MILEVA EN QUELQUES MOTS

A quoi sert Mileva ?

Mileva est une intelligence artificielle appartenant à la famille des intelligences artificielles cognitives. Comme tout système d'intelligence artificielle, Mileva travaille à partir d'un référentiel, modélisant l'univers dans lequel elle travaille, et de données qu'elle analyse.

Mileva est spécialisée dans la modélisation d'interactions entre acteurs hétérogènes, **résolvant ainsi le problème de l'analyse de situations complexes intégrées**, afin d'apporter des éléments de réponses à une question posée. Mileva permet au décideur, sur une question donnée, de connaître les forces en présence, **leurs interactions** et les signaux faibles existants. Elle a une approche globale grâce à des sources multiples fermées (les données apportées par le décideur) et l'accès à des données ouvertes comme Google. Mileva prend en compte de nombreux paramètres, détermine des proximités, des risques et ruptures.

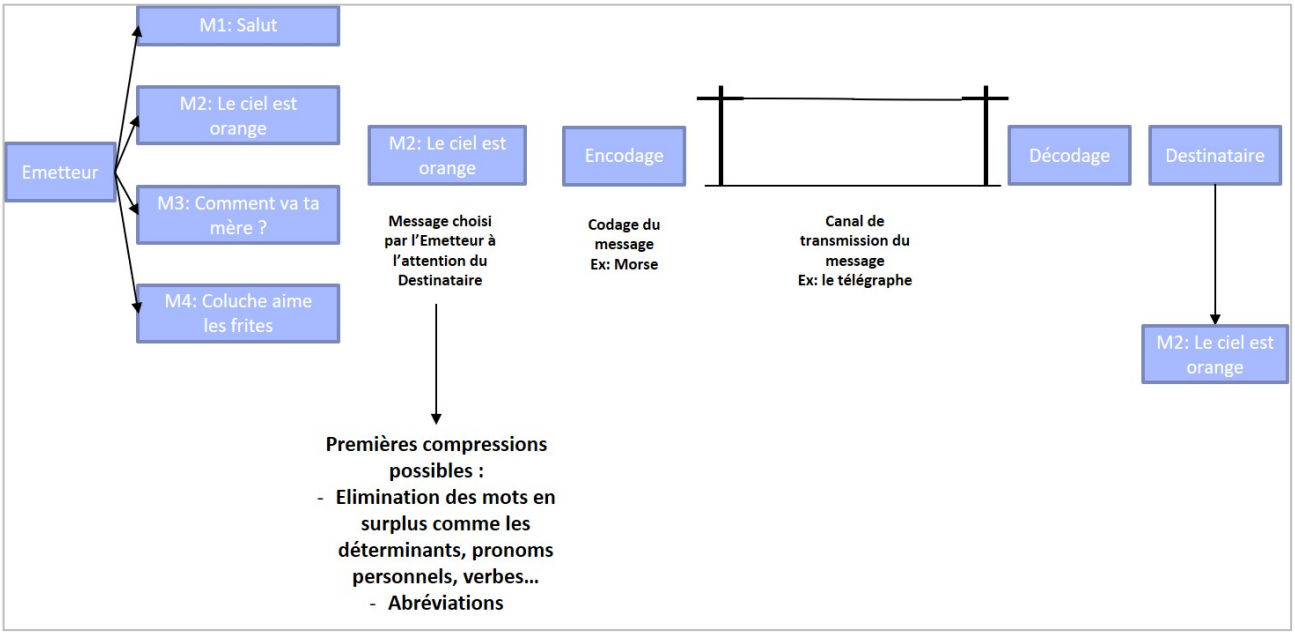
Le fonctionnement de Mileva

I La théorie de l'information

La théorie mathématique de l'information et de la communication est une théorie publiée en 1948 par Claude Shannon.

A l'origine, cette théorie s'applique à un système de communication constitué d'un émetteur, d'un canal de transmission et d'un récepteur.

Il s'agit de transmettre des informations de manière efficace, rapide et économique. Pour cela, le plus simple est de **mesurer la** longueur du message au maximum en ne conservant que les informations principales par évacuation des redondances que représentent certains déterminants (le, la, les, des…), des pronoms (je, tu, ils…) ou en utilisant des abréviations.



Système de communication de Claude Shannon

En s'intéressant à la structure de la langue, Claude Shannon observe que certaines lettres reviennent plus souvent que d'autres. Par exemple, dans la langue française, le « E » a une fréquence de 12,26% quand la fréquence du « W » est de 0,04%.

De la même manière, on constate des fréquences différentes de bigrammes, de trigrammes… Par exemple, l'enchaînement « ES » a une fréquence de 3,05 % quand l'enchaînement « DQ » a une fréquence de 0.01 %. Ainsi, les informations essentielles des textes sont extraites par le tri des lettres les plus courantes sans pour autant en perdre le sens.

Par exemple, la phrase « Le ciel est orange » devient « ciel orng ».

Les documents, ainsi compressés, contiennent uniquement les informations nécessaires pour en conserver le sens. Ces dernières peuvent être assimilées à un « niveau d'information » qui est mesuré par une formule mathématique produisant une valeur mathématique appelée « **entropie** ». Ainsi, l'entropie permet de connaître le niveau d'information d'un document, soit les informations essentielles d'un document.

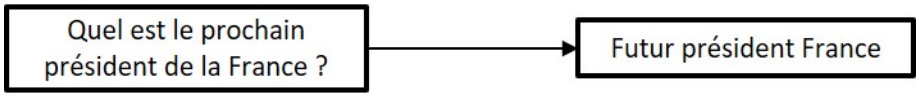
Par ailleurs, l'entropie permet de déterminer si les informations d'un document sont nouvelles par rapport aux informations préalablement collectées.

En effet, de la même manière que des lettres apparaissent plus ou moins souvent dans un texte, des informations reviennent plus ou moins souvent dans un périmètre de recherche défini. Par analogie, il est possible d'extraire les données essentielles d'un périmètre de recherche, de la même manière qu'on extrait les informations primordiales d'un document en en retirant les redondances (mots et caractères inutiles car ils reviennent fréquemment). On notera que le cas particulier de la

maximisation de l'entropie revient à calculer des probabilités bayésiennes. La théorie de l'information n'est pas très éloignée du monde des probabilités.

II La question

Mileva répond à une question en langage naturel (Ex : Quel sera le prochain Président de la République française ?) mais posée sous une forme simplifiée (Ex : Futur Président France ?).



Mileva ne répondra de manière précise à une question, que si cette dernière correspond à la problématique fondamentale ciblée. Il faut donc déterminer quel est le problème adressé afin de poser correctement la question. Le choix des mots est essentiel : deux questions similaires, presque identiques, renverront des résultats aux nuances différentes. Pour répondre à la question posée, Mileva a besoin d'un périmètre de recherche dans lequel elle réalise l'analyse des données.

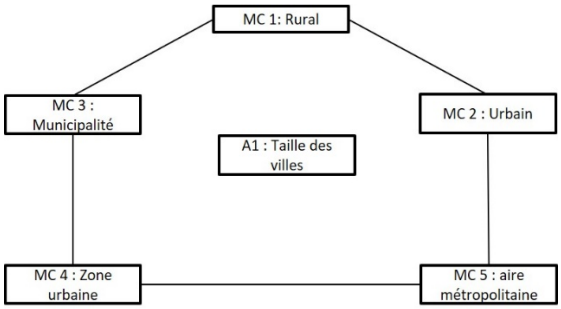
III Le périmètre de recherche

Un périmètre de recherche est un ensemble de documents collectés par Mileva pour adresser la problématique qui lui est posée, c'est une base informationnelle. Ce périmètre de recherche peut être défini « spatialement », par exemple, Google.com, Google.ru ou une base de données interne…, et temporellement, par exemple la période de 2003 à 2015. Ce sont les données contenues dans ce périmètre de recherche qui sont analysées afin de produire des éléments de réponses à la question formulée.

IV Référentiel, agent et mots-clés

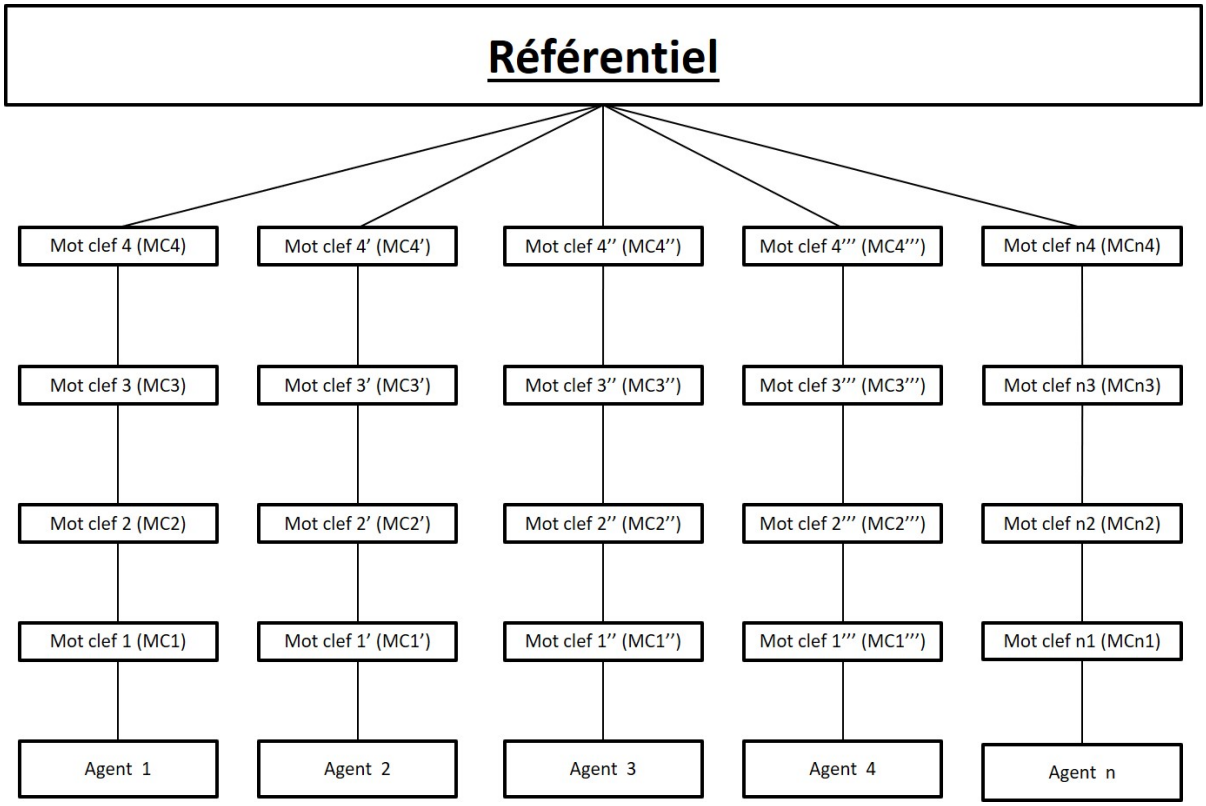
Pour répondre à la question, Mileva a besoin que lui soient transmis les éléments de compréhension du domaine de la question. Par exemple, quels sont les acteurs en présence ? Quelles sont les menaces, opportunités ? Quelles sont les méthodes de travail ?

Ces éléments sont appelés « agents » et constituent la structure du « référentiel » de Mileva. Le système n'a pas de compréhension prédéterminée de ce que recouvrent ces agents, il faut donc les lui définir à partir de mots-clés propres à chacun des agents et les caractérisant. Le référentiel est donc une matrice constituée d'agents et de mots-clés et constitue une représentation de l'univers de la question soumise à Mileva. A travers les mots-clés l'intelligence artificielle va apprendre ce que chaque agent recouvre. Chacun des mots-clés spécifie un agent en particulier.



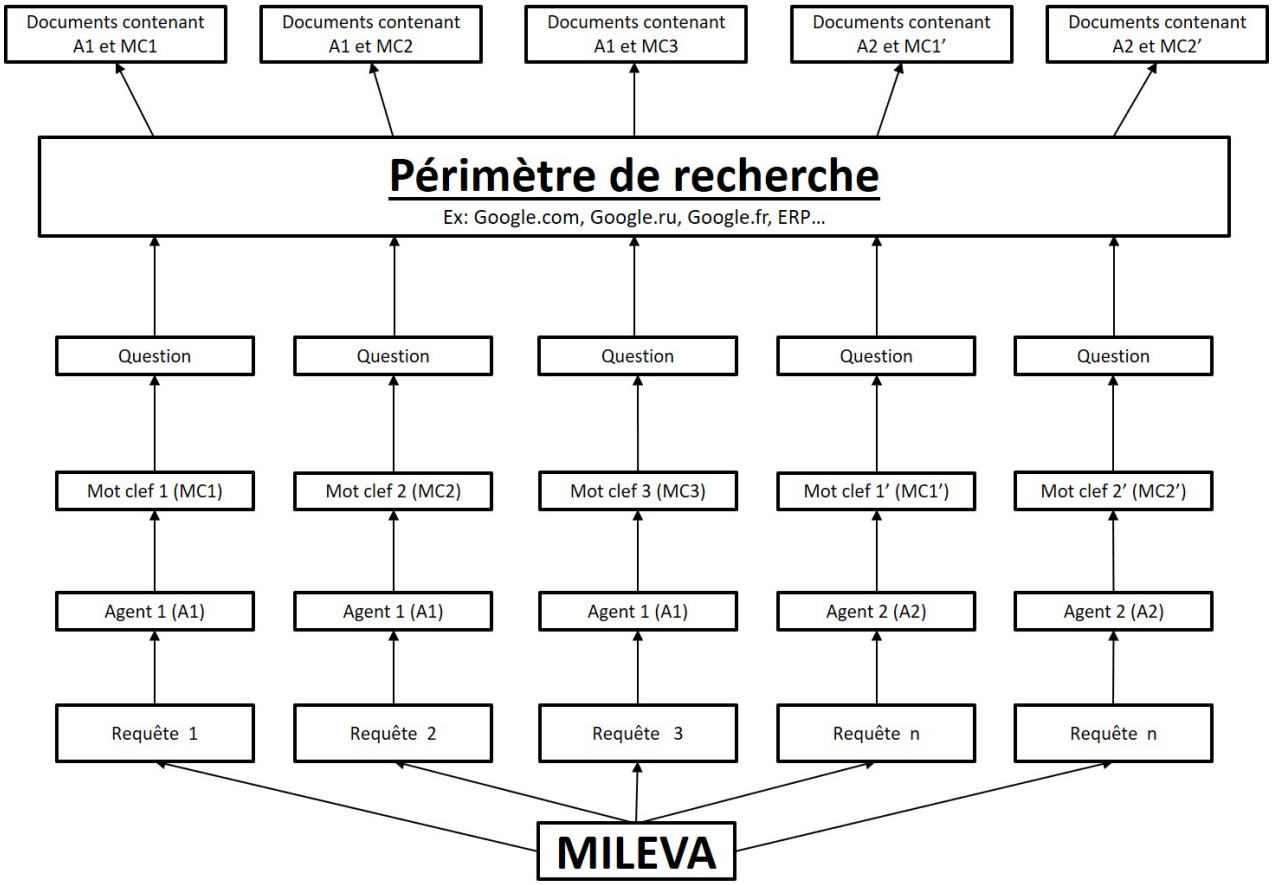
Agent 1 défini par ses mots-clés

Ce référentiel doit décrire, avec autant de précision que possible l'univers de la question posée (les acteurs humains en présence, les risques, les opportunités, les outils, le cadre légal, le cadre politique, le cadre économique, le cadre environnemental…) sous formes d'agents (concepts ou acteurs) définis par des mots-clés. Ce référentiel est construit par l'utilisateur qui peut faire appel à un expert du domaine concerné. C'est une étape-clé du travail avec Mileva. En effet, les agents sont essentiels, car ce sont eux qui, par le jeu de leurs interactions, apportent les éléments de réponse à la question.



V Les requêtes

Une fois le référentiel construit, Mileva émet des requêtes dans le périmètre de recherche afin de l'explorer et de l'analyser. Le système lance des milliers de requêtes de la forme « agent + mot-clé + question » afin de remonter les documents permettant d'identifier des liens entre agents.

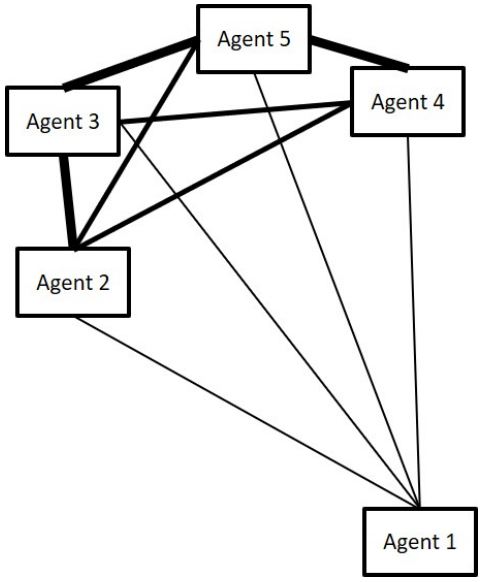


Pour cela, Mileva analyse alors son périmètre de recherche à partir de son référentiel en utilisant la théorie de l'information. Elle procède à l'extraction des informations essentielles, car nouvelles, des documents qu'elle a à disposition dans son périmètre de recherche. Elle procède, ensuite, à la détermination des interactions entre les agents (concepts) de son référentiel. Concrètement, les informations remontées à chaque nouvelle requête sont comparées aux résultats des itérations précédentes afin de déterminer si elles sont nouvelles ou non.

VI Les liens entre agents

Ces liens sont définis par la quantité d'informations échangées : plus il y a d'informations échangées, plus le lien entre les deux agents concernés est fort.

La force de ces liens permet, par la suite, de les positionner les uns par rapport aux autres et, ainsi, de déterminer quels sont les agents à prendre en compte dans la réponse à la question posée. En effet, un agent, ayant des liens forts avec de nombreux agents, est un agent apportant un élément de réponse.

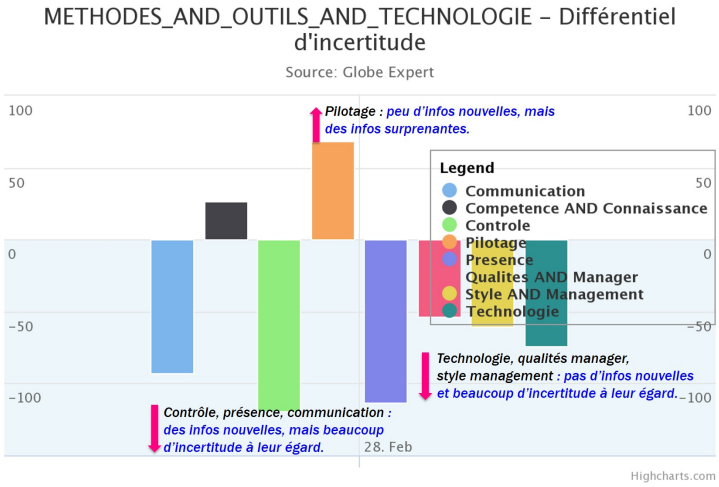


Modélisation des liens entre agents

Les relations entre agents étant calculés, la phase d'apprentissage est terminée. Mileva va maintenant pouvoir faire évoluer les agents au sein de cet univers informationnel et déterminer les comportements optimums pour répondre à la question posée ou aux décisions possibles. Dotés chacun d'une base de connaissance propre, d'un caractère spécifique, ils interagissent entre eux jusqu'à épuisement des possibilités – soit un seuil d'équilibre – et déterminent ainsi des relations de proximité, d'influence, des risques critiques, mais aussi un ensemble de comportements : surprenant, imprévisible, etc. Pour ce faire, Mileva utilise des modèles mathématiques dits évolutionnistes. L'interaction entre agents naît de ces modèles.

VI L'analyse des résultats

Mileva produit des résultats sous forme d'indicateurs (cartes, graphiques, statistiques) représentant des valeurs mathématiques. Ils constituent autant d'éléments qu'il faut analyser, croiser et mettre un perspective, à la manière d'indices dans un jeu de Cluedo. C'est de cette manière que l'on apporte la réponse à la question posée.



Exemple d'indicateur produit par Mileva

VII. Les biais : un point d'attention particulière

Comme toute construction intégrant l'intervention humaine, Mileva est soumise à des biais. Ceci ne constitue pas une faille, mais relève de la nature même d'un outil élaboré par des hommes et analysant un environnement, des situations issus de l'action humaine.

La manière dont la question est formulée, le choix des agents et mots-clés du référentiel peuvent impliquer des biais. Par ailleurs, les données sur lesquelles travaille Mileva comportent des biais liés au périmètre de recherche qui est borné dans l'espace et dans le temps.

Pour toutes ces raisons, **des modèles mathématiques statistiques extrêmement poussées sont utilisées pour mesurer ces biais** et permettre à l'utilisateur de n'accéder qu'aux résultats fiables.

Pour conclure, Mileva est un système d'intelligence artificielle appartenant à la famille des intelligences artificielles cognitives. Elle révèle la trame simple et les briques fondamentales de situations complexes. Elle permet de visualiser leurs évolutions possibles et le niveau de probabilité de celles-ci. C'est, en ce sens, un outil d'aide à la réflexion et la décision.