

## DOSSIER RECHERCHES ACTIONS (2)

### ANALYSE DE TEXTE : ÉCLAIRAGE

# LES JOURNÉES INTERNATIONALES SUR LITTÉRATURE ET INFORMATIQUE

(Littérature générée par Ordinateur).

Denis FOUCAMBERT

Les 20, 21 et 22 avril 1994 se sont tenues à l'Université Paris VII les premières rencontres entre informaticiens et littéraires pour présenter les réalisations dans ce domaine, les recherches en cours, ainsi que pour étudier les modalités de constitution d'un éventuel groupe international de recherche en ce domaine.

Les différentes interventions peuvent être divisées en trois parties :  
Les remises en cause du support-livre.  
Les recherches sur la génération automatique de textes littéraires.  
Quelques exemples de programmes existants.

#### I. Les remises en cause du livre papier

Un certain nombre de prise de paroles ont remis directement en cause le support classique de la littérature : le papier. Ces intervenants, pour la plupart poètes "confirmés", considèrent que, la forme littéraire s'épuisant, il convient de changer de médium et d'emprunter l'ordinateur. Il ne s'agit pas ici de générer des textes de façon automatique, mais de se servir de ce nouveau support pour aller vers une poésie dite "animée", "visuelle", pour tenter de créer une littérature plus interactive. On intègre donc des images (le texte devenant aussi, en lui-même, une " image visuelle "), des sons ; on travaille sur la vitesse de présentation, on module. Bref, nos chers artistes s'en donnent à cœur-joie dans la création multimédia, très au goût du jour... Deux revues françaises (Kaos et Alire) existent et proposent sur ce sujet des réalisations poétiques sur disquettes (il semble que ce soit une exclusivité française).

Rien de bien novateur dans une optique purement informatique, mais ce nouveau mode d'expression peut produire des résultats amusants et permet à leurs auteurs de se démarquer dans leur champ originel tout en découvrant les possibilités de ce nouveau mode d'expression.

Ce nouveau support apporte avec lui des possibilités nouvelles comme celle de l'hypertexte. Il en a été question durant ces journées, mais peut-on considérer ces techniques comme de la génération de texte ? Je répondrai par la négative, sans dénier un certain intérêt à un fonctionnement hypertextuel, qu'il faudrait d'ailleurs étendre à un fonctionnement hypermédiatique : le livre de Carlos Fuentes Terra Nostra sur support informatique avec un fonctionnement hypermédiatique, voilà une véritable gageure qui, si elle était menée à bien, permettrait une véritable " lecture savante " de cet ouvrage faisant appel à d'autres écrits, à des peintures, à l'histoire... De même pour le roman de A. Carpentier Concert baroque, avec ses nombreuses références musicales. Pareillement, quelle puissance apporterait dans la navigation entre les articles, l'Encyclopédie Universalis en hyper-média...

Mais pour autant, il ne semble pas que l'on puisse entrer cette catégorie dans la création, mais plutôt dans l'éclatement du support pour associer la littérature à d'autres disciplines créatives.

## II. Les recherches sur la génération automatique de textes littéraires

Ce domaine provoque le plus de recherches de pointe actuellement. Comme le terrain commence tout juste à se structurer, comme la masse de problèmes à résoudre est relativement importante, des interventions ont porté sur des tentatives de résolutions de difficultés relativement différentes. La façon d'aborder ces notions n'est, on s'en doute, pas dissociable des idéologies dominantes dans les différents pays représentés.

Le pays le plus en pointe est sans conteste les États-Unis. Des présentations de travaux exécutés au Massachusetts Institute of Technology ou au Rensselaer Polytechnic Institute essaient une approche comportementaliste (en cela, bien dans le courant dominant aux États-Unis) du fonctionnement de la génération automatique. On considère un personnage à qui on attribue un certain caractère et donc un comportement possible. Le travail de la machine consiste donc à trouver les réponses que ce personnage fictif devrait avoir dans des circonstances variées, devant des "stimuli" extérieurs. Il faut donc que la machine réagisse à des situations pseudo-réelles comme un être humain le ferait dans la réalité. On voit là poindre très fort des objectifs de l'Intelligence Artificielle, dont on sait par ailleurs que les cinq premiers plans de recherches se sont soldés par des échecs et que le sixième est en route...

Un point d'une importance pourtant primordiale semble être oublié par ces chercheurs : si on veut générer des textes, fussent-ils littéraires, on doit bien à un moment s'intéresser aux problèmes de générations de lexiques et de syntaxes et cela sans oublier le fait, bien prosaïque, que la succession de phrases ne donne pas forcément un texte...

Un intervenant de Néo-Zélandais a exposé un travail allant dans ce sens. Ayant mis au point une base lexicale et un générateur de phrases, il a fait "tourner" ces modules de production de phrases. Mais devant la pauvreté de ce que la machine produisait, il s'est orienté vers une approche plus "parodique" en examinant les schémas de phrases d'Hemingway pour construire un modèle des différentes tournures employées par cet auteur, et s'en servir comme modèles de phrases pour la génération. Cette approche plus simple, semble, au moins pour le moment, plus proche de la réalité...

## III. Quelques exemples de programmes

Il est curieux de constater que ces journées d'études sur l'informatique et la littérature ont nécessité l'emploi d'un simple rétroprojecteur bien plus intense que celui d'un ordinateur.

L'ALAMO a présenté ses trois logiciels de génération de textes, qui, même s'ils proposent une base de travail intéressante, pèchent par manque de bases lexicales et de générateurs de phrases sophistiqués (leurs concepteurs l'admettent bien volontiers).

Un autre produit, présenté par M. Brasseur, s'inspire fortement des travaux de l'OULIPO (que tous les intervenants européens semblaient connaître et à qui ils rendent un juste hommage pour leur vieil engagement dans ce domaine). Rien de bien nouveau : on peut, par exemple, recréer de façon automatique des produits textuels, le plus souvent déjà existants sur papier (ex : les mille milliards de poèmes). Un troisième logiciel créé par les Éditions Mireille et intitulé *Exégèse* de Mallarmé a été présenté par Michel Gautier de Paris V. Logiciel dit pédagogique, proposant une analyse détaillée d'un seul poème de Mallarmé (analyse fermée, où tout est prévu : le cheminement dans le texte, les commentaires asso-

ciés) qui semble en complète contradiction avec ce qu'on sait de la lecture et avec les Instructions Officielles concernant les classes de lycées pour lesquelles il est prévu.

On se rend rapidement compte de la distance énorme entre les objectifs de cette nouvelle discipline et ses réalisations.

Je voudrais rappeler ici que l'AFL, forte de ses travaux sur ELMO 2 000 peut déjà prétendre être au niveau de ce qui a pu être présenté, non sur le plan de la génération comportementaliste bien américaine, mais sur l'analyse et la génération de textes assistées par ordinateur. En effet, nous sommes déjà capables de générer un nombre infini de textes à partir d'un texte choisi comme point de départ et ce, de manière entièrement automatique. Peut-être faudrait-il mieux dire Transformations Textuelles plutôt que véritable Génération de Textes qui suppose qu'on parte de rien. Les règles de transformation s'appuient sur des contraintes oulipiennes, et une coopération est envisagée avec l'ALAMO pour développer ces produits, encore à l'état de brouillon.

Il est regrettable que les organisateurs de ces Journées n'aient pas songé à associer des linguistes. En effet, la notion de texte n'était visiblement pas entendue de la même façon par les intervenants et, parler de lexique, de phrases, de textes nécessite des définitions précises et la prise en compte des modalités propres à ces notions. On a parfois frisé le ridicule tant ces concepts étaient particulièrement malmenés. Il est par exemple frappant qu'on n'ait pas entendu parler de " progression thématique ", notion pourtant importante quand on parle de génération de texte. Ce n'est pas parce qu'on parle de littérature, de " création " qu'il faut oublier les contraintes propres au maniement de la langue écrite...

La dernière journée était consacrée aux suites à donner à cette réunion. Il semble que le pôle moteur soit situé de l'autre côté de l'Atlantique. Cette situation n'est guère favorable à la conception de générateurs de textes français : la langue dans laquelle s'effectuent les travaux n'étant pas la même, et une simple traduction des programmes absolument pas envisageable, les sociolinguistes et psycholinguistes ont montré les contraintes spécifiques à l'utilisation de tel ou tel idiome. Aussi est-il nécessaire de développer des outils français pour faire progresser les travaux envisagés, et n'en doutons pas, pour permettre aux travaux anglophones de progresser aussi...

Pour terminer, il convient sûrement de rappeler, comme certains intervenants l'ont fait, que se servir d'un ordinateur, ce n'est pas essayer de plagier le fonctionnement d'une création humaine, mais se servir des caractéristiques des ordinateurs pour ce qu'ils savent faire (où alors faudra-t-il attendre la 40<sup>ème</sup> génération de l'Intelligence Artificielle...). Cela peut sembler évident, mais certains représentants américains semblent vouloir faire de la machine un clone de notre organisation neuronale.

Denis FOUCAMBERT

Bibliographie :

- **Informatique et Littérature** (1950-1990), *Vuillemin Alain*, Coll. Travaux de Linguistique Quantitative n°47, Édition Slatkine, Genève 1990.
- **Littérature. Colloque Nord Poésie et Ordinateur**. *Collectif*. Édition Université de Lille 3
- **LINX Numéro 4 : Texte et Ordinateur : les mutations du lire-écrire**, *Centre de Recherches Linguistiques*, Paris X Nanterre.