

SIGNAUX FAIBLES ET CHAINONS MANQUANTS DANS LES SYSTEMES COMPLEXES

Une application du Calcul Littéraire au Renseignement Militaire

Jean Rohmer

Thales Communications Direction Technique

Centre d'Excellence en Nouvelles Technologies d'Analyse de l'Information (CENTAI)

jean.rohmer@fr.thalesgroup.com

Résumé

Cet article examine les relations entre les systèmes complexes et leur modélisation sous forme de réseaux sémantiques. Nous considérons ensuite le Renseignement Militaire comme une activité de modélisation d'un système complexe. Nous introduisons le concept de « *calcul littéraire* » comme ensemble d'opérations sur les réseaux sémantiques. Un exemple de calcul littéraire consiste à aider à déterminer si une nouvelle information est ou non *un signal faible, un chaînon manquant*, afin de décider si elle doit ou non être l'objet d'un traitement intensif de la part d'une cellule de Renseignement.

Abstract

In this paper, we first examine the relationship between complex systems and their modelling under the form of semantic networks. We then consider Military Intelligence activity as modelling a complex system. We introduce the concept of « *litteratus calculus* », to perform useful operations on semantic networks. An example of litteratus calculus helps to determine whether a new incoming information is or not a *weak signal, a missing link*, and if it deserves more processing from an Intelligence Cell.

Réseaux Sémantiques, Signaux Faibles et Systèmes Complexes

Les Réseaux Sémantiques sont une notion très ancienne (certains la font remonter à Aristote), qui, plus récemment, a été beaucoup étudiée et utilisée en Intelligence Artificielle, comme moyen de représentation des connaissances.

Un réseau sémantique est tout simplement un graphe dont les noeuds sont des intitulés en langage naturel d'objets, de concepts, et dont les arcs expriment des relations entre ces objets. Le tout peut aussi être vu comme un ensemble de phrases élémentaires qui partagent des termes :

Ignace de Loyola / est le fondateur de / La Société de Jésus

La Société de Jésus / est aussi appelée / Les Jésuites

D'Alembert / a violemment attaqué / Les Jésuites

Les Jésuites / ont été protégés par / Catherine II de Russie

Catherine II de Russie / a protégé / Voltaire

D'Alembert / est un acteur de / Siècle des Lumières

Voltaire / est un acteur de / Siècle des Lumières

Voltaire / a fait ses études à / Collège de Clermont

Collège de Clermont / a comme enseignants / Les Jésuites

Comme on le voit, les réseaux sémantiques sont un outil assez simple pour décrire, représenter, modéliser des phénomènes complexes ...

Si un système complexe est caractérisé par le grand nombre d'interactions entre ses composants, on peut penser qu'une représentation d'un système complexe en réseaux sémantiques va comporter de nombreux nœuds et de nombreuses arêtes.

On peut aussi espérer que l'étude du réseau sémantique correspondant peut aider à étudier de manière intéressante certains aspects du système complexe qu'il décrit.

Si nous considérons un réseau sémantique comme un graphe, nous pouvons nous intéresser à ses chemins, aux manières d'aller d'un sommet à un autre, à l'ensemble de toutes les relations entre deux sommets, etc ...

Plus un système est complexe, plus les cheminements possibles sont variés.

Dans le domaine du Renseignement Militaire, il est intéressant de considérer l'ensemble de la situation qu'une Cellule de Renseignement a à surveiller comme un système complexe.

Plusieurs expériences récentes se sont attachées à représenter l'ensemble des connaissances d'un théâtre d'opérations -au sens large : il peut s'agir d'une sphère spatio-temporelle objet de menées terroristes- sous forme d'un réseau sémantique.

A la STAT (Section Technique de l'Armée de Terre de l'Etat Major de l'Armée de Terre), le Lieutenant Colonel De Nicola a ainsi mené des expérimentations opérationnelles de grande ampleur, utilisant le logiciel IDELIANCE, de la société Thales.

Un des problèmes clés d'une activité de Renseignement est de savoir si une nouvelle information entrante est ou non importante. Cette activité d'analyse et de prise de décision est délicate, dans la mesure où la nouvelle information peut apparaître anodine tout en étant symptomatique de phénomènes importants. On appelle couramment *signaux faibles* de telles informations.

Dans cet article, nous souhaitons montrer comment l'utilisation de réseaux sémantiques pour représenter la complexité d'une situation militaire peut aider à mieux détecter ces fameux *signaux faibles*.

Le principal obstacle à la détection des signaux faibles est précisément la complexité du réseau sémantique dans lequel ils viennent prendre place. Mais la célèbre citation de Hölderlin va venir à notre secours :

«Là où est le danger, là aussi croît ce qui sauve »

Plus un système est complexe, plus il y a d'interactions, de chemins possibles entre ses composants. Mais cette multiplicité de chemins pourra aussi être mise à profit pour faciliter la circulation des facteurs de compréhension et d'intervention sur le système.

Le projet « Smallworld » de l'Université de Columbia a récemment vérifié expérimentalement l'affirmation selon laquelle deux individus pris au hasard dans le monde sont en moyenne à moins de 6 liens interpersonnels les uns des autres. Dit autrement, cela met en évidence qu'au sein d'un système complexe, il existe des *courts-circuits locaux* qui en facilitent l'appréhension. Plus un système est complexe, plus il facilite l'existence de sous-systèmes élémentaires qui aident à l'appréhender.

Nous allons dans ce qui suit appliquer concrètement cette potentialité à la détection des signaux faibles, en procédant à des opérations sur les graphes des réseaux sémantiques. D'une manière générale, nous appelons « *calcul littéraire* » de telles opérations.

Signaux Faibles ou Chaînes Manquants ? : une introduction informelle

On parle souvent et abondamment de "signaux faibles". Parfois l'usage de cette expression est censé suggérer les capacités un peu inquiétantes du locuteur à voir, par la grâce de pouvoirs ou de réseaux occultes, ce que le commun des mortels ne verra jamais, parfois l'expression stigmatise avec ironie l'incapacité de telle ou telle institution -très puissante de préférence- à percevoir ces ondes mystérieuses, aveuglement qui la conduira immanquablement à sa perte. Souvent l'évocation des "signaux faibles" est associée à l'allusion à un programme informatique réel ou fantasmé qui permettrait de détecter ce troublant *rayonnement sémantique*.

Si on se risque rarement à définir les signaux faibles, on en donne par contre fréquemment des exemples.

Jean Maurice Bruneau, chercheur à l'Institut National des Télécommunications, rappelle ainsi que la Maison Blanche avait un jour commandé une quantité beaucoup plus importante qu'à l'accoutumée de cigares de La Havane, signe que son hôte allait déclencher une action hostile au régime de Fidel Castro.

Selon une autre anecdote entendue lors d'un colloque sur l'analyse de données, une célèbre terroriste française aurait été "logée" par une observation fine sur l'ensemble du territoire des variations des ventes à des particuliers d'aliments pour hamsters, les autorités connaissant son attachement à la compagnie de ces animaux.

Regardons les signaux faibles de plus près

Ces deux exemples, différents à plusieurs titres dans leur nature, ont néanmoins un point commun, et commun selon nous à toutes les circonstances où l'on peut parler de signal faible. Ce point commun est en général ignoré tellement il tombe sous le sens, mais c'est pourtant lui qui va nous conduire à une proposition pour «détecter les signaux faibles». Trois éléments fondent l'analogie entre les deux cas:

1) Dans les deux cas, nous sommes en présence d'une problématique claire:

- les USA vont-ils lance une attaque de Cuba ?
- on cherche la localisation de la terroriste

2) L'énoncé de chacune de ces problématiques comporte au moins deux termes:

- attaque de Cuba / USA
- localisation / terroriste

3) Dans chaque cas on sait tenir deux discours qui relie le signal faible aux deux termes de la problématique

La Maison Blanche passe une grosse commande de cigares> attaque de Cuba

La Maison Blanche passe une grosse commande de cigares> USA

Chiffre des Vente d'aliments pour hamster> localisation

Chiffre des Vente d'aliments pour hamster> terroriste

La première réaction à ces observations est souvent "mais tout cela est évident!" ... c'est pourquoi, plutôt que de détailler les quatre discours juste esquissés, nous invitons le lecteur à les établir lui même avec précision, en s'aidant d'un crayon et de papier.

Une tentative pour apporter un peu de rationalité à l'analyse des signaux faibles

Ce qui nous intéresse c'est précisément que "ça a l'air évident". Si c'est si évident, ça doit être à la portée d'un programme informatique !

Notre démarche s'oppose à une autre démarche , tout aussi évidente mais plus répandue, qui consiste à dire "puisque c'est évident, il n'y a aucun besoin de faire appel à l'informatique".

Nous appelons notre démarche "*Calcul Littéraire*".

NB: L'arithmétique aussi, ça a l'air évident.

L'approche est la suivante :

Soit une *problématique P* énoncée par une phrase comportant au moins deux termes *suffisamment* significatifs, T1 et T2.

Soit une autre phrase SF dont on veut savoir si c'est un *signal faible* ou pas.

Une remarque de base est qu'un "signal faible" n'est pas "signal faible" dans l'absolu, mais seulement vis à vis d'une problématique donnée. En conséquence, si l'on veut construire un système automatique de détection des signaux faibles, il faut que ce système *possède une représentation des problématiques*.

On appelle *discours* une suite de phases dont chacune comporte *suffisamment* d'éléments communs avec les précédentes. Autrement dit, dont la suite de phases n'exprime pas des propos « décousus » (les phrases sont cousues entre elles, elles sont reliées par un fil conducteur : on retrouve naturellement la notion de toile , de réseau). *On ne passe pas du coq à l'âne entre deux phrases*.

Si on arrive à construire deux discours -où les P_i et Q_i sont des phrases- de la forme :

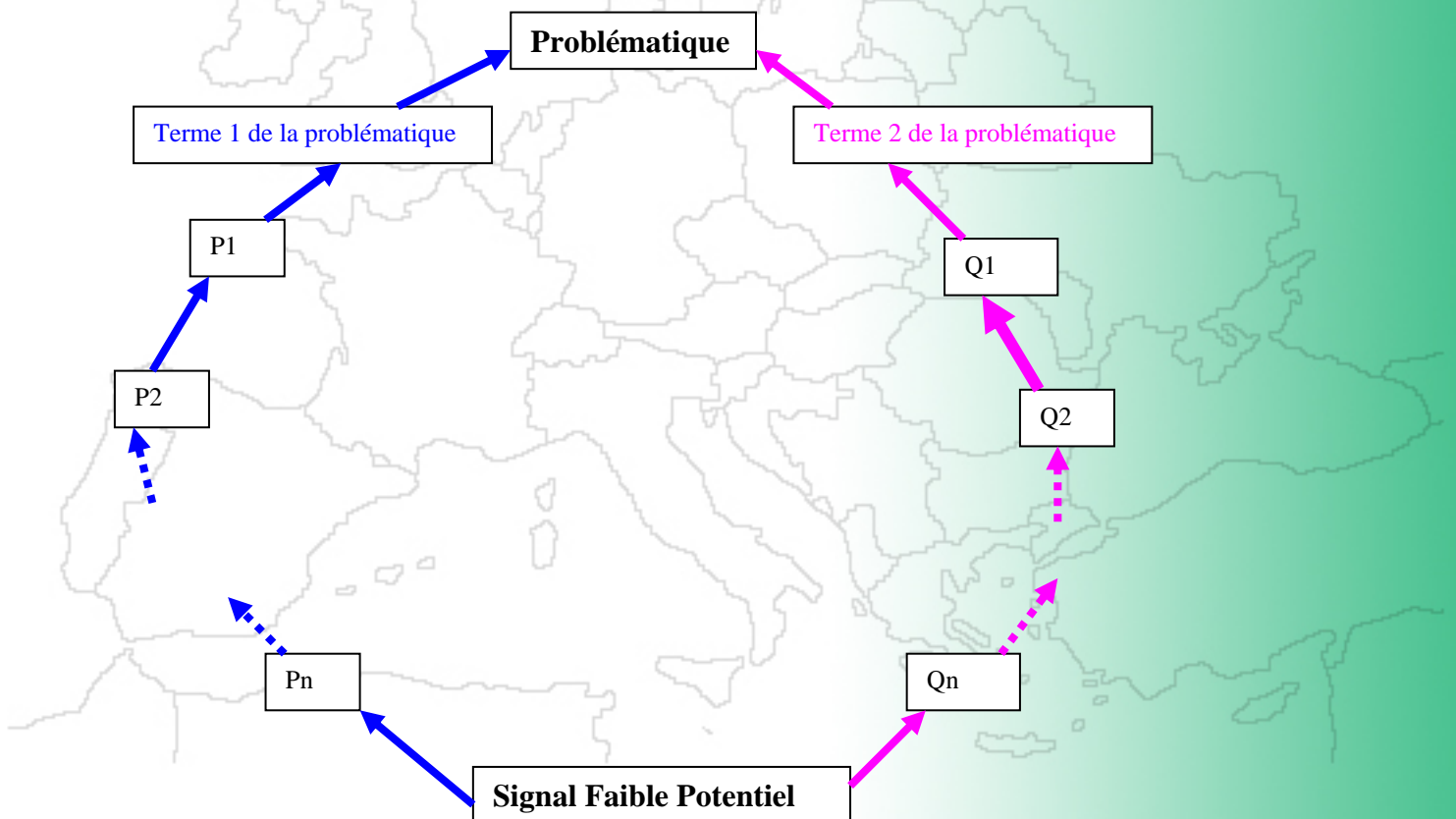
SF \Rightarrow $P_1 \Rightarrow P_2 \Rightarrow \dots \Rightarrow P_n \Rightarrow$ Terme 1 de la problématique

et

SF $\Rightarrow Q_1 \Rightarrow Q_2 \Rightarrow \dots \Rightarrow Q_n \Rightarrow$ Terme 2 de la problématique

alors SF est *peut-être* un signal faible relativement à la problématique.

La présence de ces deux discours est une condition *nécessaire* (et *non suffisante*) pour que SF soit un signal faible relativement à la problématique.



Un tel schéma met en évidence le fait que l'on est capable de tenir un discours qui relie deux termes de la problématique à notre candidat signal faible. Nous ne savons pas à ce stade si ce discours est pertinent, convaincant, mais ce que nous savons, c'est que, si nous ne trouvons

pas un tel circuit, un tel rebouclage entre la problématique et le signal, alors on est sûr que ce n'est pas un signal faible. Et inversement, si il y a un discours possible, alors nous en retrouverons le chemin. (la condition nécessaire et non suffisante).

Pour aller vers une automatisation, il faut donc procéder ainsi :

- Constituer une base de connaissances sur le domaine : c'est notre ensemble de phrases
- Représenter dans le système l'ensemble de nos problématiques (aussi sous forme de phrases)
- A l'arrivée d'une nouvelle information, rechercher, comme ci-dessus, si elle permet de boucler un ou plusieurs discours avec une ou plusieurs problématiques.
- Examiner ces discours et décider des conséquences éventuelles à en tirer

Important : la nouvelle information, qu'elle ait été ou non qualifiée de « signal faible », vient s'ajouter aux phrases de la base de connaissances.

On voit qu'une nouvelle information P1 peut ne pas être considérée comme signal faible à l'instant T, mais qu'elle pourra jouer un rôle clé à l'instant T+N lors de l'arrivée d'une information P2 qui, elle, sera en position de boucler une boucle.

Ceci met clairement en évidence que la notion de signal faible est très relative et circonstancielle : P2 ne boucle la boucle que parce que P1 était apparue précédemment et a été conservée.

*Ceci suggère d'utiliser l'expression « **chaînon manquant** » aussi bien que celle de « signal faible ».*

Autant ce dispositif permet de donner l'alerte quand P2 arrive, autant il n'alerte en rien quand P1 arrive. Il ne nous dispense pas de conserver les informations, il ne minimise pas le volume d'informations à conserver, mais il minimise le temps passé à exploiter les informations, en nous aiguillant uniquement vers des terrains potentiellement fertiles.

Un autre intérêt du système est d'indiquer quelle(s) problématiques l'information nouvelle concene éventuellement, ce qui peut d'une part aider à donner des priorités à leur examen, et d'autre part fournir des indicateurs subtils sur « ce qui chauffe en ce moment ».

Un exemple de calcul littéraire de chaînon manquant

Cet exemple a été réalisé à partir d'une variante du logiciel Idéliance.

La problématique est la suivante :

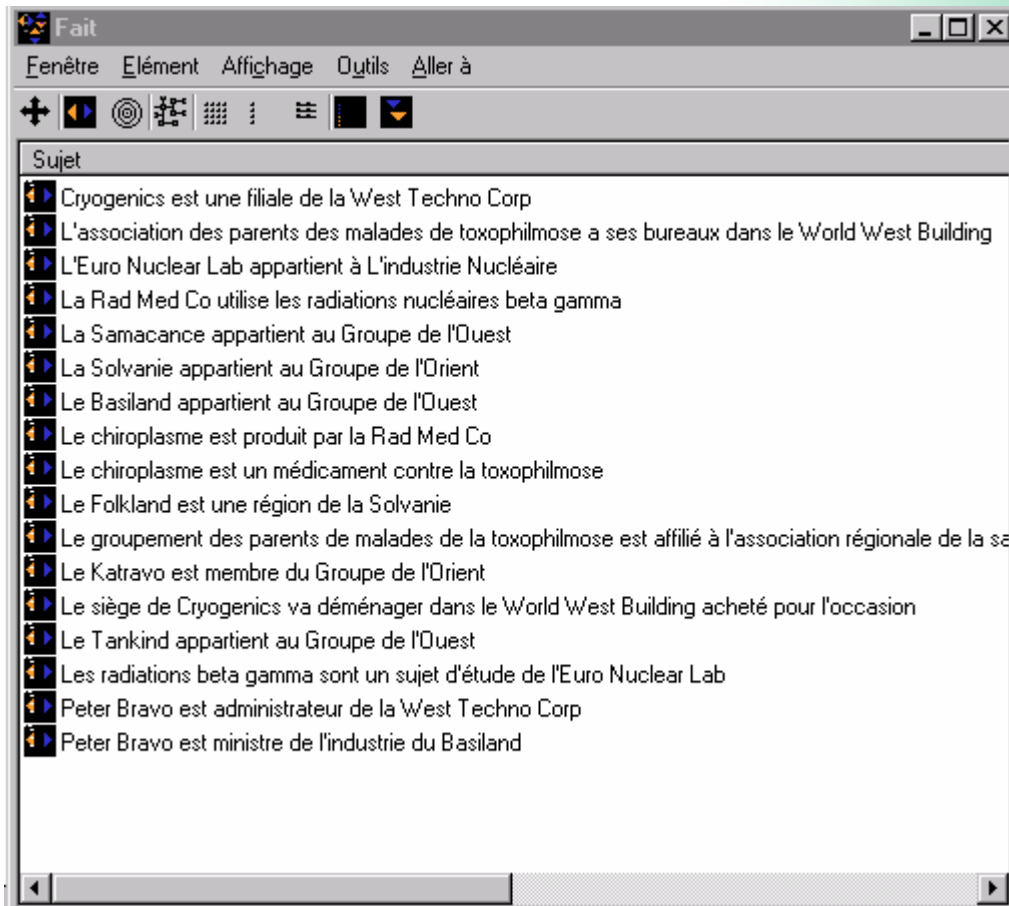
On s'interroge sur l'existence d'une connivence entre le Groupe d'Orient et le Groupe de l'Ouest (deux fédérations de pays) dans le domaine du Nucléaire.

6 ème Congrès Européen de Science des Systèmes

On part d'un ensemble de phrases :



19 - 22 septembre 2005



Et on veut savoir si la phrase :

L'association des parents des malades de toxoplasmose a ses bureaux dans le World West Building

est un signal faible relativement à notre problématique.

En suivant la procédure définie plus haut, on effectue *automatiquement* un « calcul littéraire », qui découvre effectivement trois discours qui relient cette phrase (en rouge) à trois termes (en bleu) de la problématique. Dans les graphes des discours, on a aussi indiqué les éléments de liaison entre deux phrases consécutives.

6 ème Congrès Européen de Science des Systèmes



6^{ème} Congrès Européen de Science des Systèmes



19 - 22 septembre 2005