « monde ». Mais le second est plus générique, car il désigne l'unité ordonnée d'une multiplicité d'êtres et d'objets quelconques. La multiplicité seule ne mérite pas le nom de monde, car le multiple n'est pas nécessairement doté d'organisation interne, ne " fait pas système ", selon une formulation plus moderne. La cohérence interne des éléments constituant le monde peut être plus ou moins grande. Par ailleurs, il faut encore distinguer la diversité d'éléments sans unité intrinsèque qui forme un ensemble sans former un



monde, alors que la multiplicité totalement désordonnée équivaut simplement au chaos.

L'environnement est d'abord doté d'une *ontologie*, c'est-à-dire d'un sens et d'une référence, indépendamment de ses parties constitutives et de son fonctionnement. En somme, pour construire une bonne ontologie, il est nécessaire que les réponses aux questions existentielles soient explicitées dans la représentation elle-même : qu'est-ce que cela ? à quoi cela sert ? d'où est-ce que cela vient ? qu'est-ce que cela devient ? Toute entreprise de construction d'environnement immersif devrait commencer par la spécification de son ontologie dans le scénario. Ce document servira par la suite de référence au moment de choisir la plastique de la représentation, les mouvements ou les comportements des objets et des personnages qui constituent l'environnement.

Le *contexte* de l'environnement, laissé au goût et à la fantaisie du concepteur, ne constitue pas seulement un *fond* qui permet de distinguer des objets. Il représente les *conditions* qui rendent la présence des objets de l'environnement possible. La différence entre ces deux conceptions du contexte est immense.

L'ontologie des environnements immersifs de synthèse est particulièrement importante lorsque ceux-ci sont constitués de collages arbitraires d'éléments tirés hors de leur contexte. En effet, la construction originale d'un environnement 3D demande tant de travail et d'efforts que la conception se résume bien souvent à sélectionner et intégrer des objets standardisés, des personnages et des comportements fournis par des marchands de logiciels dans des paysages fractals ou de planchers en damier, etc.

5.4.2 La fermeture

La fermeture, les limitations ou les contraintes de l'environnement immersif jouent pour beaucoup dans la manière dont les utilisateurs apprécient leur expérience. La structure joue un rôle insoupçonné dans le processus créatif. En effet, nous fonctionnons toujours à l'intérieur d'un cadre qu'impose un ensemble de contraintes de tous ordres. Ironiquement, ce cadre est essentiel à l'expression de la fantaisie,



de l'imaginaire. Les utilisateurs des environnements immersifs sont d'abord soumis aux contraintes qu'induisent les limitations du système lui-même : vitesse de traitement, dimensions de l'image, durée du délai de transmission, etc.

Lorsque les utilisateurs sont incités à faire n'importe quoi, tout ce qu'ils désirent, pourvu que ce soit créatif, ils ne sont généralement pas très satisfaits de leur expérience. Le psychologue Rollo May (1975) a étudié relation entre la *créativité* et les *contraintes* dans un texte intitulé *Le courage de créer*. Sa conclusion, c'est que la créativité surgit de la tension entre la spontanéité et les contraintes.

Paradoxalement, les contraintes permettent de canaliser les efforts créatifs en augmentant notre pouvoir imaginaire, parce qu'elles réduisent le nombre de possibilités qui nous sont offertes. Ces contraintes induisent un sentiment de sécurité qui nous permet de laisser aller notre imagination.

La *fermeture* de l'environnement fournit la même sécurité: si les individus respectent les limites en s'empêchant d'introduire de nouvelles potentialités, la compréhension de l'environnement devient possible. Les environnements sont plus satisfaisants lorsque le hasard et le bruit typique des systèmes ouverts ne viennent pas obscurcir les relations causales entre les événements comme dans la vraie vie. Les contraintes peuvent être explicites et implicites. Elles sont *explicites* lorsque, par exemple, les menus instaurent une rupture dans l'immersion par leur présence. Elles sont *implicites* lorsqu'elles sont inférées à partir du comportement de l'environnement; dans ce cas, elles interfèrent moins dans l'immersion.

Les contraintes peuvent aussi être extrinsèques ou intrinsèques, suivant le contexte mimétique, l'objet de la représentation, ou l'ontologie de l'environnement. Les contraintes sont *extrinsèques* lorsqu'elles touchent au contexte de la personne en tant qu'utilisateur du système informatique. L'*échappement* (en cas de " plantage de " l'ordinateur) est un exemple frappant de contrainte extrinsèque. Les contraintes extérieures peuvent sembler intrinsèques lorsqu'elles sont exprimées dans les termes du contexte mimétique. Par exemple, si l'échappement est défini comme un mécanisme d'autodestruction et est instancié dans un objet de l'environnement, alors l'immersion est partiellement préservée.



Les humains ont une grande compétence pour décoder les contraintes à partir des contextes, ainsi les contraintes de l'environnement et des comportements sont reconnues et intégrées au fur et à mesure que l'utilisateur se familiarise à l'environnement.

5.4.3 La modélisation 3D

L'être humain vit dans un monde à trois dimensions. Mais lorsqu'il veut le représenter, il doit le ramener à deux dimensions. Alors, deux solutions s'offrent à lui. Soit il représente uniquement une face plane des objets et des sujets, comme dans la peinture jusqu'à la Renaissance. Soit il tient compte des lois de projection, telle que la perspective. En informatique graphique, les supports matériels (essentiellement des écrans) sont bidimensionnels. Là encore, il faut distinguer deux approches qui débouchent sur deux types de modélisation, de systèmes graphiques et d'applications : les systèmes à deux ou à trois dimensions.

Les *systèmes à deux dimensions (2D)* sont dotés d'une représentation interne, elle-même en deux dimensions, à savoir une matrice de pixels. Chaque pixel se voit donc attribuer des coordonnées, qui permettent de situer sa colonne ainsi que sa rangée dans la matrice. Il est également doté d'une couleur (cf. 1.5.7). Les *systèmes à trois dimensions (3D)*, eux, sont dotés d'une représentation interne tridimensionnelle.

Cela dit, lorsqu'on voit une image générée par un ordinateur, il est impossible de déterminer si cette image a été produite avec un système à deux ou trois dimensions, car le rendu est toujours en 2D. Par exemple, une maison peut être dessinée selon les lois de la perspective et fournie à un système 2D qui se contente de la stocker. Ou encore peut-il s'agir d'une vue en perspective synthétisée par un système tridimensionnel, à partir de données tridimensionnelles.

Le *point* est l'objet graphique 3D le plus simple. Il est caractérisé par ses coordonnées sur les axes du x pour la largeur, du y pour la hauteur et du z pour la profondeur. La *droite*, limitée par deux points, est une figure très courante dans la construction de polygones. Le *polygone* est une figure plane définie par une liste de



points (les sommets), reliés par des segments de droite (les arêtes). Les polygones sont intéressants en 2D, car ils peuvent être facilement remplis de couleurs, auxquelles on peut tout aussi facilement assigner un dégradé. En 3D, les polygones constituent la base du modèle de représentation des objets. Dans ce modèle, tous les objets sont découpés en faces polygonales. Pour les objets qui ne comportent pas d'arêtes comme les sphères, il faut procéder à des approximations.

Les coordonnées des objets sont toujours relatives. Une origine est fixée à un objet, et tous les autres objets graphiques qui constituent un modèle sont référencés par rapport à celui-ci. Pour construire un objet modélisé par polygones, il faut disposer d'une liste de sommets et d'une liste de faces, avec, pour chacune d'entre elles, les références aux sommets de la face dans la liste globale.

La génération de *solides* ou de *volumes* est très importante dans la modélisation 3D, qui repose sur une géométrie des solides. Il faut être capable d'opérer des ouvertures dans leur surface ou de les combiner. Les objets se construisent à partir d'opérations ensemblistes sur des solides : union, intersection et différence.

Des *textures* sont utilisées pour éviter ou réduire l'apparence lisse, régulière et « plastique » des objets. Une texture constitue une microstructure de la surface à représenter. Différentes techniques sont envisageables pour créer la texture. Soit on applique une image provenant d'une photographie ou d'un dessin sur une surface libre. Pour ce faire, le moteur de rendu doit transformer un système de coordonnées, celui de la surface de la texture, en un système de coordonnées tridimensionnelles, celui des formes modélisées. Cette technique ne donne pas de bons résultats lorsque le nombre de points de la surface à afficher est inférieur au nombre d'éléments dans l'image à appliquer. Soit encore on altère la lumière, ou l'on utilise une texture solide indépendante de la forme de l'objet. De même, la *couleur* est indispensable pour rendre les images réalistes, tout comme les *ombrages*, ce qui requiert en tout une très grande quantité de couleurs.

5.4.4 Modélisation de la nature



Montagnes, côtes marines, galaxies, nuages: ces objets naturels semblent dépasser en complexité les lois rigoureuses de la géométrie euclidienne, à cause des formes irrégulières qui les caractérisent. Il s'agit des objets les plus difficiles à modéliser. Les artefacts, au contraire, sont mieux adaptés, tout simplement parce qu'ils ont déjà été construits une première fois. Mathématicien français aux laboratoires d'IBM des États-Unis, Benoît Mandelbrot introduit en 1975 le terme *fractal*, forgé à partir du latin " fractus " (du verbe *frangere*, qui signifie «briser»), pour décrire des objets mathématiques ou physiques et quantifier leur degré d'irrégularité. Il propose une nouvelle géométrie fractale, à partir de laquelle on peut construire des objets qui peuvent représenter aussi bien les formes des montagnes que les réseaux de fluides dans l'organisme.

Un *objet fractal* est un objet à dimension fractionnaire (une fraction étant un intervalle de valeurs entre les entiers). Cet intervalle occupe l'espace entre les lignes (selon une dimension), entre les plans (selon deux dimensions) et les volumes (selon trois dimensions). Une *courbe fractale* est une ligne qui comprend une infinité de cassures, et qui contient une variété infinie de détails à chaque point, tout le long de la courbe. Un *objet fractal* est doté d'autosimilarité: chaque petite portion de l'objet, lorsqu'elle est agrandie, reproduit exactement une plus grande portion du même objet. Les *surfaces fractales* sont générées à partir d'une subdivision récursive, à l'aide de faces triangulaires ou quadrilatères.

Certains objets, comme les nuages, le feu, la fumée ou l'eau, ont une forme irrégulière, mal définie et changeante. Ces objets sont difficiles à modéliser géométriquement: il faut plutôt utiliser des objets flous. Pour modéliser des arbres, on peut utiliser, entre autres, des dérivations fractales.

5.4.5 La lumière

La manière dont les intensités de couleurs se répartissent sur la surface des objets induit une impression de volume. Chaque point peut avoir une intensité différente, en fonction de la lumière présente. Pour synthétiser ou rendre une scène tridimensionnelle avec de la couleur, il faut aussi modéliser la lumière. Il y a



plusieurs sources de lumière différentes : 1) la *lumière ambiante* éclaire la scène uniformément, et se caractérise par son intensité ; 2) les *sources directionnelles* éclairent la scène avec des rayons parallèles à une direction donnée, et se caractérisent par leur intensité et par leur direction; 3) les *sources ponctuelles*, placées en des point précis, rayonnent la lumière radicalement ; elles sont caractérisées par leur intensité et par leur position ; et enfin 4) les *projecteurs (spots)* sont caractérisés par leur position, par leur direction et par un facteur de concentration.

Calculer la répartition de l'illumination sur une surface est très complexe et peut coûter très cher en temps de rendu, à cause des multiples réverbérations. Le réalisme des images dépend pour beaucoup des sources de lumière. À moins de placer une source unique de lumière là où se trouve l'observateur, il est indispensable de tenir compte de son ombre portée, ce qui nécessite un autre calcul onéreux.

La technique du « lancer de rayons » est extrêmement puissante. Basée sur la simulation numérique d'optique géométrique, elle permet de traiter pratiquement tous les aspects nécessaires pour susciter le réalisme. Pour chaque pixel, des rayons sont tirés à partir de l'observateur. Le calcul récursif des intersections de ces rayons avec les objets de la scène permet d'obtenir les données photométriques nécessaires pour colorier le pixel courant. L'intensité des rayons est calculée en se basant sur les propriétés de diffusion par la transparence et de réflexion spéculaire en fonction de la couleur, ainsi que sur l'intensité de la source. À partir de chaque surface touchée par un rayon, de nouveaux rayons peuvent être engendrés. Le processus récursif de la réflexion s'arrête lorsqu'un rayon quitte la scène, heurte une surface qui n'est ni spéculaire, ni transparente. La contribution du rayon devient alors négligeable. L'enjeu de l'optimisation de cette technique consiste à réduire au maximum le nombre d'intersections.

5.4.6 La caméra et le rendu



Pour que les environnements modélisés puissent s'afficher sur des écrans bidimensionnels, il faut que le modèle, décrit et manipulé au départ par des vecteurs de propriétés, soit transformé en matrices de pixels qui conviennent aux écrans. Pour afficher l'environnement sur un écran, il faut projeter les objets de l'espace sur un plan, appelé *plan de vue*, qui sera assimilé à l'image rétinale. Cette projection qui suit les lois de la perspective s'obtient de la façon suivante : à partir d'un seul *point*, appelé « centre de projection », on passe des *droites*, les projecteurs, qui vont rejoindre de tous les objets, et en cherchant l'intersection avec le plan de vue. Ce centre de projection correspond au regard (ou point de vue) de l'utilisateur. Il a un angle de vue de 45 degrés et est représenté par la caméra dans les systèmes graphiques 3D. C'est au moment du rendu que des algorithmes simulant les lois de l'optique calculent la diffraction, la réfraction et la réflexion de la lumière sur les objets du modèle, en fonction des propriétés de leur texture.

5.4.7 L'animation

L'*animation par ordinateur* est une technique qui permet de créer l'illusion du mouvement d'une scène dynamique, à l'instar de ce qu'il se passe au cinéma. La technique consiste à capter une série d'états individuels et donc distincts sur un support au moyen d'un appareillage approprié, en les reproduisant sur un écran à la vitesse qui permet de créer l'illusion de mouvement, c'est-à-dire habituellement 30 images à la seconde.

Dans l'ordinateur, une structure de données appropriée et des algorithmes de modélisation permettent de modéliser un environnement arbitrairement complexe. Celui-ci se compose d'objets fixes (comme les murs) ou d'objets libres ou mobiles (comme les chaises). Seuls les objets libres peuvent être manipulés. Les objets tridimensionnels sont constitués d'agrégats de formes géométriques, elles-mêmes constituées de polygones et recouvertes de textures. Comme on y a déjà fait allusion plus haut, les objets sont définis par les éléments suivants: a) sa position et son orientation, qui sont rendues par des coordonnées sur les axes du x pour la largeur, du y pour la hauteur et du z pour la profondeur de chaque forme géométrique qui entre dans la constitution des objets, b) certaines propriétés,



comme la taille, la forme, la couleur, la transparence et la texture c) les instructions d'assemblage et de mobilité.

L'animation, la manipulation et la transformation des objets sont effectuées à partir de calculs qui modifient les objets en assignant certaines valeurs à telle ou telle des propriétés citées plus haut. Trois propriétés peuvent être modifiées par des transformations ponctuelles, car elles s'appliquent sur tous les points de l'objet, respectivement la position, l'orientation et la taille. La position est modifiée par des translations, une translation se définissant comme un déplacement au cours duquel les différentes positions d'une même droite liée à l'objet restent parallèles.

L'orientation est modifiée par des rotations, c'est-à-dire par un déplacement de tous les points selon des arcs de cercle de même angle au sommet et de même axe. La taille est modifiée par des transformations d'échelles. Quant à la forme, elle est plus complexe à modifier, parce que cela nécessite de modifier structurellement les formes géométriques qui constituent les objets.

Un autre type d'animation est aussi possible lors du rendu, à savoir celui qui correspond au mouvement de la caméra, et donc au mouvement de notre regard dans l'environnement. L'orientation et la position de la caméra peuvent être modifiées: on peut donc l'animer et donner ainsi à l'utilisateur que l'environnement l'entoure et bouge par rapport à lui. Les sources de lumière peuvent aussi être animées (pensons au passage d'une voiture la nuit par exemple).

Il y a deux types d'animation, l'une relevant de l'approche cinématographique (il y a alors animation image par image), l'autre relevant de l'approche cybernétique (l'animation se fait en temps réel).

Dans l'*approche cinématographique*, on installe des éclairages, on décide de l'emplacement de la caméra, de la durée et du montage des séquences. Le rendu s'effectue dans un format vidéo, à raison de trente images *bitmaps* à la seconde. Le résultat peut être diffusé comme tel, ou bien s'intégrer dans une production cinématographique, ou encore dans une machine à contenu.

Dans l'*approche cybernétique*, on rend les scènes en temps réel. Cela n'est possible qu'avec un ordinateur, et, surtout, une carte graphique qui peut calculer les



mouvements et les transformations des objets, puis afficher les résultats de manière suffisamment rapide pour que les objets modélisés aient l'air réalistes. Dans l'approche cybernétique, il faut aussi que la caméra soit dirigée par l'interacteur qui se déplace dans l'environnement. Les éclairages et la position des sources de lumière de l'environnement représenté doivent être cohérents. De plus, la manière dont le système réagit aux actions de l'utilisateur doit respecter les propriétés du sens commun en général et des objets de l'environnement représenté en particulier. Par exemple, lorsqu'un personnage (éventuellement l'interacteur) heurte une surface dure, il absorbera une partie de l'énergie s'il s'agit de la surface d'un mur, mais renversera et brisera un objet mobile s'il s'agit de la surface d'un objet moins résistant, comme une table de coin.

5.4.8 Méthodes d'animation

Il y a plusieurs manières d'approcher l'animation des objets et des personnages (cf. 5.5) dans les environnements. Les méthodes d'animation peuvent être classées en fonction de la nature de l'information requise pour la manipulation (géométrique, physique ou comportementale). Dans les situations concrètes, les approches sont souvent combinées afin de maximiser l'effet de réalisme recherché.

La plus simple d'entre elles est tirée du cinéma d'animation. C'est la plus primitive et la plus coûteuse en travail humain. Tout au long d'une ligne imaginaire qui symbolise la durée de l'animation, on détermine la valeur des différents paramètres des objets impliqués, en termes de coordonnées, d'angles, de vitesses ou d'accélération. Les valeurs peuvent être assignées aux paramètres *directement* à l'écran. Cela ne peut se faire que de manière intuitive, et les positions ne sont donc définies qu'approximativement. Les valeurs peuvent être aussi assignées *rationnellement*, par le biais de zones de dialogues qui permettent d'inscrire directement les données numériques.

Les *images intermédiaires* entre deux images-clés sont calculées par interpolation, c'est-à-dire par approximation récursive de valeurs intermédiaires. L'*interpolation* sera *paramétrique* lorsqu'elle modifie la valeur des paramètres tels que la position,



l'orientation, la taille, la transparence et la couleur. L'interpolation sera *géométrique* lorsque c'est la forme qui est transformée. Cela exige de prétraiter les objets, de manière à égaliser le nombre de facettes et de sommets des deux formes. Ce procédé est appelé *morphing*.

Le second type d'animation est dit procédural ou algorithmique, parce que le mouvement est régulier et décrit par une équation. Ce type d'animation est basée sur les lois de la physique mécanique. Ces lois portent sur l'ensemble des forces en interaction et en opposition dans la mécanique qui régit le mouvement des corps et des objets: lois de la pesanteur, de la gravité, etc. Les prendre en compte ajoute au réalisme. L'algorithme d'animation calcule les trajectoires et les vitesses en résolvant les équations des mouvements à partir des données tirées du contexte.

5.5 Les personnages

Dans la sphère du spectacle, les personnages peuvent être définis *structuralement*, comme un ensemble de traits, de prédispositions et de choix qui, pris ensemble, forment une entité cohérente. Sur un plan *fonctionnel*, les personnages sont les agents des actions qui vont prendre place dans l'environnement.

Les personnages de synthèse sont des humains « virtuels », qui n'ont d'existence que dans les environnements immersifs pour lesquels ils ont été conçus. L'insertion « d'humains virtuels » à l'intérieur des environnements immersifs pose les plus grands défis à l'infographie. On réduit cette complexité en modélisant et animant les humains de manière modulaire et stratifiée en isolant plusieurs couches.

5.5.1 Typologie

Les personnages de synthèse se divisent en plusieurs catégories suivant la manière dont ils sont contrôlés. Tout d'abord, ils peuvent être contrôlés directement et intégralement par l'utilisateur. Mais alors, leurs mouvements ne correspondent pas à ceux de l'utilisateur. Deuxièmement, ils peuvent être guidés d'une part par l'utilisateur qui peut modifier les paramètres des différents moteurs de locomotion, de gestes, d'expression, etc. et d'autre part par l'assemblage d'automatismes, c'est-



à-dire par des séquences d'action préétablies. Troisièmement, les personnages de synthèse peuvent se voir attribuer de l'autonomie par rapport à l'utilisateur. Ils sont alors capables de se mouvoir par eux-mêmes à partir d'états internes construits suivant le but poursuivi et adaptés aux informations que rapportent les différents senseurs : à la façon des utilisateurs humains, ils sont dotés d'un système visuel par lequel ils perçoivent l'environnement et les autres personnages, et d'un système auditif qui leur permet de percevoir et de localiser les sources sonores de l'environnement. Quatrièmement, les personnages de synthèse, autonomes ou guidés par les utilisateurs, jouent un rôle dans un spectacle médiatique. Leur comportement est dicté par leur personnage: dans le cadre restreint des règles du jeu, ils peuvent réagir de manière autonome et communiquer interactivement avec les autres personnages, humains ou de synthèse.

Les personnages de synthèse peuvent communiquer par la voie verbale ou non verbale (par le corps et les expressions faciales). Leurs comportements peuvent se voir dépendre de leur état émotionnel et de leur mémoire, car ces personnages peuvent non seulement exprimer des émotions, mais peuvent aussi apprendre et s'adapter aux situations nouvelles. Il va sans dire que si les deux premières catégories de personnage se retrouvent actuellement dans les jeux par ordinateur actuellement, les deux dernières font encore l'objet de recherches et de développements dans les laboratoires; seules certaines caractéristiques et fonctionnalités ont déjà fait l'objet d'implémentations partielles et ponctuelles.

5.5.2 Modélisation

À un premier niveau de complexité, la *modélisation des personnages* obéit aux mêmes règles que la modélisation des objets. Donc tout ce qui a été expliqué dans la section précédente s'applique à peu près intégralement à la modélisation des personnages. Toutefois, les personnages ne sont pas construits à partir d'un agrégat arbitraire de volumes, unis afin de constituer un objet. Ils sont construits à partir d'un *squelette*. Le squelette est une structure hiérarchique, formée d'un ensemble d'os, c'est-à-dire de segments rigides connectés entre eux, correspondant aux membres et aux articulations. Le nœud pelvien est à la base de la hiérarchie.



Une *articulation* constitue l'intersection de deux segments et peut avoir au plus trois sortes d'angles de position, tous conditionnés par la forme de l'articulation: 1) la *flexion*, qui correspond à une rotation du membre qui met en mouvement tous les segments liés à cette articulation, 2) le *pivot*, qui fait tourner l'axe de flexion autour du membre et 3) la *torsion* du membre.

Un squelette humain compte environ deux cents os, assemblés en figures rigides articulées, reliés par un à trois degrés de liberté, c'est-à-dire de possibilités coordonnées de rotation ou de translation. Il contient 49 articulations avec 88 degrés de liberté. Le volume de l'humain peut être approché à partir de 15 primitives solides : deux cylindres pour le torse et le cou, un ellipsoïde pour la tête et le cou, et quatre cônes tronqués pour les bras et les jambes, deux sphères pour les mains, et, pour finir, deux cônes tronqués pour les pieds.

Les muscles constituent un système dynamique, composé d'éléments dotés d'une certaine élasticité, qui soumettent les os à des contraintes et à des forces positionnelles, par leur forme et par leurs points d'attache. Les muscles sont en fait assimilables à des ressorts dont l'élasticité et la rigidité varient, et soumis à des forces de compression et d'élongation.

La modélisation du squelette est constituée d'un ensemble de volumes indéformables articulés. On introduit des effets de cassure au niveau des articulations, effets atténués par l'ajout successif, le cas échéant, d'une peau qui entoure le squelette et d'un vêtement qui recouvre la peau. Le module d'habillage se superpose à la modélisation du squelette. Il calcule la déformation de la peau et des vêtements, en fonction des caractéristiques de la peau et des textiles, et en fonction de la configuration du squelette lors des mouvements.

Les cheveux posent un des problèmes les plus complexes à résoudre si l'on veut les représenter de manière réaliste. Une chevelure est composée d'en moyenne cent à cent cinquante milles cheveux. Au plan de la modélisation, le problème est de spécifier et par la suite d'animer un tel nombre d'éléments individuels, et de paramétrer leur géométrie, leur distribution, leur forme, leur direction. Cela est



d'autant plus compliqué que, au niveau du rendu, le diamètre d'un cheveu est souvent inférieur à la largeur d'un pixel, unité de base de la représentation.

5.5.3 Animation

Les mouvements des personnages synthétiques sont contrôlés localement de la même manière que les déplacements d'objets: on produit les données géométriques appropriées selon le déroulement de la ligne du temps. Quelles que soient les procédures d'animation utilisées, les mouvements sont dans tous les cas produits par des translations, des mouvements linéaires, des rotations du nœud pelvien (et donc du personnage en entier), et des rotations de chaque os du squelette, en fonction des contraintes qui leur sont propres.

Cela dit, il y a une différence de richesse entre les deux: les personnages de synthèses se comportent de manière plus complexe que les objets ne se meuvent, surtout si un effet de réalisme est recherché.

Les méthodes d'animation des personnages recourent en partie celles qui sont appliquées aux objets (l'animation par images-clés et l'animation algorithmique). Dans l'approche par images-clés, qui repose avant tout sur le savoir-faire d'un professionnel de l'animation, les rotations et les flexions des articulations du squelette s'ajoutent aux transformations et aux déplacements de formes. Il est important d'utiliser des points fixes qui ne sont pas animés pour éviter les sautilllements ou d'autres erreurs de registre.

L'approche algorithmique tire parti de la puissance de calcul de l'ordinateur. L'animation des objets est sous-tendue par les lois de la dynamique. L'animation des personnages, elle, met également en jeu les lois de la *cinématique*, c'est-à-dire de l'étude des mouvements humains, ainsi que celles de la *physique mécanique* (cf. 5.4.8). Pour animer les personnages, on utilise les données correspondant à un mouvement donné, défini en termes de forces, de torsions et de contraintes. Les *forces* provoquent des accélérations linéaires et angulaires des membres reliés par des articulations. L'algorithme d'animation calcule les trajectoires et les vitesses en



résolvant les équations des mouvements à partir des données tirées du contexte : les forces mises en œuvre, les contraintes et les propriétés de masse des objets.

Les contraintes cinématiques, énergétiques et dynamiques du personnage sont entièrement modélisées, tout comme les contraintes de l'espace-temps de l'environnement. De ce fait, on peut automatiser les trajectoires spécifiées et le contrôle des mouvements, aussi bien pour la préhension que pour la locomotion. Dans l'animation, les muscles jouent un rôle important, notamment dans le rendu/ la représentation du mécanisme réflexe de compensation, qui permet au corps de maintenir sa position (par exemple la position debout), malgré les forces de gravité qui s'exercent contre elle.

Les *algorithmes basés sur la biomécanique* donnent généralement des résultats réalistes au plan de la physique, mais trop réguliers pour pouvoir refléter la « personnalité » d'un personnage. Pour plus de réalisme, les coordonnées résultant des algorithmes biomécaniques sont altérées en temps réel en fonction de la morphologie du personnage et des vêtements qu'il porte. Sont même prises en compte les déformations de la peau que provoquent la flexion d'un membre ou le contact avec un objet. La personnalité des personnages habituellement spécifiée dans le scénario (cf. 7.4.2), se traduit en quelques paramètres uniques, qualifiés de « génétiques », qui déterminent la morphologie du personnage, le style des mouvements produits par algorithme.

5.5.4 Capture de mouvement

La *capture de mouvement* peut remplacer l'animation manuelle que l'on vient de décrire. Cette technique consiste à mesurer et à enregistrer les actions d'un acteur au moyen d'un dispositif basé sur minimum treize capteurs optiques ou magnétiques. Ces capteurs sont placés à des endroits stratégiques : un sur la tête, un sur chaque bras, un sur chaque main, un au centre de la poitrine, un au bas du dos, un sur chaque genou, un sur chaque cheville et un sur chaque pied. Ces capteurs sont capables de mesurer leur position et leur orientation par rapport à un transmetteur centralisé qui sert de point de référence. Pour utiliser les mesures ainsi



enregistrées, il faut les mettre en correspondance avec le mouvement du personnage de synthèse.

La capture de mouvement permet de représenter des mouvements très réalistes. Mais ceux-ci ne peuvent qu'être reproduits: les adapter à n'importe quelle variation de la situation en jeu est très laborieux, voire impossible, à cause des quantités de situations différentes et des types de « colle infographique » requises pour ce faire. Par colle infographique, on entend le mécanisme qui assemble les différentes données du mouvement captées aux articulations des différents personnages de synthèse. Pour conserver la vraisemblance, les mouvements et les gestes des personnages doivent pouvoir être facilement adaptés à toutes sortes de situations: les terrains accidentés, les objets avec lesquels le personnage peut entrer en collision, etc.

5.5.5 Comportements

L'environnement influence les mouvements du corps des personnages. Les personnages interagissent volontairement avec l'environnement. Il existe une troisième manière d'approcher l'animation des personnages, appelée *animation comportementale*, précisément parce qu'elle prend en compte les relations que chaque élément entretient avec les autres objets. Les comportements sont traduits en indications procédurales de haut niveau et les personnages (comme tout autre élément animé) sont considérés comme des entités autonomes. L'animation comportementale distingue deux niveaux : la *réalisation* des actions dans l'environnement d'une part, et la planification, la gestion de la séquence d'actions nécessaires pour atteindre un but d'autre part.

Le personnage se déplace dans l'environnement ; on appelle *espace de travail* l'espace que le personnage peut immédiatement atteindre au moment d'agir, dans n'importe quelle direction que ce soit. Les actions sont déclenchées après avoir été planifiées, et leur réalisation dans l'espace de travail est soumise à l'ensemble de contraintes qui régissent les articulations du personnage, les objets fixes et mobiles qui constituent l'environnement, et la personnalité du personnage. Comme les



actions sont effectuées dans l'environnement, il est essentiel de planifier les trajectoires et d'anticiper les collisions potentielles, pour ainsi éviter les obstacles.

Au niveau de base, l'animation se ramène toujours, en somme, à faire varier certains angles du modèle au cours du temps. À un niveau plus élevé, elle revient à planifier et à gérer des séquences d'actions en vue d'atteindre un but. Les algorithmes de planification développés par les chercheurs en intelligence artificielle (cf.2.5.3) sont mis à contribution aussi bien dans l'animation de personnages de synthèse que dans celles des robots. Le gros problème est de savoir comment décomposer une tâche donnée en une séquence de mouvements élémentaires.

Aucune des approches vues ne peut par elle-même assurer l'ensemble de l'animation des personnages de synthèse. Pour leur faire accomplir certaines tâches, il est souvent nécessaire de combiner les approches. Un exemple fera apercevoir la complexité de la réalisation d'une animation comportementale apparemment simple: un personnage se lève, marche jusqu'à la porte et l'ouvre. Pour marcher alors qu'on est assis sur une chaise, il faut se lever, pour ouvrir une porte, il faut saisir la poignée et la tourner, etc.

Tous les comportements doivent être soigneusement scénarisés et programmés sous forme de règles. Les conditions de ces règles sont basées sur les valeurs des propriétés d'objets: leur nature, leur forme, leur distance, leur texture, etc. Ces règles déclenchent différentes actions en fonction du contexte ; par exemple, face à une porte, la main se tend, saisit la poignée et la tourne. Les langages de description de comportements sont difficiles à utiliser; ils comportent des commandes scripto-iconiques, qui se présentent sous forme de boîtes reliées à des flèches, et organisées en fonction de la topologie de l'environnement simulé. Ces langages comportent aussi des commandes textuelles organisées en fonctions ou en objets.

Pour faire marcher un personnage, on définit une suite de mouvements, calculés à partir d'une modélisation de la marche en termes de changements de valeur des articulations du corps. Cette suite de mouvements sera adaptée en temps réel, non seulement, sur le plan physique, aux propriétés du sol, mais aussi, sur le plan



cinématique, à la personnalité de l'utilisateur ou aux traits du personnage de synthèse. L'orientation des mouvements sera guidée interactivement par l'utilisateur — le personnage constitue dans ce cas un avatar —, ou bien générée automatiquement par un animateur, à partir de trajectoires de référence, ou encore obtenue par acquisition de mouvements, à l'aide d'un exosquelette.

Il existe deux techniques pour implanter les comportements des personnages dans l'environnement: l'automate à états finis et le réseau neuronal. La première technique convient bien pour les comportements déterminés une fois pour toute lors de la conception. Elle repose sur deux éléments : 1) une base de connaissance, qui comporte des séries de règles permettant d'agir dans des situations particulières (monter, pivoter, courir, etc.), et 2) un automate à états finis, dont la fonction est de filtrer et de déclencher des règles de la base de connaissances en fonction du contexte. Les règles sont composées de deux parties. La prémisse spécifie le contexte sous une forme conditionnelle, et la conclusion détermine les actions élémentaires à effectuer.

Cela dit, les comportements les plus évolués ne sont pas hérités lors de la conception, mais sont acquis lors d'un apprentissage, supervisé ou non, le plus souvent par un réseau de neurones (cf. 5.2.6). Il s'agit là de la deuxième technique d'implantation des comportements.

Dans l'animation de personnages, la « mimesis », le réalisme ou la similitude avec le réel ne touchent plus vraiment à l'apparence de l'entité, mais à son fonctionnement. Pour que les comportements soient individualisés, il faut introduire des différences au niveau de la personnalité, des interactions émotionnelles, etc., en bref, au niveau du scénario.

5.5.6 La parole

Jusqu'à maintenant, on a mis l'accent sur les aspects de l'animation des personnages qui touchent aux lois de la cinématique et aux comportements des personnages. On peut se limiter à ces deux facteurs dans les environnements où l'action prime sur la communication. Mais il n'en va pas de même pour les



environnements où l'accent est mis sur une quête de sens. *L'expression des personnages* constitue un troisième aspect qui contribue tout autant au réalisme que les deux premiers. L'expression est tantôt verbale, tantôt non verbale. Dans un univers immersif, le recours à l'écriture par le clavier et l'affichage de texte est proscrit, vu l'effet de présence recherché. La communication verbale implique donc de recourir à la parole. Celle-ci se définit comme un ensemble de gestes articulatoires rendus audibles. Le chapitre suivant aborde en détails le traitement des langues naturelles, dont le traitement de la parole.

Un système générique d'interaction vocale se compose de plusieurs éléments: 1) un *module de reconnaissance vocale*, qui transforme la parole en une suite de mots ou de graphes ; 2) un *analyseur sémantique*, qui tente de déterminer la signification de l'énoncé en établissant une représentation sémantique appropriée, à l'aide de graphes contenus dans une base de données; ces représentations peuvent être multiples, dépendent largement du contexte et, bien entendu, de la performance du système ; 3) un module de *gestion du dialogue*, qui cherche à compléter les énoncés, voire à en réduire autant que possible les incertitudes, en s'appuyant sur l'historique du dialogue ou encore en interagissant directement avec l'utilisateur ; et 4) un *générateur de réponse*, qui traduit le résultat du traitement de la représentation sémantique construite en action ou en parole. Pour la reconnaissance de la parole, les réseaux de neurones est l'approche la plus couramment utilisée. Les corpus d'apprentissage et d'entraînement sont de très grande taille.

Plusieurs critères permettent de caractériser les systèmes de reconnaissance de la parole. Premièrement, la *mode d'élocution* est *continu* lorsque l'utilisateur prononce d'une traite les mots de l'énoncé, ou *isolé* quand il doit produire une pause entre chaque mot. Deuxièmement, la dépendance du système au locuteur varie. Les systèmes *monolocuteurs* ne peuvent reconnaître que la voix des personnes qu'ils ont préalablement appris à reconnaître, alors que les systèmes *multilocuteurs* sont indépendants des particularités de prononciation propre à certains locuteurs en particulier. Les premiers systèmes étaient monolocuteurs, parce qu'ils se basaient sur la reconnaissance globale des mots. Il fallait une phase d'entraînement plus ou



moins longue préalablement à l'utilisation, et les locuteurs devaient prononcer tous les mots en entier. Les systèmes plus récents sont fondés sur la reconnaissance d'unités plus petites que le mot, appelées « triphone ». Ces unités représentent la forme acoustique d'un phonème, à savoir un son élémentaire d'une langue, dans le contexte des phonèmes immédiatement voisins. Troisièmement, les systèmes se différencient suivant le temps d'adaptation du système à un nouveau locuteur, et suivant la nature de ce qui doit faire un nouvel utilisateur. Souvent, il doit prononcer un ensemble de phrases comportant l'ensemble des phonèmes de la langue, afin que le système puisse s'adapter à la voix. La plupart du temps, les systèmes demeurent *adaptatifs*, c'est-à-dire qu'aussi bien les modèles acoustiques que linguistiques sont automatiquement modifiés pendant l'utilisation réelle.

Quatrièmement, le système se distingue par la taille du vocabulaire qu'il est capable de reconnaître. Un petit vocabulaire (quelques dizaines de mots) suffisent à contrôler un processus simple. Un vocabulaire peut être moyen et comporter quelques centaines de mots, comme les systèmes d'information pour les systèmes aériens ou ferroviaires. Enfin, il peut être grand et dépasser le millier de mots, ce qui est notamment le cas pour des applications de dictée. Cinquièmement, les systèmes se distinguent par le lieu où la reconnaissance prend place: elle peut se faire en mode local, directement devant un ordinateur, ou encore au téléphone. Cela peut soulever des problèmes de bande passante, et nécessite un prétraitement, qui consiste à filtrer les interférences de manière à faciliter l'extraction de la parole.

La reconnaissance de la parole se distingue de sa *synthèse*. Il y a deux grandes façons d'approcher la synthèse de la parole: l'assemblage d'enregistrements et la synthèse directe. L'assemblage d'enregistrements en contexte de parole humaine consiste à découper et à traiter préalablement la parole de façon à ce que chaque enregistrement puisse être assemblé à n'importe quel autre, avec un minimum de coupure. L'unité de traitement est le phonème (la langue française en compte 40).

On retrouve ce type de système dans les services qu'offrent les compagnies téléphoniques (assistance annuaire, boîtes vocales, etc.). Cette approche de la synthèse de la parole est la plus simple, mais ne permet pas de rendre la prosodie



de manière réaliste, c'est-à-dire de l'adapter aux énoncés et au contexte, ni de rendre les intensités particulières susceptibles de véhiculer tel ou tel état d'esprit du locuteur, etc., bref, toute chose qui caractérise habituellement les échanges verbaux.

L'approche par la synthèse est plus laborieuse, mais beaucoup plus prometteuse, surtout dans le contexte des environnements immersifs. Suivant cette approche, les réponses orales ne sont pas pré-stockées, mais calculées en temps réel, générées automatiquement. Pour ce faire, deux modules sont requis : un *module de calcul/génération* de l'énoncé et un *organe de « vocalisation »* des énoncés calculés, qui utilise les informations sémantiques et contextuelles de l'énoncé à produire, pour le « vocaliser » avec une prosodie, un ton, une intonation, un accent et une durée adaptés au contexte et à la personnalité du personnage. La vocalisation est produite par un générateur, qui effectue une conversion du numérique vers l'analogique d'un énoncé. La parole calculée, résultat de règles de génération sensibles au contexte, prend la forme d'un vecteur de code binaire ; ces vecteurs seront transformés en une onde sonore continue, par la concaténation de plusieurs centaines de coefficients spectraux ou de paramètres de fréquences par seconde. Les sons de base sont tirés d'un dictionnaire, mais avant la génération du son, leur sont appliquées des transformations prosodiques fonction du contexte.

Pour générer la parole des avatars et des personnages de synthèse, il faut ajouter la *synthèse visuelle de la parole*, qui doit être synchronisée à la synthèse sonore. Cette synthèse visuelle s'opère en pilotant huit paramètres, dont cinq concernent les lèvres : écartement horizontal interne, séparation verticale interne, protrusion du point de contact des lèvres, protrusion des lèvres inférieure et supérieure. Un autre paramètre encore définit la position de la mâchoire, et deux définissent la position de la langue, plus précisément l'angle qu'elle forme avec la mâchoire et son avancement dans la cavité buccale.

Dans le monde des environnements textuels distribués sur le Web, on retrouve déjà des agents conversationnels appelés *chatterbots*. Ces agents logiciels ont pour objectif de simuler une conversation avec les utilisateurs, de donner l'illusion que la



conversation prend place entre deux personnes. Dans la plupart des cas, le traitement linguistique est minimal, et même rudimentaire, car beaucoup de recherches et de développement en ingénierie linguistique sont encore nécessaires pour arriver à un résultat satisfaisant. Pour l'instant, ces agents se contentent d'identifier sommairement le propos de l'utilisateur, puis de produire une réponse adéquate, en identifiant des mots-clés à l'aide de réseaux neuronaux ou en recourant à la logique floue. On recourt à des règles dont le but est de faire penser que la capacité humaine à converser est bien reproduite, mais il s'agit plutôt d'astuces visant à camoufler la faiblesse du traitement linguistique. Par exemple, l'agent logiciel tient des propos ambigus, reprend une partie de la question dans la réponse pour simuler un suivi dans la conversation, change de sujet lorsqu'il y a incompréhension, réutilise des sujets précédemment abordés, etc. Il y a même un concours — le *Chatterbox challenge* —, où un jury d'expert décide de la meilleure performance de *chatterbots* parmi un ensemble de concurrents.

5.5.7 L'expression non verbale

En psychologie, l'*expression non verbale* désigne l'utilisation du corps dans la communication interpersonnelle. L'expression non verbale peut véhiculer de l'information à propos de l'état affectif de l'individu, à propos de son identité, et à propos du monde extérieur. La communication non verbale passe par plusieurs canaux: les positions et les gestes du corps, mais surtout l'expression du visage.

La *proxémique* est la partie de la sémiotique qui étudie la manière dont les êtres animés, et les humains en particulier, utilisent l'espace environnant leur corps, et se situent par rapport aux autres entités qui occupent cet espace. La proxémique fournit des informations précieuses, car elle analyse les distances et les angles que choisissent et utilisent les individus avant et pendant leurs interactions, ainsi que le niveau d'intimité relationnelle associé à chacune des distances : proche, moyenne, éloignée, derrière un bureau, sur une tribune, etc. Elle étudie aussi les circonstances qui autorisent le toucher, etc.



Étudier le langage corporel revient à analyser les gestes, les changements de postures, les mouvements de la main, de la tête, du tronc, etc. Les auteurs s'entendent pour distinguer trois types de mouvements du corps. Premièrement, les *emblèmes* sont des gestes qui ont une signification précise, et qui peuvent généralement se traduire en un mot. Le hochement de tête pour le oui, le pouce vers le haut pour l'approbation sont des gestes emblématiques typiques. Ces gestes se comprennent à l'intérieur d'une culture, nationale ou continentale, ou d'une sous-culture donnée, par exemple le monde des adolescents ou des prisonniers. Les *illustrations*, qui constituent un deuxième type de mouvements du corps, sont directement reliées à ce qui est exprimé verbalement. Troisièmement, il y a les *postures* qui, avec le visage, constituent le principal canal d'expression non verbale des émotions et des états mentaux. Une posture se définit comme une position précise du corps ou de ses parties, associée à un système de référence. Par exemple, les épaules voûtées expriment la soumission ou l'ennui. On peut repartir en outre les gestes en deux classes. Les *gestes de régulation* gèrent la conversation, et les *gestes d'adaptation* sont dictés par des besoins individuels ou par certains états émotifs. Par exemple, celui qui est intimidé se gratte le menton.

Le visage est le lieu par excellence de l'identité. C'est par le visage que les humains se reconnaissent. C'est aussi le siège de l'expression: c'est le vecteur dynamique de signes expressifs animés, les expressions, qui transmettent un contenu émotif. Les mains sont aussi très expressives, et de ce fait, les modéliser est également très important sur le plan de l'expressivité. Plusieurs études ont démontré l'importance des émotions dans l'établissement des interactions sociales au sein des environnements immersifs.

L'animation du visage ne relève pas de la géométrie du squelette. L'expression du visage des personnages s'obtient en modélisant l'action des muscles faciaux dans la synthèse des expressions. On a constitué plusieurs fois un vocabulaire partiel d'expressions faciales. Celles-ci peuvent s'ordonner en fonction de leur impact émotionnel sur une échelle qui va d'une borne positive à une borne négative, en passant par l'indifférence absolue. Les personnages doivent aussi être se voir



attribuer des humeurs, indices bipolarisés de stress, de plaisir, etc. Ces échelles varient au fil des émotions que provoquent les interactions, et influencent les gestes qui seront réalisés.

5.5.8 L'avatar

L'avatar est essentiellement un personnage de synthèse, mais se distingue des autres personnages du même environnement en ce qu'il est le substitut, la représentation graphique de l'utilisateur, à l'intérieur des environnements immersifs. L'avatar constitue en quelque sorte le trait d'union entre l'utilisateur avec l'environnement ; c'est par lui que l'utilisateur pourra « ressentir » les différents attributs de l'environnement. C'est l'utilisateur qui contrôle les déplacements et les gestes de l'avatar à l'intérieur de l'environnement. L'avatar est ainsi influencé par les gestes qu'accomplit l'utilisateur.

L'avatar constitue une présence physique dans les mondes virtuels, et renforce l'effet de présence tant recherché. Il était déjà présent dans les tout premiers jeux sur ordinateur, comme Pac Man. Cet avatar était doté d'une représentation iconique 2D et était capable de déplacements orthogonaux.

Les *Palaces* constituent un exemple parmi d'autres d'environnements partagés ou communautaires. Il s'agit de représentations en 2D de pièces en perspective, dans lesquels des utilisateurs branchés sur Internet se rencontrent et interagissent, par le biais d'un avatar. De style caricatural ou schématique, cet avatar tient souvent de la bande dessinée. Il peut être construit à partir d'un catalogue de parties du corps et d'expression, ou à partir d'un répertoire de personnages comme Marilyn Monroe, l'haltérophile, Tarzan, Pamela Anderson, Cindy Crawford, etc. Sous la poussée de l'industrie des jeux, les avatars se sont vus petit à petit dotés d'une identité physique d'un niveau de plus en plus élevé. Ils ont été incorporés dans les environnements 3D, soumis aux lois de la mécanique newtonienne: leur corps est devenu réaliste et articulé, ce qui permet de rendre les interactions avec l'environnement et les objets qui le composent de plus en plus réalistes. L'avatar est un personnage synthétique qui évolue hors de la réalité, dans des mondes virtuels.



Cela dit, il réfère toujours à la réalité, mais de façon stylisée : certains éléments sont laissés de côté, et d'autres, au contraire, sont exagérés, ce qui leur confère un aspect caricatural.

L'avatar joue un rôle physique dans l'environnement, et social dans l'interaction. À ce titre, l'avatar le plus abouti se voit associer un ensemble de fonctions qui lui permettent d'adapter au mieux son comportement aux autres personnages, à savoir : 1) la perception (pour voir s'il y a quelqu'un dans les environs) 2) la localisation (pour distinguer l'endroit où se trouvent les autres personnages) 3) l'identification (pour reconnaître ces personnages, c'est-à-dire se remémorer les traits et les comportements) et 4) la reconnaissance et l'interprétation des gestes et des actions des autres personnages.

En tant que représentation d'un soi, l'avatar est un instrument d'expression et de communication dans les mondes virtuels. À ce titre, leur individuation est capitale. Ils servent de médiateurs dans la communication entre les individus et l'ordinateur. Les avatars permettent aussi à l'utilisateur d'altérer librement sa représentation dans un environnement immersif. Parmi les pratiques transgressives les plus fréquentes, on retrouve le changement de sexe, d'âge, de taille et d'apparence physique.

Ces avatars sont des représentations schématiques qui mettent l'accent sur des caractéristiques particulièrement saillantes des individus, comme les icônes pour les actions et les contenus. Les avatars jouent aussi le rôle de masque, en laissant à l'utilisateur la liberté de faire l'expérience d'autres identités dans les rapports avec autrui. C'est parfois en partie grâce à l'avatar qu'une conversation est entamée entre deux participants.

5.6 Effet de présence

5.6.1 Définition

Dans un premier temps, la présence virtuelle ou l'effet de présence peut se définir comme l'état de conscience et l'implication d'un individu dans un environnement qui



n'est pas directement présent à lui. Lorsque l'expérience d'une présence a lieu dans un environnement capté et retransmis, on l'appelle « téléprésence ». Cet effet n'est pas propre aux médias interactifs: il suffit de rappeler la réaction des spectateurs lorsque fut projeté le premier film des frères Lumières, « l'Entrée d'un train en gare de Laciostat »; affolés, ils se sont précipités au fond de la salle.

En fait, répondre à des stimulations audiovisuelles ou réelles provoquera les mêmes effets physiologiques sur le pouls, la pression sanguine, la tension musculaire, le diamètre de la pupille, la résistance de la peau. Des réponses corticales peuvent provoquer aussi d'autres stimulations pouvant éventuellement être mesurées par encéphalogramme, et celles-ci, à leur tour, provoqueront une réaction musculaire, une sécrétion d'hormones comme l'adrénaline ou l'endorphine. Les stimulations audiovisuelles génèrent un *feedback proprioceptif*. La proprioception désigne la sensibilité liée au sens de la position relative des segments du corps et à l'appréciation de leurs déplacements, à la régulation du tonus musculaire, à la statique et à l'équilibration.

L'illusion de se mouvoir peut provoquer ce qu'on appelle le " mal des simulateurs ", qui se caractérise par une désorientation, une démarche instable parfois accompagnée de nausée. Enfin, l'immersion provoque aussi bien l'excitation qu'elle permet la relaxation. Elle peut aussi se révéler efficace pour persuader autrui ou encore diminuer certaines phobies.

L'effet de présence a lieu ou n'a pas lieu: il ne s'agit pas d'un phénomène qui se produit à un certain degré, comme la peur ou la tristesse. De plus, suivant que cet effet prend place ponctuellement, de temps en temps ou fréquemment, l'impression sera plus ou moins intense. L'impression d'effet de présence dépend du nombre d'illusions ponctuelles. Comme le sentiment de présence est fondamental dans notre vie de tous les jours, il est difficile de le définir.

5.6.2 Suspension de l'incrédulité

À la fois physiologique et psychologique, cet effet peut se définir négativement comme l'illusion de l'absence de médiation, généralement provoquée chez le



spectateur lors d'une expérience médiatisée. L'expérience médiatisée paraît au contraire réelle, immédiate et directe, et l'utilisateur répond comme s'il n'y avait pas de média. N'importe quel système d'immersion n'induit pas nécessairement un effet de présence chez tout un chacun.

Cette illusion de non-médiation s'explique par un blocage temporaire de la fonction de « monitoring » de la réalité, à la suspension de l'incrédulité, volontaire ou non. *The willing suspension of disbelief* (Hand 1994): l'expression est empruntée au critique et poète Samuel Taylor Coleridge (1772-1834), qui inaugure le romantisme en Angleterre. À travers cette expression, il veut décrire l'état d'esprit à atteindre pour jouir pleinement de la représentation d'une action. Cet état d'esprit est essentiel pour expérimenter d'autres stimulations émotionnelles au moment de visionner l'action des personnages qui prennent part aux événements. Par exemple, lorsque l'héroïne est en danger, nous ressentons la peur pour et avec elle. Mais pas tout à fait la même peur: la peur que nous ressentons est différente, parce que la menace n'est pas réelle pour nous. Faire comme si l'action était vraie induit l'émotion de la peur, mais sans la douleur et les dommages qui y sont associés. Bien au contraire, combien frissonnent de plaisir à l'idée d'avoir peur au cinéma ou en lisant un roman. Il en va de même pour les histoires d'amour, qui nous permettent de vivre ce que notre vie quotidienne ne nous procure pas toujours.

La réussite de la substitution de la réalité par des représentations dépend de la capacité à subsumer notre connaissance qu'il s'agit de représentations. Les émotions ressenties à cause du statut spécial conféré à l'action par la représentation sont différentes, en ce qu'elles procurent beaucoup plus de plaisir que lorsque ces actions arrivent dans la vraie vie.

La même fonction est altérée lorsque prennent place les perceptions sans stimuli que sont les hallucinations. Loin de constituer un dysfonctionnement perceptuel ou psychologique, cette illusion, au vu des gratifications qu'elle fournit, semble plutôt faire partie des besoins fondamentaux de l'individu, compte tenu de l'importance que prennent les médias dans nos vies. Cela dit, la folie guette ceux qui ne font plus la distinction entre la fiction et la réalité de leur propre vie. Pensons par exemple à



ceux qui émettent des signaux vers le ciel en attendant que les extraterrestres viennent leur rendre visite.

La question du rôle de la volonté dans l'illusion est complexe. En fait, pour qu'elle fonctionne, il faut que l'utilisateur accepte tacitement les conditions médiatiques de représentation. Mais il faut en même temps qu'il y adhère en dehors de la sphère de la raison, au niveau épidermique des sensations. Spectacle, roman, promenade dans une représentation dynamique tridimensionnelle d'un bâtiment, etc. induisent le même phénomène à des degrés divers.

5.6.3 Facteurs

Plusieurs facteurs jouent dans l'effet ou le sentiment de présence :

1) La richesse, la résolution et la qualité des stimulations sensorielles. Le visuel a plus de présence que l'audio, qui a lui-même plus de présence que l'écrit. Sur le plan de l'image, la convergence, la précision, la brillance, le contraste et la justesse de la résolution des couleurs concourent tous à l'effet de présence. De plus, le photoréalisme est préférable au dessin, les images en mouvement préférables aux images fixes, la couleur au noir et blanc. Plus la surface d'affichage est grande, plus la présence est grande. Des changements rapides du point de vue suggèrent un déplacement dans l'environnement.

Le son est primordial dans la génération de l'effet de présence. Les espaces sont dotés de caractéristiques qui modifient l'intensité des sons et la réverbération. Notre expérience de l'audition nous permet de situer une source sonore dans l'espace à partir des modifications que l'on y perçoit. De même, un changement continu de l'intensité d'une source sonore nous indique que cette source se dirige vers nous ou s'éloigne de nous. La perception du son est différente de celle des images, parce que la tridimensionnalité est détectée plus rapidement. De plus, l'audition aide à diriger le regard, ce qui diminue le temps de réaction à un stimulus visuel. L'effet maximal est atteint lorsque le son et l'image sont synchronisés, comme on déjà pu le constater au cinéma. Le son couplé à une image pas forcément réaliste joue un rôle déterminant dans le réalisme de la représentation; il suffit de penser au cinéma



fantastique ou de science-fiction. La présence est plus grande dans un système qui compte plusieurs haut-parleurs au volume élevé.

Si les utilisateurs sont mis en présence d'images et de sons de qualité très pauvre, ou encore de fantômes et d'autres bruits ou interférences, leurs réactions dépendront de leur imagination visuelle.

2) La personnalité de l'utilisateur joue beaucoup dans l'effet de présence. Tout dépend de son éveil, de sa vivacité dans la perception et dans le traitement de stimuli, de son style cognitif, visuel ou auditif, de son système représentationnel préféré — concret figuratif ou abstrait stylisé — de son degré de recherche de sensation, de son besoin de combler la solitude, de son humeur, ou encore de son stress pendant et avant l'utilisation. Outre l'âge et le sexe qui semblent des critères déterminants, l'expérience préalable du média joue aussi un rôle important: l'absence de familiarité semble contrecarrer l'effet de présence.

3) La transparence du dispositif. Un dispositif visible, trop lourd, encombrant par ses dimensions, le lien avec l'ordinateur rappellent à l'utilisateur qu'il manipule un dispositif, et donc qu'il a affaire à une représentation. La transparence du dispositif s'obtient en grande partie par l'habitude d'utilisation.

4) La cohérence, la convergence et la fermeture de l'environnement. Celui-ci est cohérent lorsque les différentes stimulations produites dans le temps évoquent tous la même chose, et que les stimulations sensorielles simultanées convergent. Toute inconsistance dans la représentation la rend artificielle. La fermeture de l'environnement réfère au niveau de saturation des indices et des références du monde représenté. Leur convergence fournit l'unicité nécessaire à l'élaboration d'une carte mentale par l'utilisateur.

5) La relation entre les actions du participant et leurs effets dans l'environnement. D'une part, ce lien implique un contrôle total de l'expérience médiatisée par l'utilisateur: il doit pouvoir regarder de tous les côtés, parcourir de grandes distances dans chaque direction, se déplacer à n'importe quel rythme, aller dans n'importe quelle séquence. D'autre part, ce lien implique que l'utilisateur soit



comme prolongé dans la représentation, afin de pouvoir ramasser, palper et bouger les différents objets de l'environnement.

6) L'interaction entre les participants et l'environnement. L'exploration d'environnements à plusieurs utilisateurs renforce la présence, qui devient sociale. Transformé en une entité sociale, le média est perçu comme chaud, sensuel, personnel ou intime. Par ailleurs, même lorsque la présence d'autres personnes est uniquement stimulée par le biais de personnages virtuels, l'effet de présence sociale peut être induit par des interpellations, des adresses directes et par le regard, ou encore par un simulacre de conversation.

Il y a aussi les phénomènes de personnification où des représentations très schématiques induisent un effet de présence sociale par leur comportement. C'est le cas, par exemple, des *Tamagotchis* (mot qui signifie " petit oiseau " en Japonais). Les Tamagotchis furent lancés au Japon le 23 novembre 1996 par Bandai, la firme à laquelle on doit aussi les Power Rangers. Aux États-Unis, ils ont commencé leur épopée le 1^{er} mai 1997. Leur histoire est typiquement japonaise. Ces créatures extra-terrestres de la planète Tamagotchi se sont écrasées sur la terre. C'est un professeur et son assistant Mikachu qui les a retrouvés. Le professeur leur a construit un écrin de protection en forme d'œuf, afin qu'ils puissent survivre sur terre. Puis Mikachu en a peint quelques-uns, et les a apportés à l'école, ce qui a démarré la folie Tamagotchi. Ce qui fascinait les gens entre autres, c'est que la façon dont on soigne un Tamagotchi enfant détermine le type de caractère qu'il aura au stade final d'adulte.

Au départ, après que le Tamagotchi a été activé, un œuf apparaît. Après avoir ajusté son horloge, l'œuf éclot en 5 minutes. Les premières heures de la vie de la petite créature exigent une attention constante de la part de l'interacteur. Il faut surveiller les indicateurs des besoins et y subvenir. Il faut jouer avec lui jusqu'à ce qu'il soit content, le nettoyer après les selles. Après 20 à 30 minutes, il s'endort; après 5 minutes, il s'éveille. Par la suite, le Tamagotchi n'est plus aussi exigeant. Toutes les interactions servent à déterminer le type de caractère qu'il aura une fois la phase adulte atteinte, ce qui prend 6 à 8 jours habituellement. Durant ce temps,



il faudra discipliner le jeune qui réclame de l'attention. Si on ne fait pas suffisamment attention à lui, il pourra, par dépit, refuser de manger et de jouer. Il est aussi possible de rendre le Tamagotchi heureux en lui donnant des bonbons au lieu de jouer avec lui. Mais il est préférable de jouer le jeu selon les règles, et de garder le poids du Tamagotchi à son niveau normal afin qu'il demeure en bonne santé. Le succès fulgurant qu'ont connu les Tamagotchis vient du mélange de deux ingrédients: la simulation d'un personnage (par un microprocesseur de faible coût) et l'implication immersive des utilisateurs, et ce malgré le côté minimal de la personnification et la simplicité de l'interface. Non seulement le jeu est immersif, mais il vient aussi chercher le joueur dans sa réalité pour le ramener dans l'univers du jeu.

5.7 Bibliographie

Buci-Glucksmann, Christine, *L'art à l'époque du virtuel*, Paris : Harmattan, 2003, 243 p.

Burdea, Grigore et Philippe Coiffet, *La réalité virtuelle*, Paris : Hermès , 1993, 402 p.

Coiffet, Philippe, *Mondes imaginaires: les arcanes de la réalité virtuelle*, Paris : Hermès , 1995, 127 p

Fuchs, Philippe, Guillaume Moreau et Bruno Arnaldi, *Le traité de la réalité virtuelle*, Paris : Presses de l'École des mines, 2003 (2 volumes).

Jolival, Bernard, *La réalité virtuelle*, Paris : Presses universitaires de France (Que sais-je?), 1995, 123 p.

Mairlot, Fernand E., *La nouvelle cybernétique: essai d'épistémologie des systèmes dynamiques*, Bruxelles : Chabassol, 1982, 297 p.

Pracontal, Michel de, *L'homme artificiel : golem, robots, clones, cyborgs*, Paris : Denoël, 2002, 217 p.

Quéau, Philippe, *La planète des esprits : pour une politique du cyberspace*, Paris : Odile Jacob , 2000, 329 p.

Rheingold, Howard, *La réalité virtuelle*, Paris : Dunod , 1993, 413 p.



Vignaux, Georges, *Du signe au virtuel : les nouveaux chemins de l'intelligence*, Paris : Seuil, 2003, 216 p.



Comprendre les médias interactifs de Louis-Claude Paquin est mis à disposition selon les termes de la licence Creative Commons 4.0 : Attribution - Pas d'utilisation commerciale – Pas de modification.