

Du langage au modèle

Patrice Enjalbert , Bernard Victorri

Le travail que nous présentons ici s'inscrit dans un programme de recherche à long terme mené en collaboration par nos deux laboratoires, en vue de réaliser un système de compréhension automatique de textes capable de traiter des constats d'accidents de la route. La première étape de ce projet a consisté à définir une architecture d'ensemble, à partir d'une réflexion sur le rôle respectif que doivent jouer dans un tel système d'une part les informations purement linguistiques contenues dans les textes, et d'autre part les connaissances extra-linguistiques, spécifiques au domaine, nécessaires pour mener à bien la tâche de compréhension. L'article est consacré à la conception de cette architecture générale, préalable à la mise en place effective des différents modules qui la composent, actuellement en cours de réalisation. Ce n'est bien entendu que quand cette deuxième étape sera terminée que nous pourrons valider complètement notre approche, quitte à amender ou préciser certains aspects de l'architecture, décrite ici de manière encore prospective.

1. Sémantique linguistique et modèle du monde - architecture générale du système

Nous partons d'une conception constructive et dynamique de la sémantique: l'énonciateur vise à "mettre en scène" des événements en en spécifiant à la fois le contenu et le point de vue à partir duquel la scène doit être observée. L'auditeur/lecteur est invité à reconstruire cette scène et sa mise en perspective en utilisant le contenu des énoncés comme autant d'instructions qui le guident dans cette opération de re-présentation. Mais bien sûr la langue "ne dit pas tout": ces instructions convoyées par le matériel linguistique ne sont pas complètes, d'une part parce que tout ce qui supposé connu par les deux interlocuteurs (but de l'énonciation, connaissances communes, etc.) n'en fait pas partie, et d'autre part parce que le discours ne montre de la scène que ce qui est en est visible du point de vue que l'énonciateur a choisi de privilégier.

L'activité extra-linguistique participe donc à part entière à la compréhension de textes. La démarche n'est pas de rajouter de-ci de-là quelques connaissances et quelques inférences "au texte lui-même" pour aboutir à une bonne compréhension. On part d'un objectif, construire des représentations en accord avec la situation d'énonciation et nos connaissances sur le monde, et c'est dans ce cadre que l'on interprète au fur et à mesure les données textuelles. Mais notre conception de la sémantique ne se réduit pas non plus à une vision purement pragmatique de l'activité langagière. La structure linguistique du texte impose des contraintes fortes sur le processus d'interprétation. Ces contraintes, qui définissent ce que nous appelons le niveau "sémantique linguistique", se construisent elles aussi de manière dynamique dans l'interaction des différents éléments lexicaux et syntaxiques qui composent les phrases du texte. Elles ne se réduisent pas à de l'information que devrait prendre en compte le processus de construction de

la scène, mais elles constituent un ensemble d'instructions auxquelles ce processus doit se conformer.

Il y a donc finalement une interaction assez subtile entre le texte et les données extra-linguistiques: le texte apporte des règles de construction de la signification et ces règles s'appliquent, dans un cadre dépendant des conditions d'énonciation, à des entités et des événements intégrés dans une appréhension cohérente du "monde" dont on parle¹.

L'architecture que nous proposons est une tentative d'appliquer cette conception générale de la compréhension de textes à la situation bien spécifique du traitement automatique. Une caractéristique essentielle est de séparer nettement deux tâches: analyse linguistique du texte et construction de la scène, en chargeant deux sous-systèmes largement indépendants de les traiter (cf. figure 1). Nous pensons en effet qu'aussi bien la nature de l'objectif poursuivi que les outils à mettre en oeuvre diffèrent profondément de l'une à l'autre. Les deux termes ainsi clairement posés, l'interaction peut ensuite être étudiée proprement. Cette architecture est à rapprocher de celle que proposent Crouch et Pulman (1993), pour une application très différente (interface avec un système de planification).

Le premier sous-système ("Analyse Linguistique") intègre une analyse syntaxique *et* un certain niveau d'analyse sémantique. L'analyse sémantique "linguistique" doit s'appuyer essentiellement sur les mécanismes de la langue pour déterminer des valeurs (temporelles, aspectuelles, actanciennes, relatives à la détermination, etc.) qui, jointes aux informations lexicales, permettent de construire un ensemble de relations que doivent nécessairement vérifier les entités et événements auxquels le texte fait référence. Le résultat est un ensemble de *contraintes sémantiques* qui déterminent, mais seulement partiellement, la construction de la signification. On présentera le rôle exact de cet analyseur ainsi que les grandes lignes de son fonctionnement au §2.

Le second sous-système ("Construction de la Scène") est alimenté par trois types de données. D'abord des connaissances du domaine, qui correspondent pour la machine à ce que nous avons appelé les connaissances partagées par deux interlocuteurs. Pour notre corpus, il s'agira en premier lieu de représenter des entités géométriques (lieux), temporelles (instants et plages d'instant), matérielles (objets mobiles) et des connaissances générales sur ces entités: bref, nous avons besoin d'une *physique*, ou mieux d'une *dynamique* - sera-t-elle "naïve", "qualitative" ou "scientifique", nous reportons provisoirement le débat (voir §5). Il faut ensuite des entités (routes, carrefours, véhicules...) et des connaissances plus spécifiques de situations de conduite automobile. C'est cet ensemble de représentations et de connaissances qui sera par la suite désigné sous le terme de *modèle* (du monde de la route), présenté au §3.

Ce modèle, qui ne peut bien sûr constituer qu'une représentation très réductrice des connaissances du domaine, définit les limites de la compréhension des textes par le système. Ainsi, comme nous le verrons, nous ne comptons pas, dans un premier temps, modéliser les

¹ Que ce monde soit le monde "réel" ou un monde imaginaire: de toute manière, ce que nous appelons "la construction de la scène", c'est une opération mentale de représentation, dont la relation avec la réalité objective peut être plus ou moins complexe sans que cela change fondamentalement la nature de l'échange langagier qui vise avant tout à aboutir à des représentations mentales partagées entre les deux interlocuteurs.

intentions des conducteurs et rien de ce qui se rapporte dans les textes à ces intentions (*Je voulais éviter le véhicule...*, *Je me suis affolé*, etc.) ne sera pris en compte dans la construction de la scène, même si l'analyse linguistique pouvait les mettre en évidence.

Le deuxième type de données utilisé par ce sous-système concerne les conditions d'énonciation. C'est à ce niveau que seront fournies les informations sur le locuteur (conducteur d'un des véhicules impliqués dans un accident, identifié comme le véhicule A, etc.) ainsi que les conditions générales de la scène à construire (se situant dans le passé par rapport au moment d'énonciation, généralement sur une route, etc.). C'est aussi à ce niveau que nous pourrons, plus tard, spécifier les intentions probables du locuteur (se disculper vis-à-vis de sa compagnie d'assurance).

Enfin ce sous-système reçoit bien sûr les contraintes sémantiques produites par l'analyse linguistique. Le processus consiste alors en quelque sorte à compléter ces informations, en inférant à l'aide des autres sources de connaissance les données complémentaires nécessaires, de manière à construire *une représentation cohérente intelligible de la scène*. Le système peut alors exécuter les tâches demandées en sortie, comme de répondre à des questions sur les véhicules, leur position relative à tel ou tel moment etc., ou encore d'afficher une visualisation schématique de la scène, sous la forme d'un croquis représentant ces positions.

Une fois établie l'architecture générale, nous nous sommes attachés à définir correctement l'articulation entre les deux sous-systèmes ainsi découpés. Nous aborderons cette question dans le §4 sous deux rapports: à propos du *lexique* d'une part, qui par définition relève à la fois de la langue et du modèle; et à propos de la *représentation des procès*, qui concrétisent pour nous les contraintes sémantiques construites par l'analyse linguistique. Une autre caractéristique de notre approche doit être relevée à propos de la représentation de la sémantique lexicale: celle-ci est formulée dans les termes du modèle, c'est à dire dans un cadre conceptuel spécifique — et non dans un cadre à vocation universalisante ou "proche du langage", comme peuvent l'être les formalismes de représentation de connaissances basés sur des "primitives" conceptuelles du type proposé par Schank, Sowa ou encore dans ce même volume, par J-P. Desclés. Ce point, cohérent avec notre conception de la sémantique comme construction de représentations relevant d'un modèle du monde, sera également discuté au §4.

Certes bien des questions restent à résoudre, que seule la réalisation de l'ensemble nous permettra de mieux cerner. En particulier nous sommes conscients de ce que le découpage entre mécanismes linguistiques et connaissances du monde recèle une part d'arbitraire (voir §2.1). Il faut d'abord souligner à ce sujet que l'architecture proposée n'interdit nullement les interactions en retour du second système sur l'analyseur (traitements complémentaires "à la demande" par exemple, ou toute autre forme de coopération: les flèches du diagramme indiquent le flot de données principal, non celui du contrôle): nous verrons au cours de la réalisation la forme exacte à donner à ces interactions. Mais nous pensons absolument nécessaire, d'un point de vue méthodologique, de démêler les différentes instances qui interviennent dans l'interprétation des textes (bien mises en évidence dans ce volume par D. Estival et F. Gayral), et de faire sa part aux structures et mécanismes propres au système de la langue.

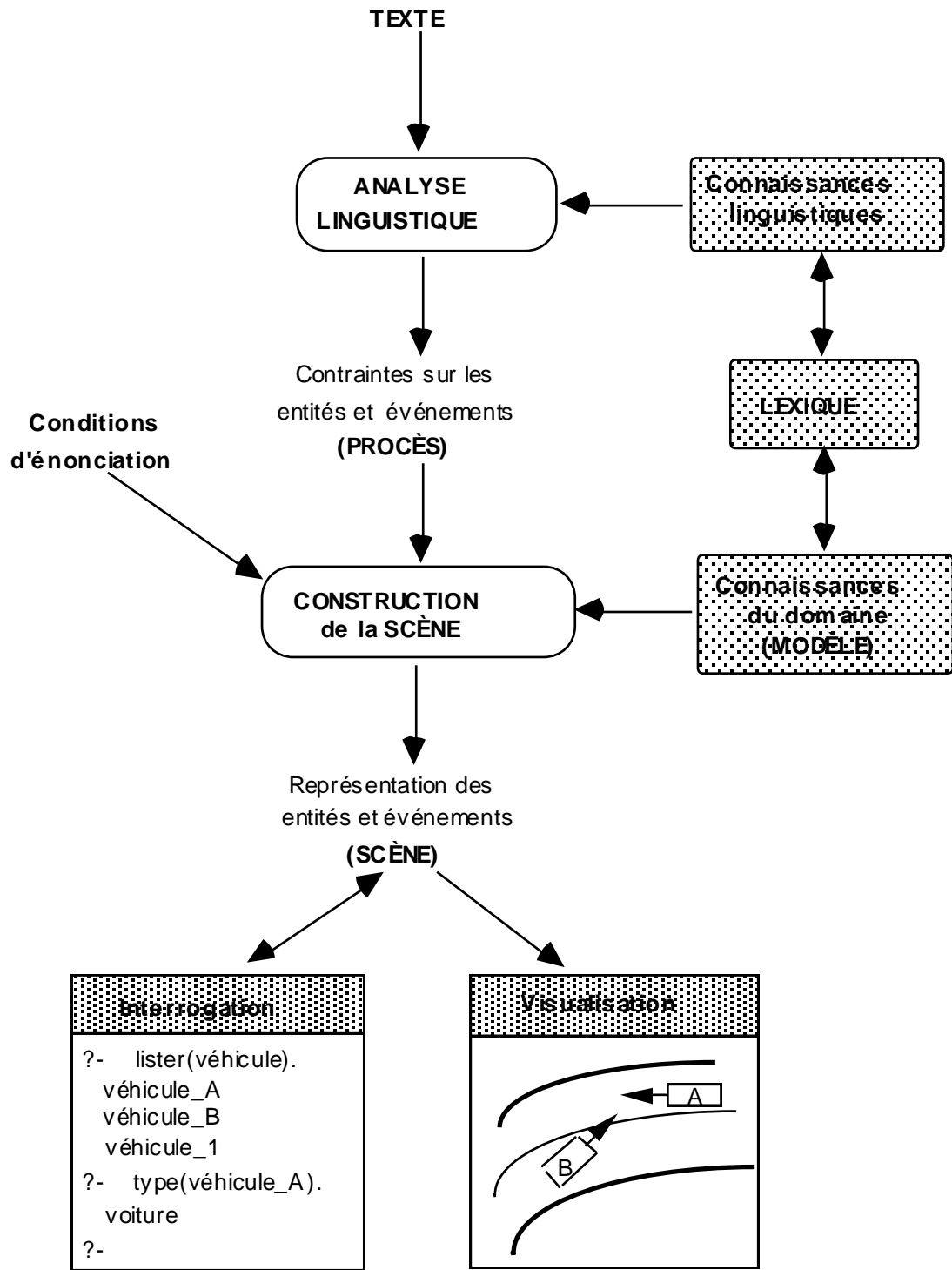


Figure 1: Architecture générale du système

2. Sémantique linguistique

2.1. Le rôle de l'analyseur linguistique

Le rôle du premier sous-système, l'analyseur linguistique, est donc de déterminer les contraintes calculables à partir des seules formes linguistiques présentes dans le texte qui constitueront l'une des entrées du deuxième sous-système responsable de la construction d'une scène cohérente. Quelques exemples permettront de mieux préciser ce point.

Soit la phrase *Le camion roulait à toute allure*. Les données linguistiques présentes ne permettent évidemment pas à elles seules d'estimer, même grossièrement, la vitesse du camion. Il faut prendre en compte des connaissances encyclopédiques (sur les camions, le code de la route, etc.), ainsi que la situation (cela se passe-t-il dans une agglomération, sur autoroute ?) pour pouvoir aboutir à une telle estimation. Cela ne signifie pas que *à toute allure* soit ambigu ou indéterminé; le sens linguistique de cette expression, valable dans toutes les situations, peut se gloser de la manière suivante: *à une vitesse qui, du point de vue de l'énonciateur, est beaucoup plus grande que celle à laquelle on pouvait s'attendre dans des conditions similaires, et proche du maximum des possibilités physiques de l'entité pour ces conditions de mouvement*. C'est ce sens, que l'on peut obtenir par des procédés linguistiques de comparaison de type paradigmatique (avec *à vive allure*, *à grande vitesse*, *à faible allure*, etc.) et de type syntagmatique (avec *l'homme courait / marchait à toute allure*, *le vélo / le train roulait à toute allure*, etc.), qui doit constituer la sortie de notre analyseur linguistique.

Prenons un autre exemple dans le domaine aspectuo-temporel: *Je traversais le carrefour quand un camion percuta ma voiture*. Là encore, rien dans les données linguistiques ne permet de dire si le procès *traverser le carrefour* (c'est-à-dire *aller d'un bord du carrefour à un bord opposé*) a vraiment eu lieu: seules des connaissances encyclopédiques conduiront à conclure que l'énonciateur n'a sûrement pas pu terminer son déplacement. Pour autant, les marques aspectuo-temporelles présentes dans cette phrase imposent des relations précises entre différents types d'intervalles de temps, les uns liés aux procès eux-mêmes, les autres liés aux instants sur lesquels portent l'assertion, même si l'ensemble de ces relations ne suffit pas à déterminer complètement à elles seules le cours des événements (cf fig. 2). Ainsi la marque de l'imparfait sur *traverser* indique-t-elle que l'intervalle de temps IA_1 sur lequel porte l'assertion est strictement inclus dans l'intervalle IP_1 correspondant à l'accomplissement du procès (ce qui explique que l'on ne puisse rien asserter sur la fin du procès). De même, si l'on appelle IA_2 et IP_2 les intervalles analogues pour le procès *percuter*, la marque du passé simple implique que ces deux intervalles sont égaux, et la construction de la proposition circonstancielle introduite par *quand* indique que les intervalles IA_1 et IA_2 sont consécutifs et adjacents (la borne supérieure de IA_1 est égale à la borne inférieure de IA_2). De plus, les intervalles IA_1 et IA_2 sont forcément antérieurs à l'instant d'énonciation, que l'on représente aussi par un intervalle, noté IE_0 . D'une manière générale, on peut, comme l'a montré Laurent Gosselin (1992, 1993, voir aussi W. Klein, à paraître, pour une approche similaire sur l'anglais et l'allemand), représenter l'ensemble du système aspectuo-temporel du français en termes de relations entre ces différents types d'intervalles: ce sont ces relations que nous inclurons dans les sorties de l'analyseur linguistique de notre système.

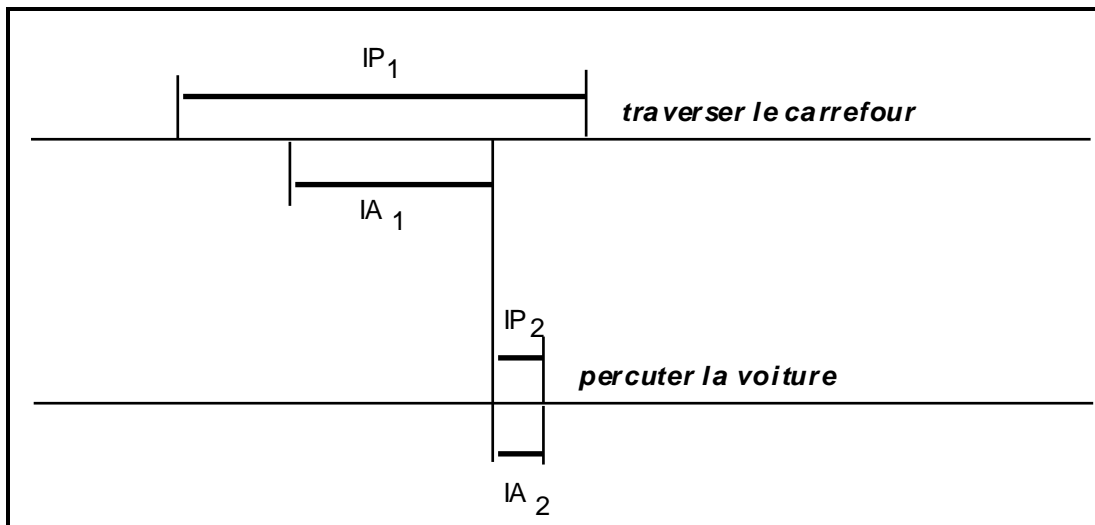


Figure 2: Relations aspectuo-temporelles correspondant à la phrase
Je traversais le carrefour quand un camion percuta ma voiture.

Troisième exemple: *Je roulais dans une rue étroite quand une voiture est arrivée en face.* Pour comprendre cette phrase il faut bien sûr représenter deux véhicules V_1 et V_2 et leurs deux conducteurs C_1 et C_2 , alors qu'à première vue, il n'y a de marques linguistiques que pour un conducteur (l'énonciateur C_1 , marqué par *Je*) et pour un véhicule (V_2 , marqué par *une voiture*). Comment obtenir dans la représentation de la scène le conducteur et le véhicule manquants ? Le problème est différent dans les deux cas. Pour l'existence du véhicule V_1 , il faut lever une ambiguïté de nature linguistique. En effet, la construction sujet humain + *rouler* + complément de lieu peut avoir deux sens: rouler sur soi-même, ou se déplacer dans un véhicule à roues (au sens large: vélo, voiture, train...). Le choix entre ces deux sens ne relève pas de l'analyseur linguistique (on reviendra plus bas sur ce point), mais on peut tirer de la forme linguistique utilisée l'existence de ces deux sens, dont l'un implique l'existence de V_1 . En revanche, rien de tel pour le conducteur du véhicule V_2 : il n'y a aucune marque linguistique de son existence, et il n'est d'ailleurs pas difficile d'imaginer l'emploi de la même phrase dans des contextes où il n'existerait pas (freins desserrés, voiture téléguidées, ...). C'est donc le système de modélisation qui devra dans ce cas inférer cette existence (à l'aide d'une règle par défaut: en l'absence d'indications contraires, les voitures ont des conducteurs dans le cadre des témoignages utilisés).

Comme on a pu le voir sur les exemples ci-dessus, notre méthodologie pour distinguer les caractéristiques qui relèvent du système linguistique de celles qui dépendent de nos connaissances sur le domaine repose sur l'application de deux critères essentiels:

1) Existe-t-il d'autres formes linguistiques que l'on peut opposer à la forme étudiée sur la caractéristique en question ? Ainsi le fait que *rouler*, dans la construction qui nous intéresse, s'oppose à *marcher/courir/voler/naviguer...* permet d'intégrer le trait 'véhicule à roues' dans sa définition linguistique, alors que ce n'est pas le cas pour des expressions comme *traverser le*

carrefour, tourner à droite, arriver à toute allure, que l'on peut trouver dans des emplois similaires.

2) Existe-t-il d'autres situations où la même forme linguistique peut être employée alors que la caractéristique en question n'est plus présente ? C'est ce critère qui nous force par exemple à renoncer à intégrer le trait 'conducteur' à une voiture en déplacement.

L'avantage de ces distinctions est double. D'une part, il permet de construire un analyseur linguistique qui ne soit pas complètement *ad hoc*, puisque cette méthodologie assure une relative indépendance de la sémantique linguistique par rapport au domaine et aux objectifs du système de compréhension. D'autre part, cela permet une grande liberté dans la construction du système de modélisation, en fonction du problème précis que l'on veut résoudre. Ainsi, si l'on décide de négliger dans un premier temps certains aspects des textes de témoignage d'accidents de la route (les aspects argumentatifs, par exemple), on peut le faire sans aucune incidence sur la conception de l'analyseur: les données linguistiques portant sur ces aspects seront tout simplement ignorées par le système de modélisation. De même toutes sortes de simplifications (sous forme de valeurs par défaut, voir §4.1) pourront faciliter le traitement à ce niveau: en somme, c'est le système de modélisation qui est *ad hoc*, ce qui est tout à fait normal puisqu'il ne s'agit pas d'un modèle général du monde, mais d'un modèle bien particulier adapté à la réalisation d'une tâche de compréhension précise et limitée.

Bien sûr cette architecture présente aussi des inconvénients. D'abord il n'est pas si simple de toujours séparer définitions lexicales et connaissances encyclopédiques, même si l'on s'appuie sur les deux critères exposés ci-dessus: en fait il existe nombre de cas intermédiaires pour lesquels la décision comporte une part d'arbitraire. Ensuite, la levée d'un certain nombre d'ambiguïtés purement linguistiques réclame la prise en compte de la situation. C'est ainsi que le choix entre les deux sens de *rouler* que nous avons relevés dans l'exemple présenté plus haut (rouler sur soi-même et se déplacer dans un véhicule à roues) ne peut s'effectuer qu'en faisant appel au système de modélisation: rouler sur soi-même n'est pas complètement impossible dans le domaine des accidents de la route, mais réclame un contexte bien précis (passager éjecté d'une voiture, piéton renversé, ...). Il faut donc prévoir une coopération souple entre les deux sous-systèmes qui élimine la plupart de ces ambiguïtés, pour éviter une lourdeur des traitements linguistiques qui rendrait le système difficilement utilisable en pratique. Cette coopération peut prendre plusieurs formes. D'une part, au niveau du lexique (cf. §4.1), on peut marquer les sens des unités lexicales qui sont les plus plausibles dans le domaine considéré: ils seront alors choisis en priorité par l'analyseur. D'autre part, comme nous l'avons déjà signalé ci-dessus, on peut mettre en place un système d'interactions entre les deux systèmes en cours de traitement: face à une ambiguïté du type de celle de *rouler*, l'analyseur peut faire immédiatement appel au système de modélisation pour trancher, en fonction des éléments contextuels déjà traités; réciproquement, si le système de modélisation tombe sur une incohérence, il peut faire appel à l'analyseur pour revoir certains choix de sens qui pourraient être cause de cette incohérence. Ainsi, les inconvénients créés par la séparation en deux sous-systèmes sont contournables, sans remise en cause de l'architecture générale : ils forcent au contraire à prendre conscience de difficultés essentielles du traitement de la langue qui sont le plus souvent éludées dans les analyseurs *ad hoc*.

2.2. L'architecture de l'analyseur linguistique

L'architecture de l'analyseur est fortement modulaire. Chaque module est chargé d'un type de problèmes précis, et il dispose pour cela de connaissances spécialisées. L'idée générale est de pouvoir utiliser de manière optimale les études "sectorielles" d'une très grande richesse menées par les linguistes depuis longtemps sur chacune des grandes questions de la sémantique linguistique: la détermination, l'anaphore, l'aspect, les modalités, le temps, l'espace... Les interactions entre ces modules sont nombreuses et relativement complexes, comme le montre la figure 3.

Le premier module occupe une place particulière: chargé de l'analyse morpho-syntagmatique, il prend en entrée le texte, et il le découpe en segments, en spécifiant des relations syntaxiques possibles entre ces segments. Il faut tout de suite préciser qu'il n'effectue pas une analyse syntaxique complète, mais un premier découpage superficiel, qui sert d'entrée aux autres modules. Ceux-ci sont donc des modules syntaxico-sémantiques, qui sont spécialisés chacun dans un type de problèmes précis, et qui coopèrent pour construire la liste d'entités et de relations qui constituent la sortie de l'analyseur. Dans la première version de l'analyseur sur laquelle nous travaillons, on compte quatre de ces modules: nous prévoyons d'augmenter ce nombre dans les versions ultérieures, pour prendre en compte des aspects que nous avons décidé de laisser de côté pour l'instant (en particulier, l'expression des modalités).

A l'heure actuelle, seuls le module morpho-syntaxique et un module syntaxico-sémantique (le module aspectuo-temporel) sont complètement spécifiés. Les autres modules, ainsi que le mode de communication entre modules, réclament encore une réflexion plus approfondie. La présentation qui suit a donc forcément un caractère quelque peu spéculatif.

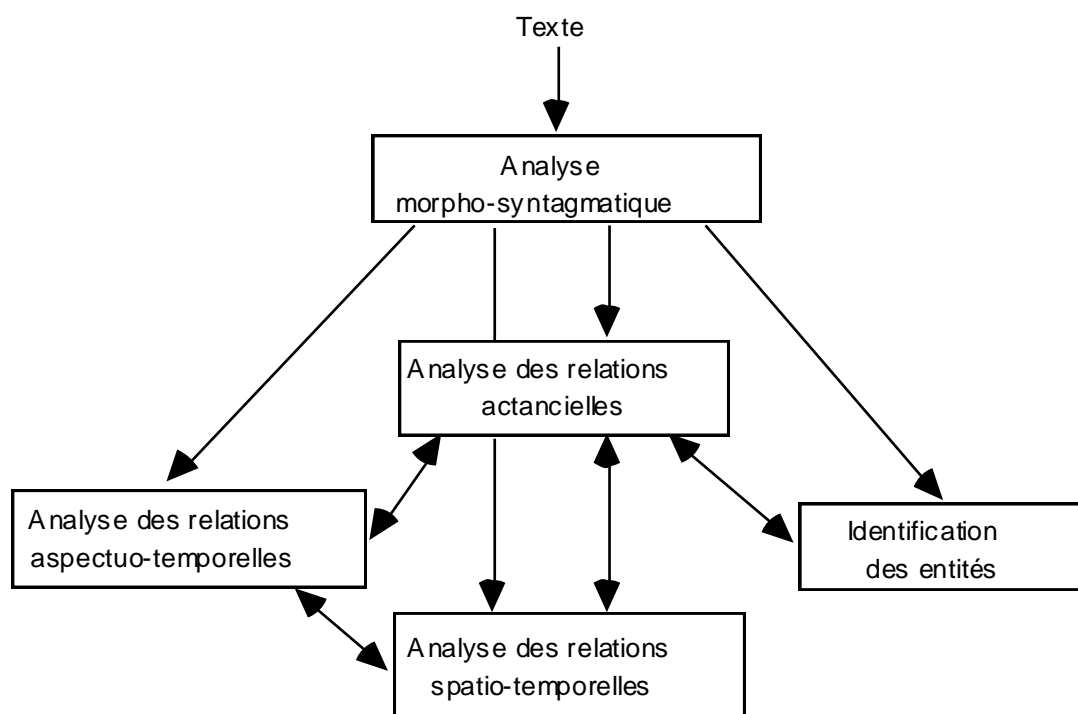


Figure 3: Architecture de l'analyseur linguistique

2.2.1 *L'analyse morpho-syntagmatique*

Ce premier module est donc chargé d'une première analyse du texte, qui n'utilise que des connaissances morphologiques et syntaxiques. Nous nous sommes largement inspirés des travaux de J. Vergne (1994) pour sa réalisation. L'idée de base est de regrouper les unités morphologiques en 'segments maximaux sûrs', c'est-à-dire en groupes d'unités dont les relations syntaxiques sont déterminées de façon univoque par la catégorie et la contiguïté des unités sur la chaîne. C'est le cas par exemple d'un verbe avec ses pronoms clitiques (*Je le lui donne*), d'un nom avec son déterminant, les adjectifs antéposés et éventuellement une préposition (*la jolie petite voiture, par un beau matin*), etc. Une quinzaine de classes de groupes ont été identifiés. Ce module ne résout pas tous les problèmes de rattachement entre ces groupes: il identifie simplement les 'relations potentielles' entre eux, c'est-à-dire les relations qui sont admissibles d'un point de vue purement syntaxique, sans choisir en cas d'alternative; ainsi dans la phrase *Je roulais sur l'autoroute à vive allure*, deux rattachements potentiels sont identifiés pour le groupe prépositionnel *à vive allure* (rattachement au groupe verbal *Je roulais*, et rattachement au groupe prépositionnel *sur l'autoroute*); ce sera la tâche des autres modules, qui disposent eux de connaissances à la fois syntaxiques et sémantiques de décider quel est le rattachement correct, si cette information est nécessaire à la poursuite des traitements.

Ce module opère en plusieurs étapes:

1. Découpage du texte en phrases
2. Reconnaissance morphologique: assignation à chaque unité d'une ou plusieurs catégories grammaticales possibles
3. Analyse en segments maximaux sûrs (avec désambiguïsation de la plupart des unités polycatégorielles)
4. Identification des relations potentielles, et regroupement (partiellement hiérarchisé) des syntagmes présentant des relations univoques.

2.2.2 *L'analyse des relations actancielles*

Ce module joue un rôle central dans la suite des traitements. En effet, d'une part il est chargé d'analyser la structure actancielle de chaque proposition (au sens large: procès, actants et circonstants), et d'autre part, il sert de contrôleur, faisant appel aux autres modules syntaxico-sémantiques pour résoudre des problèmes spécifiques. Le module prend en entrée la sortie de l'analyseur morpho-syntagmatique, et il s'appuie de manière essentielle sur des informations lexicales: en effet, pour chaque entrée lexicale (principalement les verbes, mais pas uniquement), sont recensées les constructions syntaxiques admissibles, avec indication des classes sémantiques possibles des composants (animé, humain, lieu, partie du corps, etc.). A l'aide de ces informations, le module doit construire à partir des relations potentielles entre syntagmes la structure actancielle de chaque procès: détermination du sujet, des compléments essentiels et circonstanciels de chaque verbe, rôle sémantique de chaque actant, type sémantique de chaque circonstant, etc. Comme on le voit, c'est d'abord et avant tout autour des verbes que cette construction est organisée. Il faut noter cependant que le même type de traitements s'applique aussi aux constructions à tête nominale ou adjectivale. De même il faut remarquer que ces traitements sont récursifs: la structure interne d'une subordonnée est ainsi rattachée au niveau supérieur à l'élément dont elle dépend.

La construction ainsi obtenue n'est pas complète: c'est justement l'un des rôles de ce module que de faire appel aux autres modules, dans des conditions qui seront précisées pour chacun d'eux, et ceux-ci, en retour, contribueront à affiner et à compléter l'analyse. D'autre part, comme on l'a vu, certaines ambiguïtés peuvent nécessiter un appel au système de modélisation. Ce n'est qu'à la fin de ce processus d'interactions que l'analyse sera terminée. La sortie de ce module est alors constituée par les relations entre les entités et les procès, à l'exclusion des relations aspectuo-temporelles et spatio-temporelles, prises en charge par les autres modules spécialisés, comme on le verra plus loin.

2.2.3 L'identification des entités

Ce module est chargé de la création d'entités nouvelles et du repérage d'entités déjà introduites précédemment dans le texte. Il est donc appelé par le module d'analyse des relations actanciennes pour chaque groupe susceptible de représenter une entité (en général, un groupe nominal ou un pronom). Ce module doit maintenir à jour une liste (historique) des entités qui ont déjà été créées, ainsi que des expressions linguistiques qui leur correspondent dans le texte, et il doit utiliser un ensemble de connaissances linguistiques sur les phénomènes d'anaphore pour résoudre le problème qui lui est posé. L'approche linguistique que nous avons choisie est fondée sur les travaux de M. Ariel (1990). Quand on a affaire à une anaphore, ce module doit renvoyer le symbole correspondant au module appelant, qui peut ainsi continuer et compléter son analyse, à l'aide des informations déjà obtenues sur cette entité. Dans le cas contraire, il doit y avoir création d'une nouvelle entité, à laquelle est associée un nouveau symbole typé selon la classe sémantique de l'entité. Parallèlement, l'indication de la création d'une nouvelle entité est ajoutée à la sortie de l'analyseur.

2.2.4 L'analyse des relations aspectuo-temporelles

Ce module, conçu à partir des travaux de Laurent Gosselin (1992, 1993), est chargé d'établir les relations temporelles explicitées dans le texte. Il est appelé par le module d'analyse des relations actanciennes pour chaque procès. Il utilise un ensemble de règles qui portent à la fois sur le type lexical de procès indiqué par le verbe, sur les marques de temps verbal, sur les déterminants des compléments essentiels, sur les adverbes et compléments circonstanciels temporels, y compris les propositions subordonnées. Son originalité essentielle, que nous avons déjà présentée rapidement au §2.1 sur un exemple (*Je traversais le carrefour quand un camion percuta ma voiture*), consiste à introduire deux types principaux d'intervalles de temps:

- d'une part, des intervalles, notés IP_i , liés aux bornes des procès: par exemple, pour *traverser le carrefour*, la borne inférieure de l'intervalle IP associé correspond au début et la borne supérieure à la fin de la traversée;
- et d'autre part des intervalles, notés IA_i , sur lesquels portent les assertions: dans notre exemple, à cause de l'imparfait, l'intervalle IA associé est strictement inclus dans IP ; en particulier, cela implique que l'on ne sait pas si la traversée s'est poursuivie jusqu'à son terme (la borne supérieure de IP ne fait pas partie de IA).

D'une manière générale, les relations aspectuelles sont des relations entre IP_i et IA_i , alors que les relations proprement temporelles (présent, passé, futur) sont des relations entre IA_i et le moment de l'énonciation, qui est aussi un intervalle, noté IE_0 (il peut d'ailleurs y avoir d'autres

intervalles d'énonciation IE_i , dans le cas de discours rapportés). De plus, d'autres éléments viennent complexifier le système: intervalles liés aux circonstanciels temporels, relations entre procès, etc.

Le rôle du module est donc de créer ces intervalles et d'établir les relations qu'ils entretiennent. Ces relations sont ajoutées à la sortie de l'analyseur, et l'intervalle de type IA_i est renvoyé au module appelant: en effet, toutes les relations actanciennes doivent être indexées par cet intervalle, puisque par définition elles ne sont assertées que pour cet intervalle de temps.

2.2.5 *L'analyse des relations spatio-temporelles*

Ce module est chargé des relations spatiales et en particulier des mouvements, ce qui implique aussi forcément le temps. Il est appelé par l'analyseur des relations actanciennes dès qu'un procès fait référence à l'espace, soit par le contenu lexical du procès lui-même (en particulier pour les verbes de mouvement), soit par l'intermédiaire d'un complément de lieu. Comme le module précédent, il doit alors utiliser un ensemble spécialisé de règles qui portent sur les différentes marques spatiales présentes dans la proposition étudiée pour en tirer les contraintes spatiales qui constitueront son apport à la sortie du système. En particulier, c'est ce module qui doit traiter les prépositions qui indiquent une localisation ou un mouvement (*à*, *dans*, *sur*, *vers*, etc.) pour les traduire par des relations spatiales précises, qui, on le sait, peuvent varier suivant le co-texte (groupe prépositionnel détaché ou intégré à la construction verbale, type spatial du complément nominal : ponctuel, linéaire, surfacique, tri-dimensionnel, lieu clos, ouvert, contenant, etc.)

L'une des difficultés essentielles provient du fait que les traitements de ce module sont centrés sur la notion de repère spatial, définis par une position donnée dans l'espace et une direction privilégiée: en effet, la plupart des indications spatiales (*à droite*, *en face*, *sur le côté*, etc.) sont relatives à un tel repère, qui peut être, suivant les cas, un repère lié à l'énonciateur, à l'un des actants en mouvement, à un point de vue extérieur d'où est décrite la scène, etc. Or, comme on le verra au §4.3, les données linguistiques sont en général insuffisantes pour déterminer à elles seules de quel repère il s'agit. Le module ne pourra donc que donner certaines contraintes explicites sur la nature de ce repère, et c'est bien sûr au système de construction de la scène de le spécifier complètement, en fonction de connaissances sur le domaine et de la cohérence globale de la scène.

Ce module doit pouvoir interagir étroitement avec le module aspectuo-temporel, dans la mesure où d'une part le contenu lexical des verbes de mouvement s'exprime par des relations qui font intervenir les bornes de l'intervalle IP lié au procès (et non pas l'intervalle d'assertion IA: cf. l'exemple de *traverser*), et où d'autre part la notion même de mouvement implique une correspondance entre événements spatiaux et événements temporels (comparer *Je roulai en direction de Paris jusqu'à midi* et *Je roulai en direction de Paris jusqu'au premier village*).

3. Un modèle du "monde de la route"

Au second sous-système revient la tâche de construire la scène décrite par le texte traité. Il doit s'appuyer pour cela sur une modélisation du domaine concerné, ici le "monde de la route et de la conduite automobile".

Nous partirons d'une évidence: la référence aux objets et phénomènes physiques (routes, véhicules, déplacements, chocs...) est dans ce contexte particulièrement prégnante. Or la géométrie analytique et la physique classique (cinématique et dynamique) proposent un cadre de modélisation de ces objets et phénomènes riche, bien maîtrisé et mathématisé - donc dans une forme adaptée au traitement informatique. C'est donc dans ce cadre que nous avons entrepris d'élaborer notre modèle, en nous focalisant sur la composante "physicaliste" des connaissances nécessaires à la compréhension des constats. D'autres composantes pourront ensuite être introduites: par exemple une théorie de l'action, nécessaire pour représenter de manière satisfaisante les verbes et locutions dénotant les actions de conduite automobile (cf. §4.3).

De manière générale nous voulons nous autoriser une lecture "filtrante", écartant certaines informations ou certaines "dimensions" du domaine concerné, en attente de développement d'une modélisation et de procédures de traitement adaptées. Ainsi, nous pouvons dans une certaine mesure et dans un premier temps faire l'impasse sur cette théorie de l'action. De la même manière, nous négligerons la fonction argumentative des textes du corpus, dont l'importance est à juste titre soulignée par D. Estival et F. Gayral. Nous sommes prêts à accepter une compréhension partielle, et à vrai dire nous voyons mal comment procéder autrement si l'on attend quelque résultat concret.

Nous commencerons par quelques indications (ou rappels) concernant les notions géométriques impliquées dans le modèle. Puis nous montrerons comment il est possible, sur cette base, d'élaborer une description des notions spécifiques de notre corpus (véhicules, routes avec leurs files, virages...) et de traiter les problèmes de localisation.

3.1. La base géométrique

Nous nous intéressons ici à la géométrie plane: il est rare en effet que le relief ou que les accidents de terrain interviennent dans les accidents et en tout état de cause, cela n'est pas le cas dans notre corpus. Mathématiquement, le plan est un ensemble p de points. Nous supposons connue la notion de *vecteur*: intuitivement, et pour simplifier, un couple de points M et N détermine un vecteur noté \overline{MN} , représenté traditionnellement par une flèche joignant M à N . Un vecteur possède une longueur, et un vecteur de longueur 1 est dit *unitaire*, représentant de manière très satisfaisante la notion intuitive de direction.

3.1.1 Repères et coordonnées

Un *repère* sera pour nous la donnée (A, \vec{d}) d'un point A et d'un vecteur unitaire (ou *direction*) \vec{d} . Tout point du plan est alors parfaitement déterminé par la donnée de ses *coordonnées polaires* (figure 4.a): le couple (ρ, θ) formé d'un nombre réel positif ρ égal à la

longueur de \overrightarrow{AM} (le *module*, noté $|\overrightarrow{AM}|$ et d'un angle θ (*l'argument*, $-180^\circ < \theta \leq +180^\circ$), angle formé par \vec{d} et \overrightarrow{AM} . Un tel repère possède une interprétation naturelle simple: l'origine A est la position d'un observateur, et la direction \vec{d} est une direction privilégiée du lieu ou d'un objet qui s'y trouve.

Pour repérer des vecteurs, la donnée d'une origine n'est pas nécessaire: on parlera alors de repère vectoriel, réduit simplement à la donnée d'une direction \vec{d} . Les coordonnées polaires (ρ, θ) d'un vecteur \vec{v} relativement à un repère sont sa longueur (ou module) $\rho = |\vec{v}|$, et l'angle θ entre \vec{d} et \vec{v} (figure 4 b).

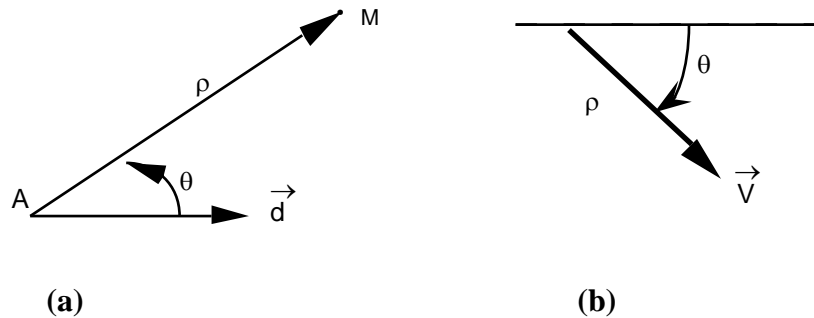


Figure 4: Coordonnées polaires d'un point M et d'un vecteur \vec{v}

3.1.2. Courbes et rubans

Rappelons d'abord qu'une courbe est, du point de vue mathématique, un ensemble de points possédant certaines propriétés de "linéarité" et de "continuité" et que tout point d'une courbe est déterminé par son *abscisse curviligne*: de même que l'abscisse d'un point A sur une droite orientée mesure la distance algébrique entre une origine donnée 0 et A, de même l'abscisse curviligne d'un point sur une courbe représente en quelque sorte le "chemin parcouru" depuis une origine donnée sur la courbe. En outre, en tout point nous pouvons définir une tangente à la courbe C, ou plus précisément un vecteur tangent (unitaire). Si C est une courbe, on notera C(s) le point M d'abscisse curviligne s et $\vec{T}(s)$ le vecteur tangent à C en ce point (figure 5).

Nous pouvons maintenant définir la notion de *ruban*, qui sera l'idéalisation géométrique d'une route. On pourrait dire qu'un ruban est en quelque sorte la donnée de deux courbes "parallèles", ce que l'on peut formaliser de la manière suivante. Un ruban r est la donnée:

- d'une courbe C que nous appellerons la *directrice* et,
- pour chaque point d'abscisse curviligne s, de deux nombres positifs $l_g(s)$ et $l_d(s)$, respectivement les *largeurs gauche* et *droite* en s — permettant d'obtenir de manière évidente les deux bords, gauche et droit, du ruban (figure 5).

Il existe une famille de repères naturels associés à un ruban r, définis comme suit. Étant donné le point A d'abscisse s de la directrice et \vec{T} le vecteur (unitaire) tangent en A, (A, \vec{T}) est le *repère axial du ruban en s*. Intuitivement, \vec{T} désigne la "direction" du ruban en A. Cette famille de repères nous fournit un moyen commode pour situer tout point P du ruban. Soit s l'abscisse curviligne du point A de C "à l'aplomb de P" — c'est-à-dire tel que la perpendiculaire à \vec{T} en A passe par P. Les coordonnées de P relativement au repère (A, \vec{T}) sont nécessairement de la forme (ρ, θ) avec $\theta = +90^\circ$ ou -90° et P est parfaitement déterminé par le couple (s, y) où $y = +\rho$ ou $-$

ρ selon le signe de θ . Clairement, si la directrice n'est pas trop courbe et la largeur du ruban point trop grande ce repérage devient biunivoque pour les points du ruban. Intuitivement, si r représente une route, (s,y) étant les coordonnées du "centre" d'un véhicule, s sera le "point kilométrique" et y déterminera la position latérale dudit véhicule sur la route. Finalement, notons que nous avons un moyen simple de caractériser un ruban "courbe" au sens usuel d'un ruban "droit": dans le second cas, $\bar{T}(s)$ est constant.

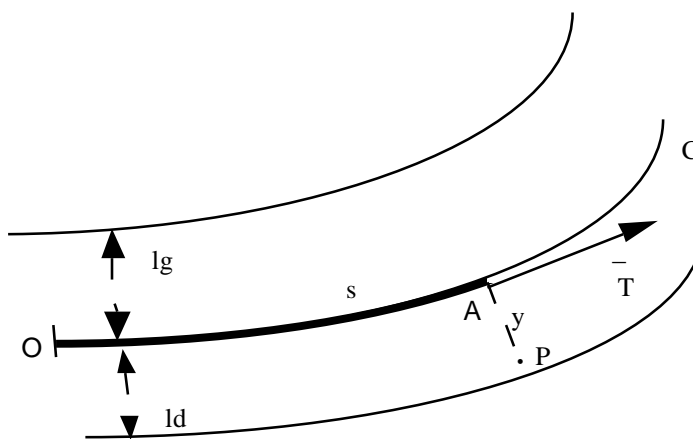


Figure 5: Abscisse curviligne et repère axial d'un ruban

3.1.3 Points et vecteurs mobiles

Il nous faut maintenant considérer des points et vecteurs évoluant dans le temps: par exemple pour repérer la position, ou la direction d'un véhicule à différents instants. Nous adoptons la formalisation classique du temps en physique: le temps est constitué par l'ensemble des *instants*, assimilés eux-mêmes aux nombres réels. La notion d'*intervalle de temps* se confond alors avec celle d'intervalle de réels: $[t_1, t_2]$ où t_1 est l'instant initial et t_2 l'instant final. Un *point mobile* est tout simplement la donnée d'un point du plan à chaque instant (dans un certain intervalle de référence), soit mathématiquement, une fonction:

$$M: [t_1, t_2] \longrightarrow p$$

$M(t)$ est la *position* du point mobile M à l'instant t . À tout point mobile est associé un *vecteur vitesse*, lui-même dépendant du temps, dont la direction représente la direction du déplacement à chaque instant et dont le module mesure la rapidité du mobile.

Finalement, un vecteur mobile est une fonction:

$$\vec{w}: [t_1, t_2] \longrightarrow \text{Vect}$$

prenant ses valeurs dans l'ensemble des vecteurs du plan. $\vec{w}(t)$ est le vecteur instantané à l'instant t . Par exemple la vitesse d'un point mobile ou, comme nous le verrons, l'orientation d'un véhicule sont des vecteurs mobiles.

3.2. Routes et véhicules

Considérons maintenant le "monde de la route" en tant que tel. Nous ne prétendons pas que ce domaine puisse être totalement "mathématisé", sur le modèle de la physique

scientifique. Mais les outils géométriques introduits dans le précédent paragraphe nous fournissent un cadre général de représentation des objets matériels et de leurs mouvements qui sera en quelque sorte le "noyau dur" du modèle. Nous allons montrer comment définir dans ce cadre les concepts essentiels de *véhicule* et de *route*, et comment traiter certains problèmes de *localisation*.

Préalablement, nous devons introduire un langage de représentation des connaissances tel que le requiert un traitement informatique. En effet, le lecteur a pu l'observer, notre discours a jusqu'ici été mathématique mais non formel —c'est-à-dire codé dans un langage formel de type logique ou informatique. Il est très important d'opérer cette distinction: on ne voit pas que des mathématiciens ou des physiciens s'expriment en formules logiques; par contre, le traitement informatique nécessite une représentation des connaissances entièrement formalisée.

3.2.1 *Le langage de représentation de connaissances*

Nous utiliserons une logique typée et un langage fortement inspirés du projet LILOG (Herzog O. & Rollinger C-R. 1991). Les objets de l'univers de référence sont répartis en classes, identifiées par un *type*. Les opérateurs du langage (symboles de constantes, fonctions et prédicats) sont eux-mêmes typés. Les règles de formation des formules sont les règles usuelles en logique, en respectant bien sûr le bon typage des arguments des fonctions et prédicats.

À chacune des notions mathématiques précédemment introduites correspond un type: *ty_point*, *ty_vecteur*, *ty_direction* (vecteurs unitaires), *ty_ruban*, *ty_instant*, *ty_point_mobile*, *ty_repère*... Nous utilisons en outre un dispositif inspiré des langages à objets pour structurer la déclaration des différents symboles du langage. De manière simplifiée, une déclaration de type aura le format suivant:

```
type ty_0
  attributs: attr_1: ty_1, ... , attr_n: ty_n
  propriétés: prop_1, ... , prop_p
```

qui intègre à la définition du type *ty_0* la déclaration d'un ensemble de symboles de fonctions: *attr_1: ty_0* —> *ty_1*... et de prédicats: *prop_1*... sur les objets de type *ty_0*. Mais on peut aussi lire la déclaration d'un *schéma*², associant à tout objet de type *ty_0* un ensemble d'attributs et de propriétés. Voici un exemple:

```
type ty_vecteur
  attributs: module: ty_nombre_positif, argument: ty_angle
  propriétés: nul, unitaire
```

² Au sens de ce terme en Intelligence Artificielle: un schéma ("frame") est un ensemble structuré de données définissant un type d'objet en même temps que ses composants ou que diverses relations définitoires avec d'autres objets. Par exemple, une voiture possède une carrosserie, un un moteur, des roues... et en principe, lorsqu'elle est "en service", un conducteur. Un schéma peut avoir une caractère de *typicité* dans sa configuration générale ou dans certaines des valeurs de ses attributs (par exemple, le nombre de roues: quatre en général).

À partir de types précédemment définis on peut construire de nouveaux types "complexes". Ainsi un type de la forme $t1 \rightarrow t2$ est celui des fonctions prenant leur argument dans les objets de type $t1$ et retournant une valeur de type $t2$. Par exemple, on peut définir:

type $ty_point_mobile = ty_instant \rightarrow ty_point$

Notation: Si V est un terme de type $ty_vecteur$, la notation $V.module$, $V.argument...$ désigne les attributs correspondants, soit formellement: $module(V)$, $argument(V)...$

3.2.2 Véhicules

Nous pouvons appréhender le concept de véhicule de plusieurs points de vue, notamment:

- comme objet matériel (solide) avec un encombrement (longueur, largeur, hauteur) et une masse —distinguant par exemple une voiture d'un camion;
- comme objet mobile, avec une position, une orientation et une vitesse instantanées. La position et la vitesse seront pour nous celles d'un point déterminé du solide, par exemple son centre de gravité (G), tandis que l'orientation \vec{o} sera la direction correspondant à la plus grande dimension (figure 6.a).
- comme moyen de transport, avec un conducteur et des passagers.

Dans le "schéma" associé au concept de véhicule nous allons regrouper les attributs et propriétés sous différentes rubriques correspondant aux points de vue répertoriés. Nous aurons alors le type suivant:

type $ty_véhicule$

point_de_vue matériel

attributs longueur, largeur, hauteur: $ty_longueur$
masse: ty_masse

point_de_vue cinématique

attributs position: ty_point_mobile
orientation: $ty_direction_mobile$
vitesse: $ty_vecteur_mobile$

point_de_vue moyen de transport

attributs conducteur: $ty_instant \rightarrow ty_personne$
passagers: $ty_instant \rightarrow ensemble\ de\ ty_personne$

Remarque:

Il ne s'agit là que d'une partie des opérateurs faisant intervenir le type des véhicules. Toute fonction ou relation, avec un nombre quelconque d'arguments peut être définie.

Par exemple: $est_dans(V:ty_véhicule, L: ty_lieu)$ pour exprimer qu'un véhicule se trouve en un lieu donné (route, carrefour...), etc. Nous avons simplement voulu insister ici sur la partie de ces déclarations qui peut être intégrée dans le "schéma" typique d'un véhicule. Divers axiomes peuvent également compléter la définition d'un type d'objet (par exemple ici: $\forall X: ty_véhicule \rightarrow à\text{-roue}(X)$).

3.2.3 Routes

De la même manière, une route peut être appréhendée de différents points de vue, c'est-à-dire s'intégrer dans différents réseaux ou systèmes conceptuels. On distinguera notamment:

- un point de vue *géométrique*: c'est celui pour lequel nous avons introduit plus haut un certain nombre d'objets mathématiques.
- un point de vue *cartographique* ou *géographique*: une route permet de relier un ensemble de villes ou villages, on "prend" une route pour se rendre en un lieu donné... Voir par exemple: *Me rendant à Beaumont sur Oise depuis Cergy. Je me suis retrouvée à un carrefour juste après la sortie Beaumont sur Oise* (Texte 1) ou: *le chemin de Condos* (Texte 9). Néanmoins cet aspect est secondaire dans les constats et nous ferons l'impasse dans un premier temps. Nous aurons alors le schéma suivant:

type ty_route

point_de_vue géométrique

attributs bord_droit, bord_gauche, directrice: ty_courbe

largeur_gauche, largeur_droite: ty_longueur

est_droite, est_courbe: ty_nombre → ty_booléen

Commentaires:

- les attributs directrice, largeur_gauche, largeur_droite correspondent aux primitives mathématiques de description d'un ruban. Dans le cas d'une route à double sens, la directrice est la médiane (matérialisée ou non).
- *bord_droit*, *bord_gauche* sont en un sens redondants, mais correspondent à des notions importantes en situation réelle de conduite.
- *est_droite*, *est_courbe* sont des propriétés de la route dépendant de *l'abscisse curviligne* du point considéré. D'où leur définition comme "attributs" et non comme "propriétés".
- D'autres attributs ou propriétés peuvent bien sûr ajoutés comme le nombre de files, la présence d'un médiane matérialisée, le caractère "à double sens" ou non de la circulation etc.

3.2.4 Localisation

Nous considérons ici la classe d'expressions du type: devant, derrière, en face, à droite, à gauche... (qualifiées de "projectives" dans Briffault X., 1992). Le problème crucial est de déterminer le référentiel utilisé par le rédacteur. Première remarque: la notion mathématique de *repère* nous fournit une formalisation de celle intuitive de *référentiel*. Soit alors un point M, et (A, \vec{d}) un repère; nous obtenons les caractérisations suivantes (figure 6. b):

M est *devant* par rapport à (A, \vec{d}): $|\theta|$ petit c.a.d. $|\theta| < \theta_{\min}$.

M est *à droite* par rapport à (A, \vec{d}): $\theta < 0$.

et de manière symétrique: M est *derrière* ou *à gauche*.

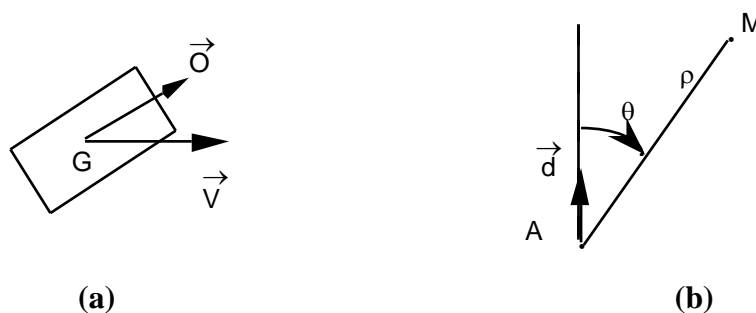


Figure 6: Orientation d'un véhicule et référentiel de relations spatiales

Remarques:

- Observez une certaine *compositionnalité*. Par exemple: "en face sur la droite" se code: $\theta_{\min} > \theta > 0$.
- On peut adjoindre une notion de "champ de visibilité": quelque chose comme " ρ petit".
- En pratique les prédicats *à droite*, *à gauche* etc. sont graduels: une position peut être "un peu", "plutôt", "franchement"... à gauche, à droite, en face etc.. On pourrait envisager d'en rendre compte à l'aide de prédicats *flous*. C'est une souplesse que nous permet le recours à des quantités continues.

Lorsque dans un texte on parle d'une position *à droite* ou *à gauche*, il faut donc chercher *le repère le plus plausible*, et interpréter cette déclaration de position comme relative à ce repère.

Considérons ainsi un véhicule B sur une route. nous avons trois repères "naturels" (figure 7):

- le *repère propre* du véhicule: (P_B, \bar{o}_B) : c'est le repère de référence dans: "je roulais devant/derrière le véhicule B", "un véhicule arrive sur ma gauche" (si le véhicule B est celui du locuteur) etc.
- le *repère lié à la route*: $(C(s_B), \bar{T}(s_B))$: c'est une manière d'interpréter "le véhicule B roule à droite": à droite par rapport au repère lié à la route et porté par la médiane.
- un repère lié au *mouvement* du véhicule: (P_B, \bar{u}_B) , où \bar{u}_B est le vecteur unitaire associé à la vitesse \bar{V}_B : $\bar{u}_B = \bar{V}_B / |\bar{V}_B|$.

En général, repère propre et repère lié au mouvement sont confondus, mais non si le véhicule chasse ou s'il roule en marche arrière (dans ce dernier cas ils sont opposés). La distinction permet par exemple de rendre compte de l'ambiguïté de l'énoncé: *J'effectuais une marche arrière quand un camion arrivait en face...*

À titre d'illustration, nous pouvons définir des *conditions normales de circulation*:

- (1) le véhicule ne chasse pas: $\bar{u}_B = \bar{o}_B$
 - (2) le déplacement se fait parallèlement à la direction de la route: la composante latérale $\bar{v}_{B,trans}$ du vecteur vitesse dans le repère lié à la route est petite
 - (3) on roule à droite: $y_B < 0$ (si $\bar{T}(s_B)$ est orienté dans le sens du parcours de la route par B).
 - (4) on reste sur la route: $|y_B| < l_d(s) - l_0$, pour un certain l_0 correspondant à la largeur du véhicule.
- Ces conditions peuvent être totalement formalisées dans notre langage.

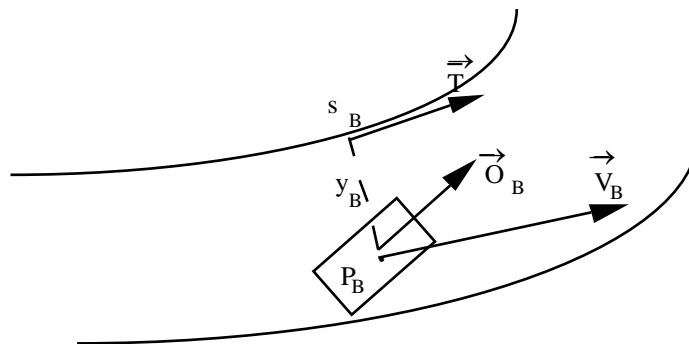


Figure 7: Repères "naturels" associés à un véhicule sur une route

4. Du langage au modèle

Nous abordons maintenant la question cruciale de l'interface entre les deux tâches: analyse (sémantique) linguistique et construction de la scène dans le cadre du modèle. La notion centrale est celle de *procès*. Associé en général à chaque proposition syntaxique du texte, un *procès* spécifie un événement (au sens large: action, état...) qui implique un ensemble d'objets: c'est en général la description d'un ensemble de relations entre ces objets, valide dans des conditions temporelles données. Insistons sur le fait qu'il ne s'agit que d'une spécification partielle, établie à partir de l'analyse des formes linguistiques et du lexique. D'autre part, ces objets et relations relèvent du modèle; ils sont représentés et formulés dans les termes prévus par celui-ci. L'ensemble des *procès* construits à partir d'un texte constitue la forme concrète des "contraintes" qui, intégrées avec les connaissances "du modèle" et les conditions d'énonciation, vont permettre de construire la signification des phrases et des textes - en l'occurrence la scène de l'accident. C'est cette notion que nous allons présenter, après avoir indiqué comment nous entendons représenter la sémantique lexicale.

4.1 Du lexique aux *procès*

Le lexique joue un rôle capital dans tout système de compréhension automatique de textes. C'est en effet à partir de l'information lexicale que l'on doit construire les entités et les événements dont il est question dans le texte, et que l'on doit déduire des constructions syntaxiques entre unités lexicales, les relations sémantiques entre les entités et événements correspondants. Le formalisme de représentation lexicale que nous avons choisi est basé sur la notion de *schéma*. Un schéma est associé à une unité lexicale et il est constitué d'une liste d'indications pour l'analyseur linguistique. Cette liste est composée de trois parties : une partie 'entités', qui donne la liste des entités dont la construction (création ou identification d'une entité déjà créée) est imposée par l'unité lexicale elle-même ; une partie 'relations' qui décrit les propriétés et relations entre ces entités qui sont impliquées par l'unité lexicale ; enfin une partie 'instructions' où l'on trouve les indications sur les relations qu'elle entretient avec les autres unités lexicales qui lui sont syntaxiquement liées. Les types des entités, les propriétés et les relations sont exprimés en termes du modèle du monde de la route : le lexique est ainsi l'élément de base qui va permettre les interactions entre les deux sous-systèmes.

Nous allons donner quelques exemples concrets de ces schémas pour des verbes de mouvement, sur lesquels nous avons particulièrement travaillé en raison de leur importance dans ces textes. Mais auparavant, il nous faut discuter brièvement du problème de la polysémie de certains de ces verbes, comme *rouler*, *déporter*, *doubler*, *dépasser*... Soit par exemple *rouler*. Le *Petit Robert* donne entre autres sens: "déplacer (un corps arrondi) en le faisant tourner sur lui-même" pour la construction transitive, "avancer au moyen de roues, de roulettes", "avancer, voyager dans un véhicule à roues" pour l'intransitif... sans parler des "duper" et autres expressions telles "rouler les hanches", "les mécaniques", "les 'r' ", etc. Deux attitudes sont possibles :

1) On cherche un sens "premier" ou "plus abstrait" — un "archétype cognitif" selon J-P Desclés, ou un "noyau" de sens selon d'autres auteurs — à partir duquel les autres sens pourraient se dériver en fonction de spécificités contextuelles. Cette position présente plusieurs difficultés. L'existence d'un sens "premier" n'est à l'heure actuelle qu'une hypothèse parmi d'autres. D'autre part en admettant cette hypothèse, il reste à trouver des mécanismes généraux de dérivation, et pour une exploitation en traitement automatique, à les formaliser. Au bout du compte, pour comprendre "je roule sur la chaussée" à partir (par exemple) de la rotation des roues du véhicule, la... route risque d'être longue!

2) On décrit plusieurs schémas de sens, qui devront être sélectionnés en fonction du contexte ou de constructions grammaticales. Cette position peut relever d'options théoriques sans doute tout aussi sujettes à débat. Elle peut être aussi dictée par des raisons purement pragmatiques : disposer de représentations partielles mais extensibles à volonté, plus directement exploitables en traitement automatique — sans préjuger du fonctionnement cognitif ou linguistique plus profond et sans anticiper sur une compréhension plus sûre du phénomène.

C'est cette seconde option que nous avons choisie. Voici, sur l'exemple du verbe *rouler*, le format proposé.

rouler

construction 1: voix active; transitif; sujet GN1:humain; complément direct GN2: objet physique

—> *schéma 1* % correspondant à "déplacer quelque chose en le
% faisant tourner sur lui-même"

construction 2: voix active; intransitif; sujet GN1: humain; complément prépositionnel GP1 facultatif: lieu

—>*schéma 2* % "se déplacer à l'aide d'un véhicule à roues"

—>*schéma 3* % "se déplacer en tournant sur soi-même"

construction 3: voix active; intransitif; sujet GN1: véhicule; complément prépositionnel GP1 facultatif: lieu

—>*schéma 4* % "se déplacer, en parlant d'un véhicule à roues"

(...)

Ainsi, pour une phrase comme *Je roulais dans une rue étroite*, l'analyseur linguistique peut sélectionner directement, à partir de la construction syntaxique, les deux sens linguistiquement possibles qui sont représentés ici par les schémas 2 et 3. Reste à effectuer le choix entre ces deux sens, ce qui, comme nous en avons discuté au §2.1, nécessite une coopération avec le système de modélisation, sous une forme ou sous une autre.

Revenons maintenant au formalisme de représentation des schémas. Voici à titre d'illustration le "schéma 2" pour *rouler* :

rouler

schéma 2

entités: V: ty_véhicule

relations:

à_roues(V)

sur IP: V.vitesse ≠ 0

instructions

contrainte_identification (GN1, V.conducteur,IP)

contrainte_spatiale(GP1, V.position, IP)

Muni de ces indications, l'analyseur linguistique peut construire les différents éléments du procès qui dépendent directement de ce sens de *rouler* : existence d'un véhicule (à roues), et prise en compte du fait que le véhicule est en mouvement sur tout l'intervalle IP, associé au procès (cf. la définition de l'intervalle IP au §2.1). En outre, l'instruction "contrainte_identification(...)", destinée au module "Identification des entités" (cf. §2.2.3), impose une contrainte sur le second argument (V.conducteur) en fonction de l'analyse du sujet GN1: en principe une identification avec le référent calculé pour GN1. Cette contrainte est relativisée à l'intervalle IP. De même, la seconde instruction est destinée au module "Analyse des relations spatio-temporelles" pour l'analyse du complément prépositionnel de lieu GP1 (s'il est présent) et produit une contrainte sur V.position.

Ainsi pour la phrase *Je roulais à droite sur la chaussée...* (texte 8 modifié, voir § 4.2), S₀ désignant le locuteur, et R₁ la route référent de *la chaussée* nous aurons les contraintes:

sur IP : V.conducteur = S₀ & V.position ∈ R₁

Donnons encore, comme second exemple, le codage du schéma d'interprétation pour *déporter*, que l'on trouve aussi dans le texte n°8. Le sens correspondant à la construction présente dans ce texte (voix passive avec sujet de type 'objet mobile') est, selon le *Petit Robert*: "être dévié de sa direction, entraîné hors de sa route".

déporter

construction 1: voix passive; sujet GN1: objet_mobile

—> *schéma 1* % correspondant à "être dévié de sa direction"

entités: X: ty_objet_mobile, \bar{D} : ty_direction_mobile

relations:

sur IP : dir_normale(X, \bar{D})

& |composante(X.vitesse, trans, \bar{D})| > k |X.vitesse|

instructions:

contrainte_identification(GN1,X,IP)

où k est un paramètre à fixer, et "dir_normale" désigne la direction dans laquelle devrait "normalement" s'effectuer le déplacement (cf. §4.3).

4.2 Les procès

Un exemple, accompagné de quelques commentaires (à propos d'expressions linguistiques dont nous n'avons pas eu l'occasion de discuter dans ce qui précède), devrait suffire à présenter le format que nous proposons pour le codage des procès. Il s'agira de la première phrase du texte n° 8, légèrement modifiée en:

Je roulais à droite sur la chaussée quand un véhicule arrivant en face dans le virage a été déporté.

de manière à éviter le traitement des locutions *la partie droite de* et *complètement*, qui alourdirait quelque peu la présentation.

Un ensemble d'intervalles temporels sont systématiquement introduits, et ne figureront pas dans la rubrique "objets" des procès. Il s'agit (cf. §2.2.4) de l'intervalle d'énonciation (IE₀) et des

intervalles "liés aux procès" (IP_i) et "d'assertion" (IA_i). Il en est de même pour l'énonciateur, noté S₀ qui est de type ty_personne. Les relations d'adjacence, d'antériorité et d'inclusion entre intervalles (notée respectivement |, < et ⊂) sont définissables par des relations d'ordre entre leurs bornes.

Procès n°1

TEXTE: "Je roulais à droite sur la chaussée"

OBJETS: V₁: ty_véhicule, R₁: ty_route, Rep₁: ty_repère

RELATIONS ASPECTUO-TEMPORELLES:

IA₁ < IE₀ & IP₁ ⊂ IA₁

RELATIONS EVENEMENTIELLES:

sur IP₁: V₁.conducteur = S₀ & V₁.vitesse ≠ 0 &
V₁.position ∈ R₁ & à_droite(V₁.position, Rep₁)

Commentaire:

L'analyse linguistique stricte laisse inspecifié le référentiel Rep₁ auquel réfère l'expression "à droite", et qui doit donc être actualisé en fonction d'autres informations, comme on le verra au §4.3.

Procès n°2

TEXTE: "quand un véhicule (...) a été déporté"

OBJETS: V₂: ty_véhicule, D̄₂: ty_direction_mobile

RELATIONS ASPECTUO-TEMPORELLES:

IA₁ | IA₂ & IA₂ = IP₂ & IA₂ < IE₀

RELATIONS EVENEMENTIELLES:

sur IP₂: V₂.vitesse ≠ 0 & dir_normale(V₂, D̄₂)
& |composante(V₂.vitesse, trans, D̄₂)| > k |V₂.vitesse|

Commentaire:

On remarquera que si lexicalement *être déporté* nécessitait (dans cette acception) une entité de type 'objet mobile', la valeur lexicale du sujet (*un véhicule*) a permis de spécifier davantage le type de cette entité.

Procès n°3

TEXTE: "(un véhicule) arrivant en face dans le virage"

OBJETS: SR₂: ty_segment_route, Rep₂: ty_repère

RELATIONS ASPECTUO-TEMPORELLES:

IP₃ ⊂ IA₃ & IA₃ = IA₂

RELATIONS EVENEMENTIELLES:

sur IP₃: V₂.vitesse ≠ 0
& lié(Rep₂, S₀)
& devant(V₂.position, Rep₂)
& devant(S₀.position, (V₂.position, V₂.vitesse))
& courbe(SR₂)
& R₁ ⊃ SR₂

Commentaires:

- Nous introduisons un type `ty_segment_route` dont les éléments sont de la forme: (R:route , [s₁,s₂]) où [s₁,s₂] est un sous-intervalle de l'intervalle de définition du ruban R.
- Rep₂ est le repère requis pour interprétation de "en face", parmi les repères liés à l'énonciateur. De nouveau un calcul de détermination est nécessaire pour le préciser (cf. §4.3).
- "V arrive" est ici interprété comme: "je suis en position 'devant' par rapport au repère lié au mouvement de V."
- La dernière relation suppose que l'analyseur linguistique ait pu résoudre l'anaphore associative reliant la route "support" de SR₂ et la route R₁ introduite par le premier procès.

4.3 Des procès à la scène

L'exemple que l'on vient de voir montre bien en quoi les sorties de l'analyseur linguistique ne sont que des spécifications partielles de la scène, qu'il s'agisse d'identification d'entités, de relations temporelles, spatiales, etc. Le sous-système "Construction de la scène" va utiliser ces sorties, en les combinant avec d'autres informations, pour remplir un objectif précis : représenter une scène d'accident de la circulation dont l'énonciateur (S₀) est l'un des protagonistes. Pour cela, le mécanisme de raisonnement s'appuie sur des connaissances a priori liées à la nature de la tâche, qui permettent de compléter le texte et de construire l'enchaînement des événements de manière cohérente.

D'abord, ce système peut compléter la description grâce à des valeurs par défaut. Supposons par exemple que l'on traite la phrase *Je roulais à faible allure*, dans laquelle le verbe *rouler* est utilisé dans la construction 2 (cf. §4.1), mais sans groupe prépositionnel de lieu. Du point de vue strictement linguistique, il n'y a aucune raison de supposer que ce lieu est une route (on peut rouler dans le désert, sur la plage, sur une place, ...). Par contre, la connaissance du contexte de ces témoignages permet d'imposer des règles par défaut qui consistent à affirmer qu'en l'absence de précisions explicites, les acteurs de la scène se déplacent sur des routes dans des conditions normales de circulation (à droite de la chaussée, etc.). On voit là un avantage de notre approche : Il serait difficile, à partir d'une description lexicale par des primitives conceptuelles générales, de traiter correctement ce problème des valeurs par défaut. Il faudrait en effet, soit "encombrer" l'analyse linguistique d'une foule de règles par défaut censées couvrir tous les contextes imaginables, soit se contenter d'un analyseur ad hoc, dédié à ces témoignages. Ici, on assure un certain degré de généralité de la partie linguistique du système, tout en se donnant des moyens efficaces de remplir l'objectif de compréhension fixé, en faisant intervenir ces règles par défaut au niveau extra-linguistique où elles sont justifiées.

De même, il revient au module de construction de la scène d'inscrire sur un axe temporel unique les événements et autres relations entre entités. Les sorties de l'analyseur donnent, comme on l'a vu, un certain nombre de relations temporelles entre des intervalles de nature diverse. Un raisonnement temporel doit permettre d'en déduire comment les événements relatés s'agencent, en utilisant à la fois ces relations et des connaissances encyclopédiques. Par

exemple, s'il est asserté qu'un véhicule était en mouvement pendant un certain intervalle IA, strictement inclus dans l'intervalle IP correspondant, et qu'aucun élément ne permet de déduire que ce mouvement s'est interrompu (collision, feu rouge, etc.), le système peut asserter que le véhicule a continué son chemin après l'intervalle IA, alors que cela n'est pas explicitement mentionné. D'une manière générale, ce raisonnement temporel s'effectue en référence à un cadre spatio-temporel global, dans lequel toute entité est repérée spatialement à tout instant de l'intervalle total pendant lequel se déroule la scène.

Ce repérage spatial nécessite aussi que le système résolve les indéterminations produites par l'analyse linguistique sur les différents repères décrits dans les procès. Ainsi, pour la phrase *Je roulais à droite sur la chaussée* (§4.2 procès n°1), on a vu que l'analyseur laisse complètement inspecifié le repère par rapport auquel doit être comprise l'expression *à droite*. En effet, rien dans la forme linguistique ne permet de choisir entre les trois repères principaux mentionnés au §3.2.4 pour un véhicule en mouvement. C'est donc le système de construction de la scène qui devra effectuer ce choix. En l'occurrence, des connaissances encyclopédiques en rapport avec l'expression "rouler à droite" pourront être utilisées, ou un raisonnement reposant sur le fait qu'un véhicule ne peut pas être à droite de lui-même. Rep₁ sera ainsi actualisé en:

$$\text{Rep}_1 = \text{repère_route}(V_1)$$

(qui spécifie le repère lié à la route sur laquelle circule V_1 , ayant pour origine le point de la médiane à la hauteur de V_1 .position, et pour direction la direction de même sens que V_1 .orientation — cf figure 7, § 3.2.4).

Bien entendu, si l'on avait l'expression *à droite du camion*, l'analyseur linguistique aurait pu formuler une contrainte supplémentaire sur ce repère, en indiquant qu'il devait être en rapport avec l'entité représentant le camion : il faut noter que cela ne suffit pas pour autant à spécifier complètement le repère, puisqu'un objet peut être à droite d'un autre par rapport à l'orientation de ce deuxième objet ou bien par rapport à l'observateur. Le même phénomène de spécification partielle intervient pour l'expression *en face* (§4.2 procès n°3), où l'on sait grâce à l'analyse linguistique que le repère est lié à l'énonciateur, sans pour autant pouvoir choisir entre les différents repères possibles (orientation de l'énonciateur ou direction de son mouvement). Un dernier exemple de travail à effectuer par le système de construction de la scène dans le domaine spatio-temporel est fourni par la notion de "direction normale" (§4.2 procès n°2). Ici, il s'agit de déduire des conditions normales de circulation (cf §3.2.4) que cette direction est la direction de la route sur laquelle circule V_2 (dans le sens du mouvement de V_2), et un autre raisonnement, relativement complexe, permettra d'identifier cette route avec R_1 .

Notons enfin que l'interprétation purement spatiale et cinématique ("physicaliste") à laquelle nous nous limitons ici n'est pas totalement satisfaisante. En particulier les verbes de mouvement qui apparaissent dans les textes font référence, à des degrés divers mais la plupart très fortement, à des *actions de conduite*, répondant à des buts bien précis, intégrées dans des plans, ou nécessitant elles-mêmes un plan complexe: doubler un véhicule, traverser un carrefour, négocier un virage. Notre modèle devra donc être complété dans ce sens, de manière à intégrer une *théorie de l'action*. Nous pensons dans un premier temps nous limiter à quelques idées simples, et caractériser une action A par une *précondition*, un *but* et une information sur la

transition ou le scénario standard. Il conviendra également d'établir le lien avec la notion de procès dans la théorie de la temporalité de référence et les divers intervalles temporels qu'elle stipule. Nous devons ensuite sans doute tenir compte des intentions de l'énonciateur quand il écrit son témoignage (se disculper vis-à-vis de sa compagnie d'assurances) pour pouvoir mieux interpréter la manière dont il présente ses actions de conduite.

Il faut souligner que ces extensions sont facilitées par l'architecture globale du système, puisque d'une part chacun des schémas d'interprétation est formulé dans le cadre d'un "modèle" du monde (ou plus précisément d'ailleurs, de différentes "régions" de ce modèle: ce n'est pas le même corps de concepts qui est à l'oeuvre pour interpréter *rouler sur une route* ou *rouler* dans le sens de *duper*), et que d'autre part l'organisation modulaire de l'analyseur linguistique permet aussi d'ajouter les composantes d'analyse sémantique que nous avons laissé de côté pour l'instant (comme les modalités, qui jouent un rôle important dans l'expression des actions de conduite) sans remettre en cause la structure du système.

5. Conclusion

Ainsi l'architecture que nous venons de présenter s'appuie à la fois sur une conception théorique de l'activité cognitive de compréhension de textes et sur des impératifs de "faisabilité" et d'efficacité dans le traitement automatique. Plutôt que de rechercher un petit nombre de primitives "universelles" à partir desquelles on pourrait exprimer intégralement le sens d'un texte, nous plaidons pour des univers conceptuels riches et variables d'un domaine à l'autre, dans le cadre desquels on peut interpréter les données issues d'une analyse sémantique des énoncés, en considérant que ces données sont intrinsèquement insuffisantes pour permettre à elles seules une compréhension.

Une objection pourrait porter sur le manque de plausibilité cognitive de cette modélisation en termes de "physique scientifique" —par opposition à une "physique naïve". Notre réponse est double. Premièrement, il est question ici de *traitement automatique* du langage et nous ne sommes pas tenus, en principe, à cette forme de plausibilité. D'autres critères, relatifs à l'ingénierie informatique, entrent également en ligne de compte. En second lieu, la critique devrait être soigneusement argumentée: par exemple le principe de repérage présenté dans ce texte —relatif à une *direction* et un *point origine*— ou encore la représentation du sens lexical de *être déporté* —grâce à la représentation d'une direction par un vecteur— nous paraissent assez satisfaisants au sens où ils ne font que formaliser l'intuition que l'on peut avoir de ces notions.

Enfin, une confusion doit être écartée. Notre modèle fait référence aux nombres réels, à des systèmes de coordonnées, à des fonctions à valeur réelle, etc.; pour autant, on ne doit pas le qualifier de "numérique". Le lecteur a pu observer que nous n'avons pas écrit une seule équation de courbe! Notre représentation est parfaitement symbolique (comme la plupart du temps en mathématiques). Les procédures de raisonnement mises en oeuvre seront essentiellement des inférences logiques —même si nous pensons qu'une simulation peut le cas échéant être un procédé pertinent et efficace.

Références

- HERZOG, O. et ROLLINGER C.-R. (1991): *Text Understanding in LILOG*, Lecture Notes in Artificial Intelligence n° 546, Springer Verlag.
- BRIFFAULT, Xavier (1992): Modélisation informatique de l'expression de la localisation en langage naturel, Thèse d'Université, Paris 6.
- GOSSELIN, L. (1992): Une grammaire calculatoire du temps et de l'aspect en français, Rapport interne de l'ELSAP, Université de Caen.
- GOSSELIN, L. (1993): Subjectivité et temporalité : une approche cognitive, *Elseneur*, 9, Presses Universitaires de Caen.
- KLEIN, W. (à paraître): *Time in Language*, Cambridge University Press.
- CROUCH R.S. et PULMAN S.G. (1993): Time and modality in a natural language interface to a planning system, *Artificial Intelligence*, 63, 265-304.
- VERGNE J. (1994): A non-recursive sentence segmentation, applied to parsing of linear complexity in time, *NeMLaP 94, International Conference on New Methods in Language Processing*, Manchester.
- ARIEL M. (1990) *Accessing noun phrases antecedents*, Routledge, London.