

***Avant-propos***



Si je devais avoir un regret concernant le libellé de cette thèse, entreprise voici huit ans, il porterait sur l'impression qu'il donne que la modélisation du droit est accomplie et qu'il s'agit d'en déterminer l'apport sur le plan de la codification et de la simplification du droit. Or, la question fondamentale est celle de la modélisation, et ce n'est qu'après avoir tenté d'apporter une réponse sans doute partielle à cette question considérable, que l'on peut être en mesure d'apprécier ses implications possibles pour la codification et la simplification des textes normatifs.

Sous réserve de cette précision introductive relative au risque d'une interprétation un peu décalée du libellé lui-même, disons que la recherche a été conforme au projet initial. Nous voudrions rappeler les trois hypothèses qui sont à la base de la démarche entreprise, la méthode de recherche employée, les difficultés rencontrées, les principaux résultats obtenus et leur signification dans le contexte scientifique auquel ils se rattachent qui est celui des sciences cognitives et de l'intelligence artificielle.

Nous sommes partis d'une première hypothèse qui est celle de la possible modélisation du droit.

Cette hypothèse pourrait apparaître comme un truisme se passant de démonstration dès lors que comme le disait Valéry, on raisonne toujours avec des modèles. Le droit n'aurait aucune raison d'échapper à la règle. Toutefois, cette pétition de principe, que nous faisons nôtre, une fois posée, nous n'aurions rien résolu, car la question revêt une double dimension. La première, superficielle mais non triviale, est celle de la détermination du modèle, voire des modèles et de leur organisation. Et cette première dimension renvoie à la seconde plus profonde, qui est celle de définir cet objet que l'on souhaite modéliser en évitant toute approche indûment réductrice. Modéliser impose une certaine forme de dépouillement, de simplification, de systématisation, car cela est dans la nature même de l'activité de conceptualisation, dans la nature même des opérations cognitives de reconnaissance, d'identification, de conceptualisation et de modélisation. Comment donc modéliser sans mutiler, c'est-à-dire en préservant et en respectant l'intégrité, c'est à dire la complexité de l'objet que l'on cherche à modéliser. Or, au-delà des débats épistémologiques sur la modélisation analytique et la modélisation systémique, nous ne pouvons nous engager dans une démarche de modélisation du droit sans une profonde réflexion sur le droit en tant que système normatif.

Malgré la difficulté de l'entreprise, il est utile de signaler que le droit présente certaines caractéristiques qui le rendent propice à une modélisation.

Il y a d'abord le fait que le droit moderne est essentiellement écrit et que, bien que les sources en soient multiples, hétérogènes dans leur forme et dans leur portée, ces sources sont néanmoins accessibles.

Par ailleurs, le droit, même s'il évolue avec la société, au gré des mœurs et des technologies, se caractérise par une stabilité relative et contribue de ce fait à modéliser les comportements dans des structures répétitives.

En troisième lieu, le droit est régulé. Il a ses instances normatives de premier niveau qui constitue le législateur au sens large, incluant le pouvoir constituant, le pouvoir législatif et le pouvoir réglementaire. Il a ses instances de second niveau constitué par les juridictions avec les juridictions suprêmes chargées essentiellement d'assurer la cohérence

de l'ensemble du système appliqué par tous les administrés et justiciables dans leurs rapports respectifs et dans leurs rapports avec les autorités politiques et administratives.

En quatrième lieu, le droit est objectif dans ce sens que la règle n'est pas attachée à la personne, mais s'applique à la généralité des administrés et justiciables ou à des catégories objectivement définies. La règle est générale et impersonnelle, et si l'on doit reconnaître aux décisions individuelles ou d'espèce un caractère normatif, c'est dans la mesure où ces décisions sont toujours prises sur la base de règles générales et impersonnelles.

Enfin, le droit est décontextualisé, dans ce sens que la règle, sauf exception, n'est pas liée à un contexte ou une situation particulière. En tout cas, si la règle est souvent due à des circonstances particulières, la règle n'est qu'exceptionnellement temporaire et survit tant qu'elle n'est pas abrogée, annulée ou modifiée.

Ces grandes caractéristiques du droit emportent certaines conséquences au plan linguistique qui confèrent à la langue du droit, et particulièrement à la langue des normes, une certaine spécificité par rapport aux autres formes d'expression. En particulier, les dimensions argumentative et rhétorique très prégnantes dans le langage courant, dans le langage littéraire, journalistique ou publicitaire pour ne prendre que ces exemples, ne sont pas absentes du langage des normes, mais elles s'y trouvent codifiées, figées et nivelées de telle sorte qu'elles sont des dimensions sur lesquelles l'analyse pourra très largement faire l'impasse.

Il est clair que ces caractéristiques ouvrent de vrais espaces pour une modélisation susceptible d'avoir de retombées opérationnelles.

A l'inverse, le droit apparaît évolutif, souvent instable dans son interprétation, fondés parfois sur des concepts à la teneur indéterminée, il peut être incomplet, contradictoire, et nous pourrions pousser le paradoxe jusqu'à dire qu'il est structurellement contradictoire. En somme, le droit est structurellement complexe. Voilà qui invite à une certaine circonspection et qui pose un réel défi au modélisateur.

La seconde hypothèse de travail est celle de l'utilisation de l'outil linguistique à fin de modélisation des systèmes et sous-systèmes normatifs. Cette hypothèse est-elle scientifiquement légitime ? Et dans l'affirmative quels moyens retenir offerts par la linguistique qui est loin de constituer une discipline unifiée, cohérente, malgré des progrès prodigieux accomplis dans les dernières décennies.

Le choix que nous avons fait de lier la modélisation à l'analyse linguistique est né de l'observation de l'utilisation d'une méthode de conception de systèmes experts, la méthode K.O.D., pour l'automatisation de tâches de gestion de procédures d'attribution de prestations familiales (Claude Vogel, 1988, Beate Lippold, 1991, Christian Tremblay, 1992 p. 409 à 419). Au terme de l'analyse et de son implémentation informatique, il apparaissait que le système expert était capable de restituer les textes normatifs appliqués d'une manière qui n'altérait pas leur contenu. A partir du moment que l'on est capable, à partir d'une analyse textuelle, le cas échéant éclairée par des commentaires d'expert, de générer un texte conforme dans sa signification au texte original, on est en droit, sous réserve de mobiliser des outils linguistiques adéquats, de modéliser les textes originaux et d'en restituer le sens textuellement et sans trahison.

Le choix de lier la modélisation à l'analyse linguistique est par ailleurs sous-tendu par un choix épistémologique fondamental consistant à nier l'autonomie du niveau conceptuel par rapport au niveau linguistique. On peut certes imaginer une telle autonomie. Et c'est ce que nous faisons quand nous avons recours à des procédés graphiques ou à des conventions logico-mathématiques pour représenter des connaissances ou des raisonnements. Mais nous ne faisons par-là qu'inventer de nouveaux langages, artificiels et plus spécialisés que ne le sont les langues naturelles, mais ce sont des langues quand même qui ne peuvent être décrites qu'au moyen d'une langue naturelle. Ce qui veut dire que l'activité conceptuelle ne se développe et ne se réalise qu'au travers d'un langage, peu importe au demeurant qu'elle le soit dans différentes langues. C'est ce qu'exprimait avec force Lev Vygotski en disant que « la pensée ne s'exprime pas mais se réalise dans le mot » (1997, p. 431). Sans cette solidarité consubstantielle entre la pensée et le langage, il est clair que notre démarche serait vouée à l'échec. Reste qu'un échec toujours possible, pourrait s'expliquer non par une erreur sur cette hypothèse fondamentale, mais du fait de l'inadéquation ou d'une mauvaise utilisation des outils linguistiques disponibles.

La troisième hypothèse de travail, sous-tendue par les deux premières, est évidemment que l'analyse soit menée avec des moyens informatiques. A quoi bon en effet requérir toutes les ressources de la linguistique si ce n'est dans une perspective d'automatisation, alors qu'une analyse « conceptuelle » plus directe, déjà lourde en elle-même, permettrait d'atteindre plus facilement l'objectif. Le recours à la linguistique correspond à ce que l'on appelle en économie un détour de production. Il s'agit d'un investissement lourd qui est censé, en cas de succès, d'augmenter sensiblement notre capacité de traitement et de dégager des forces pour investir ensuite dans l'accessibilité et la simplification du droit. Cela dit, la question est de déterminer, s'ils existent, les outils informatiques de conception, de développement et d'exploitation, qui permettront de traiter ce type de problèmes.

En ce qui concerne la méthode employée, deux méthodes étaient a priori envisageables.

La première eut consisté à prendre un corpus prédéterminé constitué d'un ou de plusieurs textes, à en faire l'analyse linguistique, à en produire une modélisation et à soumettre celle-ci à différents tests.

La seconde eut été de partir d'une étude générale sur le droit et des différents types de textes normatifs, de prendre des exemples, de rechercher ensuite les outils linguistiques et informatiques, et de tester les résultats.

En fait, ni l'une, ni l'autre n'était réellement adapté à un sujet dont les trois hypothèses de base nécessitaient des recherches parallèles et en interaction réciproque. C'est donc une troisième voie que nous avons empruntée sans plan rigoureux préétabli, mais en adoptant une progression circulaire, itérative, couvrant un large champ, suivant les détours et parfois les caprices, que nous espérons maîtrisés, de nos propres interrogations et de nos accès de curiosité, et tournant le dos à une démarche analytique, purement séquentielle et linéaire.

Nos connaissances linguistiques étaient au départ faibles pour ne pas dire inexistantes. C'est donc par une exploration du champ de l'analyse linguistique que nous avons commencé.

Après que nous eûmes trouvé des outils scientifiques qui nous paraissaient présenter la rigueur conceptuelle et les qualités opératoires que nous recherchions, nous avons entrepris d'en faire usage pour analyser un texte de loi, la loi du 10 juillet 1989 sur l'éducation, et nous avons réalisé en partie le développement informatique correspondant. Et ce sont les nombreuses difficultés rencontrées au cours de cette première analyse sur les trois plans du droit, de la linguistique et de l'informatique, qui nous ont déterminé à engager une phase d'approfondissement dans chacun de ces domaines.

La thèse comprend quatre parties :

- le droit
- la langue
- la langue et le droit
- l'application

Si nous avons suivi la chronologie nous aurions dû commencer par la linguistique. Cependant, cette position aurait pu être interprétée comme une position a priori, alors que le caractère incontournable mais non exclusif de la linguistique pour la démarche entreprise, qui était loin d'être évident, de simple intuition au départ s'est affirmé au cours de la recherche comme un vrai paradigme.

La linguistique et l'informatique n'étant en définitive que des moyens, il paraissait plus logique de commencer par notre objet principal, le droit, et de poser le problème autant de possible dans toute son étendue et sa complexité.

En ce qui concerne donc le droit, nous nous sommes aperçu que, sauf à se cantonner dans des sujets très spécialisés, peu représentatifs et donc peu pertinents au plan heuristique, une analyse ne pouvait porter sur un seul texte sans considération du système juridique global dans lequel il s'insère. Un nouveau texte analysé doit donc être considéré comme un ensemble d'éléments nouveaux venant s'ajouter à, compléter ou modifier une base de connaissances préexistante.

Par ailleurs, le droit, considéré dans toute sa profondeur, ne peut être limité à sa base textuelle. La règle ne devient norme qu'après application, toute application, par ses destinataires directs comme par le juge, impliquant un acte d'interprétation. On peut, pour des raisons pratiques, ne considérer dans un premier temps que le texte, mais à la seule condition de lier potentiellement la règle à son interprétation, ce qui laisse entendre qu'alors que la règle peut être textuellement invariable, la norme qui lui correspond peut être évolutive. Cette idée n'est pas admise par tout le monde, et il serait méthodologiquement critiquable qu'au détour d'une recherche qui n'a pas pour objet la philosophie du droit, une option aussi fondamentale soit prise de façon non explicite. Il est clair que cette option n'est pas de nature à simplifier le travail du modélisateur.

Nous avons ensuite, au-delà des modes de raisonnement globaux développés dans le cadre l'interprétation, ébauché une analyse des raisonnements juridiques en considérant que la logique juridique embrasse aussi bien les opérations logico-discursives qui président l'édiction de la règle de droit et son application par ceux auxquels elle s'impose que le raisonnement tenu par le juge pour régler litige. On peut mettre ainsi en évidence des modes de raisonnement qui sont parfois de type déductif, mais sont très souvent non

déductifs. Ce qui est important de noter c'est que l'on a généralement affaire à des raisonnements complexes mettant en œuvre plusieurs modes de raisonnement que l'on pourrait qualifier d'élémentaires. Il faut également souligner que les raisonnements ne sont rien sans la base de connaissance qui doit être constituée de manière à fournir aux raisonnements tous les éléments nécessaires à leur fonctionnement.

Enfin, le travail d'analyse nous est apparu comme de nature essentiellement linguistique, mais non exclusivement linguistique, puisque qu'il faut être en mesure d'intégrer les cas d'application qui n'auront pas forcément une forme leur permettant d'être intégrés comme n'importe quel texte normatif.

Reste la question du rapport entre l'analyse linguistique et l'interprétation. Question redoutable parce que si la norme résulte de l'interprétation de règles, et si l'analyse linguistique est de même nature que l'interprétation, comme le pense Riccardo Guastini, qui voit dans l'interprétation une activité linguistique, cela veut dire l'analyse linguistique est susceptible de modifier la norme. Dire que l'analyse linguistique relève d'un niveau méta-juridique ne règle rien, car parler du droit, c'est aussi contribuer à le modifier ou à changer la vision qu'on en a. Nous avons cependant tenté de répondre à la question en distinguant dans l'acte d'interprétation une part strictement linguistique, fixant des interprétations évidentes ou décrivant des interprétations faites par d'autres, et une part créatrice variable qui change, ne serait-ce que marginalement, l'interprétation et qui n'est en rien de nature linguistique.

C'est à ce prix que le modélisateur, sans légitimité juridique particulière, au même titre que la doctrine, peut échapper au reproche déontologiquement inacceptable d'être un créateur de droit inavoué.

En ce qui concerne le volet linguistique, nous avons trouvé dans l'œuvre de Bernard Pottier et dans celle de François Rastier un ensemble rigoureux, puissant et cohérent permettant des développements pratiques à la hauteur de notre objet.

Plus précisément, nous avons repris les principes de la micro-sémantique avec la structure du sémème, la hiérarchisation des sémèmes au sein des taxèmes dans le cadre de domaines d'expérience, et les concepts dérivés d'isotopie et de polysémie.

Au niveau de la phrase, que nous rattachons au niveau d'analyse mésosémantique, la théorie des voix constitue un niveau d'analyse intermédiaire permettant l'identification de macro-opérations logico-discursives qui structurent les textes normatifs.

Nous rejoignons ici le niveau d'analyse textuelle et nous avons abouti à des résultats auxquels est parvenu J-M. Adam dans le domaine des textes narratifs et des textes descriptifs

Nous avons dans notre troisième partie tenté de caractériser de manière plus méthodique que dans la partie précédente les textes normatifs au plan linguistique, et nous avons retenu pour ce faire une grille d'analyse tenant compte de la triple dimension cognitive, argumentative et rhétorique

Les macro-opérations normatives que nous avons identifiées relèvent de sept grandes familles :

- les opérations définitoires

- les opérations déclaratives : principes généraux, déclarations de droit (droit de..., droit à..., garanties...)
- les opérations institutives
- les opérations explicatives (objet, objectif, champ d'application d'une réglementation)
- les opérations attributives (mission, attributions, compétences, droit, possibilité, etc.)
- les opérations descriptives (de structures, de procédures, d'activités)
- les opérations prescriptives (obligation, interdiction, autorisation, sanction, règle applicable)

Ces opérations ont leurs propres structures de données et leurs propres méthodes, ce qui permet de les considérer comme des objets à part entière, des sortes d'agents intelligents, d'interprétants, dont le rôle sera d'enrichir la base de connaissance et d'en permettre l'exploitation.

En ce qui concerne les aspects applicatifs, ceux-ci sont essentiellement d'ordre informatique, mais il nous paraît nécessaire d'en tirer les conséquences au plan de la codification et de la simplification.

Au plan informatique, nous arrivons, sans surprise, à un système d'un très haut niveau de complexité que l'on peut décomposer en deux sous-systèmes :

- le sous-système linguistique qui comprend lui-même quatre couches :
  - o une couche syntaxique
  - o une couche microsémantique (gestion des sémantèmes)
  - o une couche mésosémantique (gestion des taxèmes, gestion des quinze modèles sémantico-syntaxiques de la théorie des voix)
  - o une couche macrosémantique (gestion des sept familles d'opérations logico-discursives)
- un sous-système de gestion de la base de connaissance qui comprend lui-même deux sous-sous-système :
  - o la base de connaissance qui ne peut qu'être une base de données objet
  - o le sous-sous-système de raisonnement dont nous avons seulement amorcé un inventaire des modules possibles et qui est susceptible de faire appel à diverses technologies de programmation : procédurale, déclarative, connexionniste, objet. En fait ce sous-système sera essentiellement fondé sur une programmation orientée objet pour des raisons que nous ne pouvons expliciter ici, mais qui ont été excellemment exposées par Jean-Pierre Bourgois (1996, p. 433 à 467). Seule l'approche objet peut en effet permettre de répondre
    - à l'évolutivité du droit

- à l'incomplétude du droit
- au flou du droit
- aux contradictions du droit
- et d'intégrer les divers formalismes propres aux autres approches.
- Il convient également d'ajouter que seule l'approche objet autorise des développements parfaitement modulaires et distribués.

Compte tenu de l'étendue des analyses complémentaires à conduire et des développements informatiques à entreprendre, on ne s'offusquera pas du fait que la thèse ne soit pas accompagnée d'une application clé en main. Nous avons seulement mis au point un analyseur syntaxique complété d'un analyseur sémantique implémentant la théorie des voix et la gestion taxinomique. Nous nous sommes arrêtés à l'identification des sept structures normatives de base.

C'est dire que l'étude de l'impact sur la codification de la démarche que nous nous sommes employés à mettre en œuvre relève présentement d'une prospective conjecturale.

Il faut, nous semble-t-il, attendre des retombées tant en ce qui concerne la conception des codes qu'en ce qui touche à leurs fonctionnalités. L'impact peut être sensible à plusieurs points de vue :

- Du point de vue de l'élaboration des codes, le système doit, à partir d'une quantité variable de textes, produire une organisation globale cohérente et complète, et l'on peut concevoir que, sans altérer la liberté d'appréciation du codificateur, le système apporte à ce dernier une assistance bénéfique en terme de productivité et de sécurité.
- Plus généralement, c'est l'écriture du droit qui peut se trouver passablement transformée du fait de l'utilisation d'outils de cette nature. De même que dans un droit codifié, le rédacteur ne devrait penser les modifications réglementaires ou législatives que par rapport au code, on peut également imaginer qu'en présence d'un outil tel que celui ici envisagé, le rédacteur serait amené à tester l'impact des modifications textuelles sur l'architecture et la cohérence des normes existantes.
- Du point de vue de l'interrogation, on peut envisager que sur la base d'une question formulée en langage naturel, le système produise sous forme textuelle le droit applicable assorti des références textuelles. Il s'agit d'une perspective de développement qui demandera des mises au point qui ne sont pas sans analogie avec les techniques mises en œuvre par certains logiciels d'importation de site Internet et où l'utilisateur doit préciser le nombre de niveaux de dépendance qu'il souhaite inclure dans sa recherche. Toutefois, une question fermée, la réponse devrait tout simplement être la réponse à la question.
- Du point de vue de la résolution de problème, il est possible d'envisager une modélisation de décisions juridictionnelles types qui permettraient de traiter de façon quasi automatique au moins 90 % des recours qui ne soulèvent pas de question de principe nécessitant une interprétation créatrice du juge. Le mot « automatisation » ne doit pas inquiéter : il ne s'agit pas de remplacer le juge, mais de lui préparer le travail en respectant bien entendu sa pleine liberté d'appréciation et le principe de l'examen

particulier de chaque affaire qui s'impose tant à l'administration qu'au juge. On peut imaginer que l'outil, mis entre les mains des justiciables ou de conseils juridiques, pourrait simuler les décisions de justice, mieux qu'aucun recueil de jurisprudence existant ne le ferait, et pourrait peut-être constituer un moyen de prévention des litiges.

- Du point de vue de la décision administrative, on est en droit de penser que dans les administrations pourvues d'un service juridique, l'outil envisagé pourrait améliorer sensiblement la productivité. Dans les administrations non pourvues d'un service juridique étoffé, l'outil pourrait offrir le moyen aux petites cellules juridiques sans effectifs de remplir plus efficacement leur rôle de conseil préalable à la décision.
- Du point de vue de la formation, on peut imaginer de même que le même instrument mis à la disposition des professeurs et des étudiants pourrait constituer un puissant instrument de formation.

Ceci est présentement de la prospective, mais nous avons personnellement la certitude qu'il n'existe aucun obstacle scientifique ou philosophique à ce que cette prospective ne se réalise. Dans ce contexte, il est évident qu'il ne s'agit plus tellement d'envisager l'impact de ces développements sur la codification. En réalité, c'est d'un bouleversement des conditions même de l'accessibilité au droit qu'il s'agit, et de l'apparition d'une codification électronique, navigationnelle et hypertextuelle, qui n'aura plus qu'un lointain rapport avec la codification manuelle que nous connaissons.

Quant à la simplification, elle ne saurait résulter de manière automatique de la modélisation. Mais il est clair que la modélisation systématique du droit à laquelle tend le système ici décrit, est de nature à fournir, du fait de la mise en valeur des redondances, des contradictions et des zones d'ombre ou d'incertitude, des arguments et des moyens d'études à ceux qui sont en charge ou qui ont la volonté de simplifier un droit structurellement complexe.

En 1987, Hervé Gallaire (1987, p. 267) observait que la représentation des connaissances restait pour l'intelligence artificielle un sujet majeur, loin d'être résolu. « On ne dira jamais de quelqu'un qu'il connaît un sujet s'il n'est capable de répondre qu'à une liste prédéfinie de questions. La vraie connaissance suppose donc l'utilisation à bon escient des informations dont on dispose et le problème qui se pose aux informaticiens est de trouver des structures qui permettront non seulement le stockage des informations, mais aussi l'utilisation de celles-ci par la machine elle-même ». Si les recherches sur la logique résolvent assez bien le problème du raisonnement, elles restent un peu faibles du côté de la description des objets manipulés et des relations existant entre eux. Notre effort s'est porté pour l'essentiel sur le problème de la représentation des connaissances, et le fait que nous soyons arrivés à la conclusion que la modélisation qui répond le mieux à cette préoccupation est la modélisation objet, impliquant au niveau applicatif base de données objet et langages orientés objet, était une conclusion logiquement prévisible. Non seulement, le concept d'objet permet une représentation naturelle du monde réel, mais il est capable d'intégrer tous les autres formalismes qui se sont révélés adaptés à des modes spécifiques de raisonnement et non au raisonnement en général. De surcroît, en terme de conception et de réalisation, la conception et la programmation orientées objet, sont les seules à supporter, du fait de leur parfaite modularité des projets coopératifs hétérogènes de grande envergure.

La recherche ici présentée s'inscrit clairement dans cette problématique.

