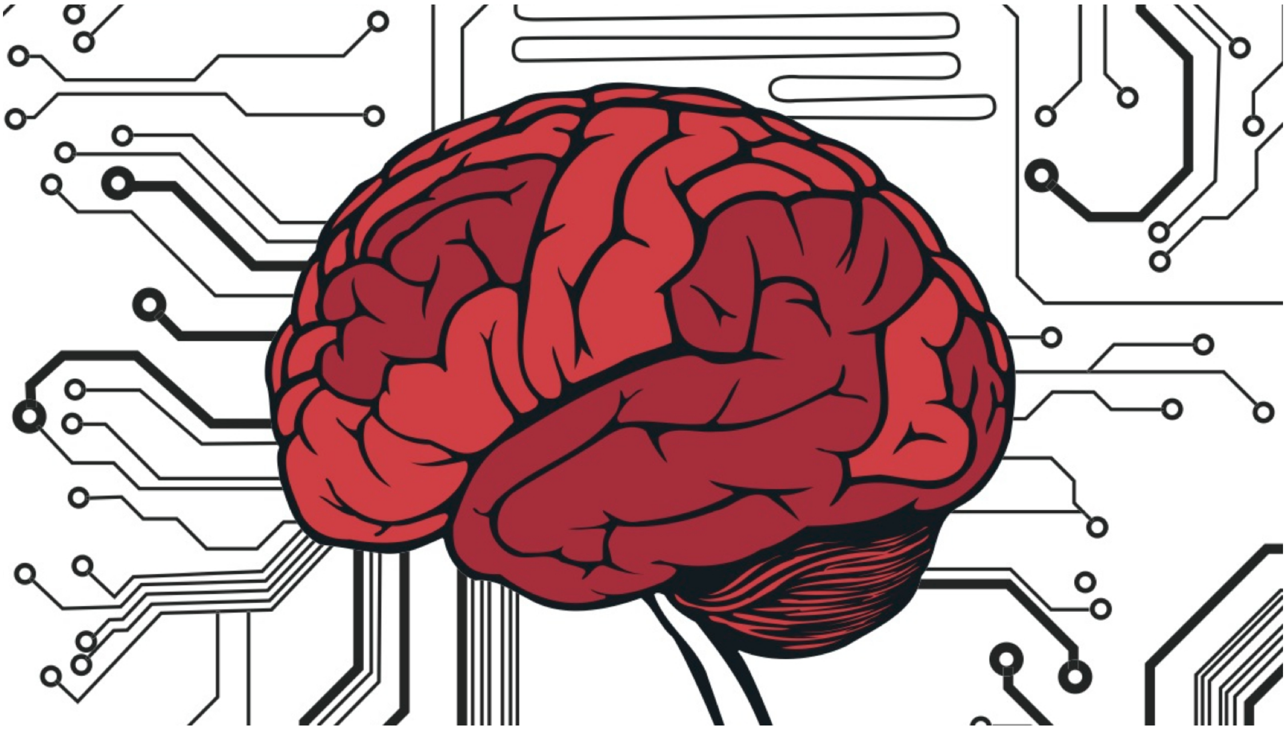


Espaces Conceptuels et Intelligence Artificielle II

L'inférence inductive

Salim Lardjane

Université de Bretagne Sud

Laboratoire de Mathématiques de Bretagne Atlantique

« A computer program capable of acting intelligently in the world must have a general representation of the world in terms of which its inputs are interpreted. Designing such a program requires commitments about what knowledge is and how it is obtained. Thus, some of the major traditional problems of philosophy arise in artificial intelligence. »

J. McCarthy & P. J. Hayes (1969)

Cet exposé est basé sur le Chapitre 6 de l'ouvrage

Conceptual Spaces : The Geometry of thought

de Peter Gärdenfors
MIT Press, 2000

L'induction

Une des caractéristique les plus impressionnantes de la cognition humaine est la capacité d'effectuer des *inférences inductives*.

Sans effort particulier, nous somme prêts, parfois avec un degré de confiance absolu, à généraliser à partir d'un nombre très limité d'observations.

Toutefois, nous n'effectuons pas nos inférences inductives de façon arbitraire.

Peirce (1932) note qu'il y a certaines formes de *contraintes* qui délimitent la classe des inférences possibles et évoque une explication *évolutionniste* de la raison pour laquelle « l'intellect humain est particulièrement adapté à la compréhension des lois et des faits de la nature » (Peirce 1932).

L'induction

D'après Quine (1969) : « To trust induction as a way of access to the truths of nature...is to suppose, more nearly, that our quality space matches that of cosmos ».

Dans la suite on va utiliser la théorie des espaces conceptuels pour développer une théorie des contraintes portant sur l'inférence inductive.

On mettra l'accent sur sur le problème de *projectibilité* c'est-à-dire de la nature des propriétés et des concepts susceptible d'être utilisés dans une induction.

L'induction

L'objectif constructif des sciences cognitives en général, et de l'intelligence artificielle en particulier, est de fournir des modèles computationnels de différents aspects de la cognition humaine.

Comment, donc, *mécaniser* l'induction ?

Comment peut-on espérer capturer la facilité et l'assurance de la compétence humaine d'induction dans un modèle contraint par la rigueur et la précision du calcul ?

L'induction

Il est entendu que l'induction *consiste à passer d'observations particulières à des généralisations*.

Cet énoncé perd son air d'évidence si on s'attaque sérieusement à la question de savoir comment une *observation* est représentée.

Notre leitmotiv sera qu'il n'y a pas de façon unique de caractériser une observation.

Selon la stratégie de représentation des connaissances retenue on peut distinguer *trois niveaux* pour rendre compte des *observations*.

L'induction

1. *Le niveau symbolique* : Cette façon de représenter les observations consiste à les décrire dans un langage spécifié. Le langage est supposé doté d'un ensemble fixe de prédicats et les symboles de ces prédicats sont supposés connus.

L'approche symbolique est au coeur du *positivisme logique*.

2. *Le niveau conceptuel* : Les observations ne sont pas définies en relation avec un langage mais caractérisées par un espace conceptuel sous-jacent. L’induction est alors fortement reliée à la *formation de concepts*. Selon l’approche conceptuelle, les concepts obtenus à l’aide d’inférences inductives présentent des effets prototypiques, par contraste avec l’approche symbolique qui repose sur des concepts aristotéliciens.

L'induction

3. *Le niveau sub-conceptuel* : Les observations sont caractérisées par les entrées de récepteurs sensoriels. Elles sont donc vues comme antécédentes à la conceptualisation. Le processus inductif consiste à établir des connections entre différents types d'inputs. Une façon de plus en plus répandue de modéliser ce type de processus est d'utiliser des réseaux de neurones artificiels.

L’induction

Notre principal objectif dans la suite sera de montrer que, selon l’approche de l’observation qui est retenue, des aspects très différents des inférences inductives entreront en jeu.

L'induction

On abordera quatre questions sur les processus inductifs :

1. Quels concepts (propriétés) peuvent être utilisés pour la généralisation inductive ? (Problème de projectibilité, Goodman 1955)

Cette question se pose en raison de paradoxes liés à l'induction mis au point par Goodman, Hempel et d'autres.

L'induction

2. Comment peut-on *généraliser* en passant d'observations particulières à des *lois* générales ?

On a vu une solution, dans le cadre des espaces conceptuels (Exposé 1) : elle correspond à la façon dont les concepts sont appris. Des problèmes connexes se posent toutefois aux autres niveaux de représentation.

L'induction

3. Comment effectuer des *inférences* à partir d'information limitée sur un objet ?

Ce qu'on entend par "inférence" dépend du niveau de représentation retenu (on le verra dans la suite). En un certain sens, cette question se ramène à la question 2.

L'induction

4. Quelles *connections* peut-on établir entre différents domaines ?

Contrairement aux trois premières questions, celle-ci porte sur la façon dont on peut découvrir de nouvelles relations entre des propriétés ou des concepts de différents domaines. La notion de domaine n'intervient pour ainsi dire pas dans l'approche symbolique; cette question n'est donc importante qu'aux niveaux conceptuel et sub-conceptuel.

Au niveau symbolique, le problème correspondant est de formuler des *lois* générales, typiquement sous la forme d'énoncés universaux ou d'expressions de probabilités conditionnelles.

L’induction

Les quatre questions précédentes sont interdépendantes. Elles relèvent de différents aspects de ce qui est appelé communément "raisonnement inductif".

Elles ont différentes interprétations selon le niveau de représentation choisi.

On va donc discuter du processus d’induction séparément pour chacun des trois niveaux.

INiveau symbolique

Pour les *positivistes logiques*, les objets d'étude fondamentaux sont des énoncés dans un langage plus ou moins formel. Bien souvent, le langage retenu est une version de la logique des Prédicats, où les prédicats atomiques représentent des propriétés "observationnelles".

Le principal outil utilisé pour étudier les expressions symboliques est l'analyse logique. Sous sa forme la plus extrême, le positivisme logique ne s'autorise que cet outil.

Niveau symbolique

Les inférences inductives sont importantes pour les positivistes logiques, car elles constituent un élément cardinal de leur programme.

Il est apparu bien vite, toutefois, que cette approche conduisait à des paradoxes.

Les plus connus sont le *paradoxe de confirmation* de Hempel (1965) et l'*énigme de l'induction (riddle of induction)* de Goodman (1955).

Niveau symbolique

Le paradoxe de confirmation de Hempel traite du problème de savoir quelles observations peuvent être prises en compte comme support inductif d'une loi générale.

Supposons qu'on s'intéresse à une loi de la forme $\forall x : Cx \rightarrow Nx$ (par exemple : "tous les corbeaux sont noirs").

Les instances confirmatoires les plus évidentes sont des énoncés de la forme $Ca \wedge Na$ (corbeaux noirs).

Niveau symbolique

Toutefois la loi générale est équivalente à $\forall x : \neg Nx \to \neg Cx$.

Pour des raisons de symétrie, les observations confirmant cette loi sont de la forme $\neg Na \wedge \neg Ca$ (non-corbeaux non-noirs).

Mais si cela est vrai, on peut confirmer la loi énonçant que tous les corbeaux sont noirs en observant des pommes vertes, des chaussures marron et des poissons rouges, ce qui est clairement paradoxal.

Niveau symbolique

Le paradoxe de Goodman part de l'énoncé que toutes les émeraudes (examinées jusqu'à présent) sont vertes.

La propriété "vleu" est définie comme la couleur de ce qui est vert avant 2019 et bleu après le début de 2019.

De même, "blert" est définie comme bleu avant 2019 et vert après le début de 2019.

Selon cette définition, toutes les émeraudes examinées jusqu'à présent étaient "vleues".

INiveau symbolique

Alors, pourquoi n'accepterait-on pas l'inférence inductive selon laquelle toutes les émeraudes sont "vleues" comme tout aussi valide que celle apparemment plus naturelle selon laquelle toutes les émeraudes sont vertes ?

On retrouve la question 1 énoncée précédemment : Quels concepts (propriétés) peuvent être utilisés pour la généralisation inductive ? ou "Problème de projectibilité" (Goodman 1955).

Niveau symbolique

Notons qu'invoquer la *simplicité* de "vert" par rapport à "vleu" ne résoud pas le problème. Bien que "vert" ne fasse pas intervenir de condition temporelle, contrairement à "vleu", on peut définir "vert" comme "vleu" avant 2019 et "blert" après le début de 2019.

D'un point de vue purement logique, "vleu" et "blert" sont des Prédicats parfaitement symétriques.

Or le point de vue logique, basé sur les représentations symboliques, est le seul retenu par le positivisme logique sous sa forme orthodoxe. D'où le paradoxe.

Niveau symbolique

En plus de ces paradoxes, se posent d'autres problèmes.

Supposons qu'on ait examiné un grand nombre de F (par exemple, des ours bruns) et qu'on ait observé que ce sont tous des G (par exemple, en hibernation).

On a alors tendance à tirer la conclusion inductive que tous les F sont des G.

Niveau symbolique

Mais supposons que tous les F observés soit également incidemment des H (par exemple, observés hors du Canada).

Alors, l'inférence énonçant que toutes les choses qui sont F et H sont également G est au moins aussi justifiée que l'inférence initiale.

Ici également, il n'y a pas de raison logique pour distinguer entre les prédicats dans la tête de l'énoncé universel, en disant par exemple que F est un prédicat plus simple que $F \wedge H$. En effet, on peut partir d'un prédicat K signifiant $F \wedge H$ et d'un prédicat L signifiant $F \wedge \neg H$ puis définir F comme $K \vee L$.

Niveau symbolique

Supposons à présent qu'un grand nombre de F aient été observés comme étant tous des G. Si aucun des objets examinés n'a été un H (un renne, par exemple), alors les observations vont dans le sens de l'énoncé général "Toute chose qui est F ou H est G" (tous les ours bruns et les rennes hibernent) autant que dans le sens de l'énoncé "Toute chose qui est F est G".

La logique ne suffit pas en elle-même à distinguer entre ces généralisations.

Niveau symbolique

Enfin, supposons que toutes les instances examinées de F aient été des G et que toutes les instances examinées de H aient été des K, où F et H sont des propriétés dans des domaines différents, de même que G et K.

Pourquoi estime-t-on que "Toute chose qui est F est G" et "Toute chose qui est H est K" sont de meilleures généralisations que "Toute chose qui est F ou H est G ou K" ?

En effet, ce dernière énoncé possède une classe de confirmation plus importante que celles des énoncés précédents.

La logique inductive traditionnelle, dans le cadre du paradigme symbolique, nous conduirait à dire que le dernier énoncé est le plus valide.

Niveau symbolique

Si on se contente d'utiliser les relations logiques pour déterminer quelles inductions sont valides, le fait que tous les prédicats soient traités de façon identique induit des *symétries* qui ne sont pas préservées par notre compréhension des inductions : "Corbeau" est traité de la même façon que "Non-Corbeau", "vert" de la même façon que "vleu", F de la même façon que $F \vee H$, etc.

On a donc besoin d'une façon *non logique* de distinguer les prédicats qui peuvent être utilisés pour l'inférence inductive de ceux qui ne le peuvent pas.

Niveau symbolique

On trouve plusieurs suggestions de telles distinctions dans la littérature. Goodman qualifie les prédicats pouvant être utilisés pour faire de l'induction de "projectibles".

Equipé de ce terme, la solution de Goodman consiste à introduire la notion d' enracinement ("entrenchment") : un nouveau prédicat d'une théorie en développement peut être utilisé pour faire plusieurs inductions couronnées de succès, et devenir ainsi "enraciné".

C'est une sorte d'induction au second ordre.

D'après Quine (1969) : « We newly establish the projectibility of some predicate, to our satisfaction, by successfully trying to project it. In induction nothing succeeds like success. »

INiveau symbolique

Une autre idée consiste à avancer que certains prédicats sont des "espèces naturelles" ou correspondant à des "propriétés naturelles" alors que ce n'est pas le cas des autres. Ce sont alors seulement les prédicats "naturels" qui peuvent être utilisés dans l'induction.

On peut donner à la notion de "naturel" une interprétation *réaliste*. Celle-ci consiste à dire que les "espèces naturelles" existent dans le monde externe indépendamment de la pensée de quiconque.

Niveau symbolique

Quine (1969) montre que les notions d'espèces naturelles et de similarité sont fondamentales pour notre mode de pensée alors qu'elles sont complètement étrangère à la logique et à la théorie des ensembles.

Ainsi, la méthodologie stricte du positivisme logique succombe sous les problèmes de caractérisation de notions telles que "projectibilité", "espèces naturelles" ou "similarité".

Or, comme le montrent les paradoxes précédents, ces notions sont essentielles pour distinguer les inférences inductives acceptables de celles qui ne le sont pas.

Machine Learning

En sciences cognitives et en Intelligence Artificielle, la distinction entre prédicats projectibles et non projectibles doit pouvoir être implémentée dans un processus computationnel.

Ainsi, le problème des inductions projectibles, comme tant d'autres problèmes en IA, est fondamentalement un problème de représentation de l'information.

Si l'on veut qu'un programme informatiuqe piosse effectuer des inductions basées sur des concepts naturels, il n'est pas suffisant que ceux-ci existent "dans le monde extérieur"; on a besoin de les spécifier dans des expressions computationnelles.

Ainsi, une analyse *conceptuelle* des espèces naturelles est beaucoup plus utile à l'IA qu'une analyse réaliste.

Machine Learning

Historiquement, le type de représentation le plus commun en IA a été *symbolique* en ce sens qu'il a été basé sur un ensemble de règles et d'axiomes, joints à une base de données.

Dans cette représentation les "faits" dans la base de données correspondent à des observations.

Les règles et la base de données sont combinées à l'aide d'un prouveur de théorème ou d'un autre mécanisme d'inférence pour produire de nouvelles règles ou de nouveaux faits.

La "connaissance" initiale ou dérivée constitue le matériel sur lequel un programme peut opérer.

Machine Learning

La forme propositionnelle (symbolique) de représentation utilisée la plupart du temps en IA est bien adaptée à la tradition positiviste.

Lorsqu'il s'est agi d'implémenter des mécanismes d'inférence inductive sur ordinateur, cela a été la méthodologie dominante.

Un exemple typique de la perspective symbolique en IA est le chapitre sur l'induction de Genesereth & Nilsson (1987).

Le problème qu'ils traitent est celui d'obtenir des énoncés généraux à partir d'observations particulières (Question 2 : Comment peut-on *généraliser* en passant d'observations particulières à des *lois* générales ?).

Machine Learning

Pour ce faire, ils supposent qu'il existe un ensemble d'énoncés Γ constituant la *théorie de base* (background theory) et un ensemble Δ d'énoncés de données (à généraliser).

Il est requis que Γ n'implique pas logiquement Δ (sinon, les énoncés de données n'apporteraient aucune information).

Il définissent alors un énoncé ϕ comme étant une *conclusion inductive* si et seulement si ϕ est consistant avec $\Gamma \cup \Delta$ et *explique* les données en ce sens que $\Gamma \cup \{\phi\}$ entraîne Δ.

Machine Learning

Genesereth & Nilsson (1987) voient l'inférence inductive comme un problème de *formation de concept* : « The data assert a common property of some objects and deny that property to others, and the inductive hypothesis is a universally quantified sentence that summarizes the conditions under which an object has that property. In such cases, the problem of induction reduces to that of forming the *concept* of all objects that have that property. »

Machine Learning

Ils définissent un *problème de formation de concept* comme un quadruplet $\langle P, N, C, \Lambda \rangle$, où P est un ensemble d'instances positives d'un concept, N est un ensemble d'instances négatives, C est un ensemble de concepts susceptibles d'être utilisés pour définir le concept à former et Λ est le langage à utiliser pour énoncer la définition.

Machine Learning

Considérons le problème consistant à identifier une classe de cartes, par exemple dans un jeu standard de 32 cartes.

Le langage retenu est un langage du premier ordre, avec un ensemble de prédicats de base comme "numéroté", "face", "impair", "valet", "quatre", "rouge", "pique".

L'ensemble P consiste en les cartes dont on sait qu'elles appartiennent à la classe.

L'ensemble N consiste en les cartes dont on sait qu'elles n'appartiennent pas à la classe.

Le "biais conceptuel" C détermine lesquels parmi les prédicats de base sont autorisés à apparaître dans la règle inductive déterminant la classe.

Par exemple, on peut n'autoriser que "numéroté", "face", "noir" et "rouge" et exclure "corné" et "jouée de la main gauche", entre autres.

Λ est le "biais logique" qui restreint la forme que peut prendre la règle inductive. Par exemple, on peut n'autoriser que les définitions consistant en conjonctions de prédicats de base.

Machine Learning

En partant de la notion de problème de formation de concept $\langle P, N, C, \Lambda \rangle$, Genesereth & Nilsson développent un algorithme effectuant une inférence inductive en respectant les contraintes posées par C et Λ.

Une notion centrale dans leur construction est celle d'*espace de version* (version space) du problème de formation de concept; celui-ci consiste en l'ensemble des règles satisfaites par toutes les instances dans P mais par aucune instance dans N.

Machine Learning

L'algorithme fonctionne en élaguant l'espace de version au fur et à mesure que des instances positives et négatives sont ajoutées.

Dans l'interprétation des résultats, la référence des prédicats de base (*i.e* l'ensemble des objets auxquels ils s'appliquent, Carnap 1947) est considérée comme acquise.

Le résultat de l'algorithme est une combinaison logique de ces prédicats.

Ainsi, la référence des concepts qui sont générés sont restreintes par les références des prédicats de base et les références de ceux-ci ne changent jamais. Par conséquent, l'algorithme ne permet pas d'obtenir de concepts radicalement nouveaux.

Machine Learning

Bien que les chercheurs en IA aient obtenu quelques succès dans leur tentatives de mécaniser l'induction, il apparaît que leur méthodologie souffre des mêmes problèmes que le niveau symbolique en général.

Les paradoxes de Hempel, Goodman et d'autres, s'appliquent tout aussi bien aux programmes d'induction de l'IA traditionnelle.

Machine Learning

Essayer de capturer les inférences inductives par un algorithme souligne également certaines des limitations générales de la perspective symbolique.

Les programmes procèdent en considérant diverses combinaisons logiques de prédicats atomiques, mais les *origines epistémologiques* de ces prédicats ne sont jamais discutées.

Epistémologie : étude de la constitution des connaissances valables (Piaget).

Machine Learning

Bien que les chercheurs en IA ne défendent pas activement la méthodologie positiviste, ils la suivent implicitement en traitant certains prédicats comme donnés de façon observationnelle ou externe.

Le fait que les prédicats atomiques sont pris comme acquis dès le début signifie que de nombreuses inductions ont été effectuées en amont.

Machine Learning

On peut être d'accord avec Genesereth & Nilsson (1987) sur le fait que l'induction symbolique est *une forme* de formation de concept, mais la signification qu'il donnent à la formation de concept est beaucoup trop étroite.

On veut savoir, non seulement la façon dont des prédicats observationnels doivent être combinés au vu d'évidence inductive mais, de façon plus importante, *la façon dont les prédicats de base sont inductivement établis* en premier lieu.

Ce problème a été négligé par les positivistes logiques et leurs épigones dans la recherche traditionnelle en IA.

Machine Learning

L'utilisation de l'analyse logique, principal outil du positivisme et de l'IA traditionnelle, ne permet pas de traiter ces formes de formation de concept.

Pour résumer, l'approche symbolique de l'induction n'autorise pas d'induction créative, ne permet pas d'obtenir de connaissance radicalement nouvelle ni d'effectuer de découverte conceptuelle.

Pour y parvenir, on doit se placer *en deçà* de l'approche symbolique, donc au niveau *conceptuel*.

Niveau Conceptuel

Dans la perspective conceptuelle, l'induction est reliée à la *formation de concepts*.

Le rôle essentiel de l'induction est d'établir des *connections* entre concepts ou propriétés de différents domaines.

Niveau Conceptuel

On a vu que la façon traditionnelle de manipuler les prédicats dans un langage symbolique les traitait tous de façon symétrique. Ainsi, la notion de *domaine* ne joue pour ainsi dire aucun rôle dans l'approche symbolique.

Rappelons qu'au niveau conceptuel, nous avions défini (Exposé 1) un domaine comme un n-uple de dimensions intégrales, c'est-à-dire observées holistiquement.

Niveau Conceptuel

On va montrer dans la suite que l'approche conceptuelle permet de répondre à la question 1 : *Quels concepts ou propriétés peuvent être utilisés dans des généralisations inductives ?*

En particulier, on proposera une solution au paradoxe de Goodman.

Niveau Conceptuel

Commençons par considérer la façon dont les *observations* sont identifiées au niveau conceptuel.

Dans notre contexte, une observation peut être définie comme *l'affectation à un objet d'un emplacement dans un espace conceptuel*.

Généralement, il s'agit simplement d'un *point*, l'espace conceptuel étant doté d'une structure géométrique (Exposé 1).

Niveau Conceptuel

Par exemple, l'observation décrite au niveau symbolique par "x est jaune" est représentée au niveau conceptuel en assignant à x un point de la région jaune de l'espace colorimétrique.

Comme les langages naturels ne divisent le domaine des couleurs qu'en un nombre fini de catégories, l'information contenue dans l'énoncé "x est jaune" est beaucoup moins précise que l'information fournie en assignant à x un emplacement dans l'espace des couleurs.

En ce sens, le niveau conceptuel autorise des instruments beaucoup plus riches pour représenter les observations que le niveau symbolique.

Niveau Conceptuel

Intéressons-nous à la façon dont le prédicat "vleu" de Goodman peut être modélisé dans un espace conceptuel.

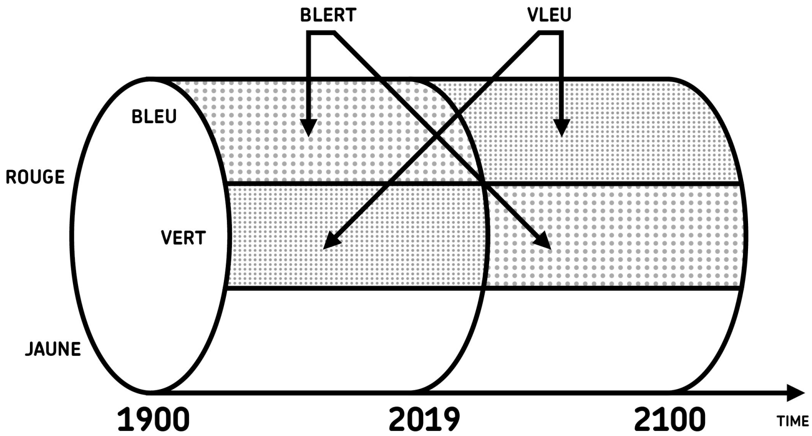
Etant donné les représentations standard des couleurs, vues dans l'exposé 1, "vert" et "bleu" sont des propriétés naturelles au sens de la définition P : On appelle *propriété naturelle* toute région convexe (ou étoilée) d'un domaine d'un espace conceptuel.

Ce n'est pas le cas de "vleu" et "blert". "Vleu" présuppose deux domaines : la couleur et le temps pour sa description.

Pour modéliser le prédicat, on peut considérer l'espace cylindrique engendré en prenant le produit cartésien de la dimension temporelle et du cercle des couleurs.

Ce cylindre est représenté ci-après.

Niveau Conceptuel



Graphique : Alan Souquet, UBS

Niveau Conceptuel

Dans cet espace, "vleu" ne correspond pas à une région convexe.

Carnap (1971) exclut les prédicats du type proposé par Goodman en distinguant entre prédicats "locationels" (locational) et "non-locationels" (nonlocational). "Vleu" est un prédicat locationel car il réfère à une localisation temporelle particulière.

La solution de Carnap est toutefois ad hoc puisqu'elle n'explique pas pourquoi un prédicat non-locationel tel que "vert ou orange" n'est pas considéré comme projectible.

Niveau Conceptuel

Mormann (1993) suggère une solution différente au paradoxe de Goodman, basée sur la notion d'ensemble fermé.

Plaçons-nous dans le cadre du cylindre temps-couleur. Un ensemble *fermé* est défini comme un produit cartésien $A \times B$ où A est un segment de la dimension temporelle et B un arc du cercle des couleurs.

Ainsi, ce qu'il appelle "ensemble fermé" correspond aux "rectangles" du cylindre temps-couleur (à distinguer de la notion d'ensemble fermé en Topologie).

Mormann propose le critère suivant : une propriété naturelle sur le cylindre temps-couleur est un ensemble fermé.

Selon ce critère, "bleu" et "vert" sont des propriétés naturelles, alors que "vleu" et "blert" ne le sont pas.

Niveau Conceptuel

Considérons à présent le paradoxe de confirmation de Hempel (1965).

Non-noir correspond à une région convexe de l'espace colorimétrique puisque toute couleur comprise entre deux couleurs non-noires est également non-noire.

Il est difficile de dire si "non-noir" est projectible ou pas. Par exemple l'énoncé "tous les corps non-noirs reflètent la lumière" apparaît être une généralisation valide.

Toutefois, le concept "non-corbeau" peut difficilement être considéré comme un concept naturel selon le critère de convexité :

Définition C : On appellera *concept naturel* la donnée (i) d'un produit cartésien de régions (convexes ou étoilées) de domaines différents ; (ii) d'un n-uple de *poids* quantifiant la *saillance* des différents domaines; (iii) d'informations sur la *dépendance* entre les différents domaines.

La classe des objets qui sont des "non-corbeaux" recouvre des domaines non reliés entre eux et, du coup, il est difficile, voire impossible, de spécifier les régions associées, de façon à satisfaire le critère de convexité (Gärdenfors 2000).

INiveau Conceptuel

Nolan (1994) remarque que les prédicat de type "vleu" ne peuvent pas être *appris* de la façon avec laquelle on apprend un langage.

Cela s'accorde bien avec notre définition d'un concept naturel.

Etant donné un espace conceptuel, il est beaucoup plus facile d'apprendre un prédicat qui correspond à un concept naturel que d'apprendre une région irrégulière de l'espace.

Niveau Conceptuel

On peut par exemple suivre le mécanisme d'apprentissage consistant à former un prototype à partir d'instances d'une propriété et ce prototype peut être utilisé pour générer une catégorisation de Voronoï (Exposé 1).

Les seules choses dont doit se *souvenir* un agent sont les prototypes des différentes propriétés qu'il considère.

INiveau Conceptuel

Pour revenir aux problèmes liés à l'approche symbolique, notons que même si F et H sont des propriétés naturelles, la disjonction $F \vee H$ n'en est pas nécessairement une.

Si F est une propriété naturelle mais que H ne l'est pas, alors $F \wedge H$ n'en est pas nécessairement une.

Toutefois, si F et H sont naturelles, alors la *conjonction* $F \wedge H$ est naturelle.

Ce résultat est vrai lorsqu'on retient l'interprétation de "propriété naturelle" comme une région convexe.

Niveau Conceptuel

Mentionnons enfin que si des prédicats comme "vleu" ne correspondent pas à des propriétés naturelles dans les espaces conceptuels standard, il est concevable qu'ils correspondent à d'es propriétés naturelles dans d'autres espaces conceptuels où, par conséquent, les prédicats "bleu" et "vert" ne seraient pas naturels.

Ce qui est considéré comme propriété naturelle *est relatif* à l'espace conceptuel sous-jacent.

Le rôle des concepts théoriques

Une autre indication de l'importance des espaces conceptuels est que la *théorisation scientifique* a lieu à ce niveau.

Déterminer les dimensions pertinentes pour expliquer un phénomène est une activité scientifique essentielle.

Une fois l'espace conceptuel d'une théorie fermement établi, des théories, mises sous la forme d'*équations* qui relient les dimensions de l'espace, peuvent être proposées et testées.

Le rôle des concepts théoriques

Un mythe du paradigme symbolique a été de présenter les théories scientifiques comme des ensembles d'énoncés, principalement sous la forme de "lois".

Les défenseurs de l'approche symbolique affirment que des équations telles que la seconde loi de Newton ou la loi de Boyle-Mariotte sont intrinsèquement des expressions symboliques et que la Science ne peut se passer de telles expressions.

Le rôle des concepts théoriques

Deuxième loi de Newton : Dans un référentiel galiléen, la somme vectorielle des forces appliquées à un objet ponctuel est égale au produit de la masse de l'objet par son vecteur accélération.

Loi de Boyle-Mariotte : Le produit de la pression par le volume est constant pour un gaz parfait à température constante.

Le rôle des concepts théoriques

Les équations sont une construction très utile, mais cela ne signifie pas que leur *mode de représentation* est symbolique.

La plupart des scientifiques qui utilisent des équations le font en faisant des *calculs* ; c'est-à-dire qu'ils traitent les variables apparaissant dans es équations comme des variables sur des dimensions particulières dotées d'une structure mathématique (typiquement celle de la droite réelle).

Très peu de scientifiques traitent les équations comme des expressions symboliques, c'est-à-dire comme faisant partie d'un système formel doté d'axiomes et de règles d'inférences.

Le rôle des concepts théoriques

Ainsi, le principal usage des équations a lieu au niveau conceptuel de représentation et non au niveau symbolique.

Ceci dit, si les théories scientifiques sont vues comme des représentations au niveau conceptuel, on doit se poser la question fondamentale : *d'où viennent les dimensions et leur géométrie?* En particulier, *quel procédé permet d'obtenir de nouveaux domaines sans correspondance directe avec nos mécanismes de perception ?*

Ce problème est lié à celui de *l'origine des concepts théoriques* dans les théories scientifiques.

Le rôle des concepts théoriques

En fait, le terme de "concept" théorique est mal choisi, car ce qui est introduit est un *domaine théorique*, qui génère alors de nombreux concepts.

Des exemples classiques de tels domaines sont la dimension de masse de Newton, les principes de Mendel en génétique et la structure atomique de Rutherford.

Le rôle des concepts théoriques

Par exemple, l'espace conceptuel de la mécanique Newtonienne du point est basés sur des dimensions théoriques et pas sur des dimensions phénomenales (psychologiques).

Les dimensions de cette théorie sont l'espace ordinaire (espace euclidien à trois dimensions), le temps (isomorphe à la droite réelle), la masse (isomorphe aux réels positifs) et la force (espace euclidien à trois dimensions).

Le rôle des concepts théoriques

Une fois qu'on a affecté une valeur à une particule sur ces 8 dimensions, elle est entièrement décrite pour ce qui est de la mécanique Newtonienne. Dans cette théorie, un objet est représenté par un point dans un espace à 8 dimensions.

Le rôle des concepts théoriques

En créant sa théorie, Newton a introduit la distinction entre "poids" et "masse". Aujourd'hui cette distinction est omniprésente en Physique bien qu'elle n'aie pas de base sensorielle.

La "masse" est une quantité théorique qui ne peut être déterminée que par des moyens indirects.

Selon Sneed (1971), les quantités théoriques sont celles qui *ne peuvent être mesurées sans appliquer la théorie elle-même*, pour la masse les 3 lois de Newton.

Le rôle des concepts théoriques

Selon la Philosophie des Sciences (par exemple, Popper 1959), l'introduction de concepts théoriques faisant intervenir de nouveaux domaines n'est pas un problème d'induction mais de "contexte de découverte".

Au sein de la Philosophie des Sciences, il a été en général considéré comme futile de vouloir mécaniser le découverte scientifique.

Le rôle des concepts théoriques

Toutefois, vu comme un processus cognitif, le problème n'est peut-être pas sans espoir et peut être vu comme une forme d'induction.

Par exemple, tant le Multidimensional Scaling (MDS) que les cartes auto-organisatrices de Kohonen sont des méthodes qui permettent de générer de nouvelles dimensions qui organisent des données apparemment non structurées.

Le rôle des concepts théoriques

Gärdenfors (2000) va juqu'à proposer que l'introduction de termes théoriques dans les théories scientifiques revient fondamentalement à comprendre comment de nouveaux domaines sont ajoutés à un espace conceptuel phénomenal.

Il suggère que l'introduction de termes théoriques en Science est analogue à l'apparition de nouveaux domaines lors du développement psychologique de l'individu.

Ce type de processus peut expliquer pourquoi un enfant est comme un "petit scientifique" (Gopnik et Meltzoff, 1997).

Le rôle des concepts théoriques

Mais introduire de nouveaux termes théoriques et donc étendre le nombre de domaines pertinents pour un concept signifie que les représentations des concepts deviennent de plus en plus compliquées.

On doit donc se poser la question : *qu'apportent les concepts théoriques ?*

Une réponse succincte est qu'ils autorisent des *nouvelles prédictions*.

Le rôle des concepts théoriques

Ainsi, même sans la mécanique newtonienne, on peut observer le mouvement des corps physiques.

On peut même formuler des lois "observationnelles" comme la loi de la chute libre de Galilée : *en un même lieu et en absence de résistance de l'air, tous les corps ont le même mouvement de chute libre s'effectuant avec la même accélération, quel que soit le corps pesant*.

Toutefois, c'est seulement lorsqu'on introduit la masse d'un objet comme magnitude constante associée à l'objet, qu'on peut prédire la façon dont l'objet va se mouvoir dans des circonstances différentes.

Le rôle des concepts théoriques

Une autre façon d'exprimer l'économie cognitive des termes théoriques est que les magnitudes théoriques peuvent être davantage corrélées avec les caractéristiques usuelles d'un objet que ces caractéristiques entre elles; ainsi la dimension théorique accroît l'homogénéité d'un groupe (cluster) de caractéristiques usuelles, ce qui peut améliorer les possibilités de communication à propos d'une classe d'objets.

Le rôle des concepts théoriques

Un autre point à relever à propos des concepts théoriques est que, dans différents domaines, il y a une différence entre la façon dont la majorité des gens utilisent les concepts et la façon dont les scientifiques les utilisent.

Les baleines ont été historiquement catégorisées comme des poissons alors que les zoologistes disent à présent que les baleines ne sont pas des poissons mais des mammifères.

Le rôle des concepts théoriques

Ceci peut être analysé comme un différence entre la *saillance relative* des domaines impliqués dans le concept de Baleine.

Ainsi, un mode de développement des théories scientifiques consiste à réaliser qu'un domaine particulier est plus important pour la catégorisation cohérente d'une classe de phénomènes que les domaines considérés comme essentiels par la plupart des gens.

Dans ce cas, les espaces conceptuels sont inchangés; c'est juste les saillances des domaines qui sont modifiées.

Le rôle des concepts théoriques

Enfin, un mode plus radical d'évolution des théories scientifiques est celui du changement de paradigme (Kuhn, 1970) où une théorie est remplacée par une autre.

Selon Gärdenfors (2000), ces changements de paradigme peuvent être analysés comme des changements d'espace conceptuel.

Il ne voit pas de différence essentielle entre ce type de changement et le développement de l'espace conceptuel de l'enfant, par exemple.

L'introduction de la distinction entre Hauteur et Volume à 5 ans (Piaget) relèverait du même type de processus que l'introduction de la distinction entre Poids et Masse (Newton).

Niveau sub-conceptuel

En Sciences, l'introduction de nouvelles dimensions théoriques est une découverte qui ne peut être effectuée par des méthodes mécaniques.

En ce qui concerne l'apprentissage humain et la formation de concept - les aspects cognitifs -, par contre, les perspectives de comprendre la genèse des espaces conceptuels ne sont peut-être pas sans espoir car il existe des modèles de la façon dont ces espaces sont formés.

Dans la suite, on va s'intéresser aux processus inductifs au niveau sub-conceptuel.

Niveau sub-conceptuel

Au niveau le plus basique de la modélisation, une observation est ce qui est *transmis* par nos organes sensoriels.

En ce sens, une observation peut être identifiée avec les réactions d'un ensemble de *récepteurs*.

Pour les êtres humains, ces entrées sont fournies par les récepteurs sensoriels, mais on peut également considérer que les machines font des observations de ce type via des instruments de mesure servant de récepteurs.

Niveau sub-conceptuel

Les récepteurs fournissent des données "non interprétées" en ce sens que l'information n'est traitée en aucune manière, ni dans un cadre conceptuel, ni sous la forme d'une expression symbolique.

En Philosophie des Sciences, une distinction est parfois faite entre *perception* et *observation*.

Niveau sub-conceptuel

Shapere (1982) relève que le terme "observation" joue un rôle double pour le philosophe de sciences traditionnel. Il écrit ainsi « On the one hand, there is the *perceptual* aspect : "observation", as a multitude of philosophical analyses insist, is simply a special kind of perception, usually interpreted as consisting in the addition to the latter of an extra ingredient of focussed attention... On the other hand, there is the *epistemic* aspect of the philosopher's use of "observation": the *evidential* role that observation is supposed to play in leading to knowledge or well-grounded belief or in supporting beliefs already attained. »

Niveau sub-conceptuel

La distinction est analogue à celle entre les interprétations scientifiques et phénomenales des espaces conceptuels.

Au sein de la tradition empiriciste de la Philosophie des Sciences, les deux usages du terme "observation" ont été confondus.

Dans la Science moderne, toutefois, il est clair que c'est l'aspect epistémique de l'observation qui est important.

Niveau sub-conceptuel

Shapere (1982) le formule de la façon suivante : « Science is, after all, concerned with the role of observation as evidence, whereas sense-perception is notoriously untrustworthy; ... Hence, with the recognition that information can be received which is not directly accessible to the senses, *science has come more and more to exclude sens-perception as much as possible from playing a role in the acquisition of observational evidence;* i.e., it relies more and more in other appropriate, but dependable, receptors. »

Niveau sub-conceptuel

Une question importante dans ce contexte est : comment obtenir de l'information pertinente à partir de ce qui est reçu par une ensemble de récepteurs ?

C'est la question 3 de l'induction formulée au niveau sub-conceptuel : *Comment effectuer des inférences à partir d'information limitée sur un objet ?*

Niveau sub-conceptuel

Une question supplémentaire concerne la façon dont les trois niveaux de représentation peuvent être reliés les uns aux autres : comment a lieu la transition du niveau sub-conceptuel au niveaux conceptuel et symbolique ?

Ces questions mettent en évidence le type de problèmes inductifs rencontrés au niveau sub-conceptuel.

Niveau sub-conceptuel

Le principal problème est que l'information reçue par les récepteurs est trop riche et trop peu structurée.

Ce dont on a besoin est une façon de transformer et d'organiser l'entrée sur un mode qui peut être géré aux niveaux conceptuel ou symbolique.

Niveau sub-conceptuel

Ceci requiert de trouver un mode plus *économique* de représentation : aller du niveau sub-conceptuel au niveau conceptuel implique habituellement une *réduction du nombre de dimensions* qui sont représentée.

Il existe plusieurs méthodes permettant de traiter ce type de problème. Lors du premier exposé, on a parlé du MDS (Multidimensional Scaling).

MDS

Par exemple, dans l'algorithme de Shepard (1962), les données d'entrée sont supposées contenir de l'informations sur les distance relatives de n points dans un espace inconnu.

Les distances entre les points ne sont pas exprimées sous forme métrique, mais seulement données comme un rang parmi les $n(n-1)/2$ distances parmi n points.

Tout rang de ce type peut être représenté dans un espace à $n-1$ dimensions (Shepard, 1962).

MDS

L'algorithme de Shepard part d'une représentation dans un tel espace puis réduit itérativement la dimension jusqu'à ce qu'aucune dimension ne puisse être éliminée sans qu'il y ait un désaccord substantiel entre le rang déduit des distances métriques et le rang initial.

Dans de nombreux domaines d'applications empiriques, les données initiales peuvent être ramenées à un espace de dimension deux ou trois. Ces dimensions peuvent alors être utilisées comme base pour la formation de concepts.

RNA

Dans la suite, on va présenter une autre méthode que le MDS pour réduire la dimension d'une représentation.

Cette méthode est basée sur un Réseau de Neurones Artificiels (RNA).

RNA

Dans un RNA, les récepteurs et l'information qu'ils reçoivent peuvent être identifiés à un ensemble de *neurones d'entrée* et leur *valeurs d'activité*.

Cet ensemble de valeurs sera appelé *vecteur d'entrée* (ang. input vector).

RNA

Dans les applications concrètes, il y a un grand nombre de neurones d'entrée, ce qui signifie que la dimensionnalité du vecteur d'entrée est très grande.

Le modèle de RNA qu'on va présenter est celui des *cartes auto-organisatrices* de Kohonen (1988, 1995).

Le but de la méthode est de *réduire la complexité de la représentation* des entrées d'une façon efficace et systématique.

Ainsi, la méthode proposée peut-être vue comme une façon de répondre à la Question 2 de l'Induction au niveau sub-conceptuel : *Comment peut-on généraliser en passant d'observations particulières à des lois générales ?*.

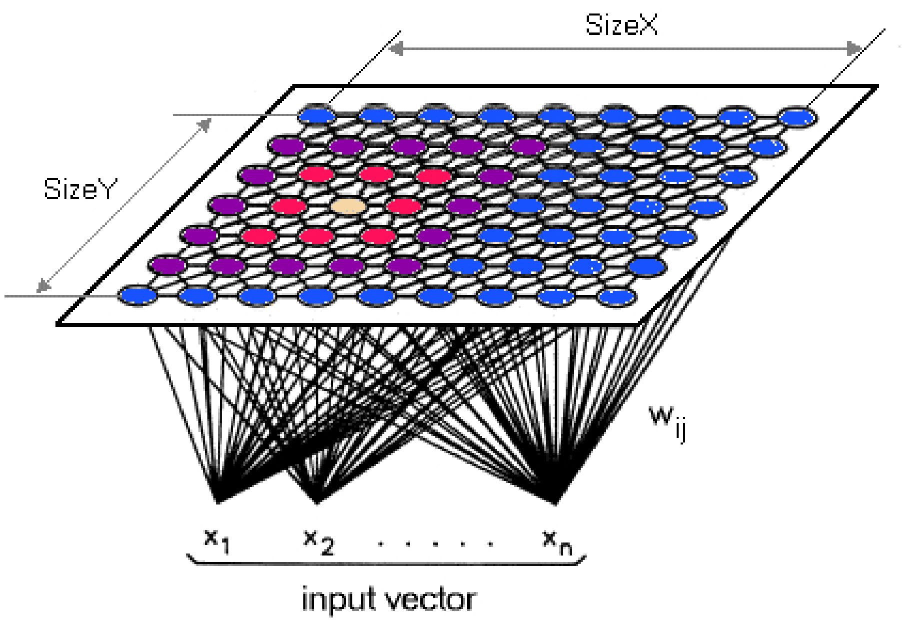
La propriété caractéristique de ces cartes est qu'elles sont à même de décrire la topologie et les relations de similarité des signaux présents dans le vecteur d'entrée, en exploitant un espace conceptuel ayant un faible nombre de dimensions.

Le but de Kohonen en utilisant ces cartes, n'est pas limité à l'inférence inductive mais porte sur la représentation de l'information en général.

Il écrit (Kohonen, 1988) : « Economic representation of data with all their interrelationships is one of the most central problems in information sciences, and such an ability is obviously characteristic of the operation of the brain, too. In thinking, and in the subconscious information processing, there is a general tendency to compress information by forming *reduced representation* of the most relevant facts, without loss of knowledge about their interrelationships. »

RNA

Une carte auto-organisatrice est un RNA consistant en un vecteur d'entrée connecté à une table de neurones de sortie. Dans la plupart des applications, cette table est uni-dimensionnelle ou bi-dimensionnelle, mais elle peut en principe être de n'importe quelle dimension.



RNA

La propriété essentielle du réseau est que les connexions entre les neurones de la table et la fonction d'apprentissage sont organisées de telle façon que des similarités entre vecteurs d'entrée sont en général, *préservées* par l'application, en ce sens que des vecteurs d'entrée ayant des caractéristiques communes sont appliqués sur des neurones *voisins* sur la table de sortie. Le degré de similarité entre deux vecteurs d'entrée est déterminé par une distance.

RNA

L'application du vecteur d'entrée vers la table préserve la plupart des relations topologiques tout en réduisant la dimensionnalité de l'espace de représentation. Préserver la topologie signifie que des points proches dans l'espace à grande dimension sont appliqués sur des points proches de l'espace à faible dimension.

Comme la dimensionnalité est réduite, cela entraîne que des régions entières de l'espace de grande dimension sont appliquées sur des points de l'espace à faible dimension.

RNA

L'application correspondante peut être vue comme une forme de *généralisation* et donc comme une réponse à la question 2 de l'Induction.

La *carte* en faible dimension qui est obtenue peut être identifiée à un espace conceptuel.

RNA

L'application est *générée* par le réseau lui-même via le mécanisme d'apprentissage du réseau.

Un inconvénient de la méthode est qu'elle requiert un ensemble d'apprentissage de taille élevée pour que le réseau se stabilise.

Modèle de Kohonen

Plus précisément, une carte de Kohonen peut être vue comme un réseau de neurones à deux couches de neurones et *deux couches d'interconnexions*.

La première couche de neurones est une *couche d'entrée* dont l'unique fonction est de relayer les entrées au reste du réseau et qui est donc composée de neurones ayant pour fonction d'activation la fonction identité (neurones *linéaires*).

La deuxième couche de neurones est une couche dite d'*adaptation*, composée de neurones positionnés sur un treillis régulier (une "grille") uni-, bi- ou, dans certains cas, tridimensionnelle.

Modèle de Kohonen

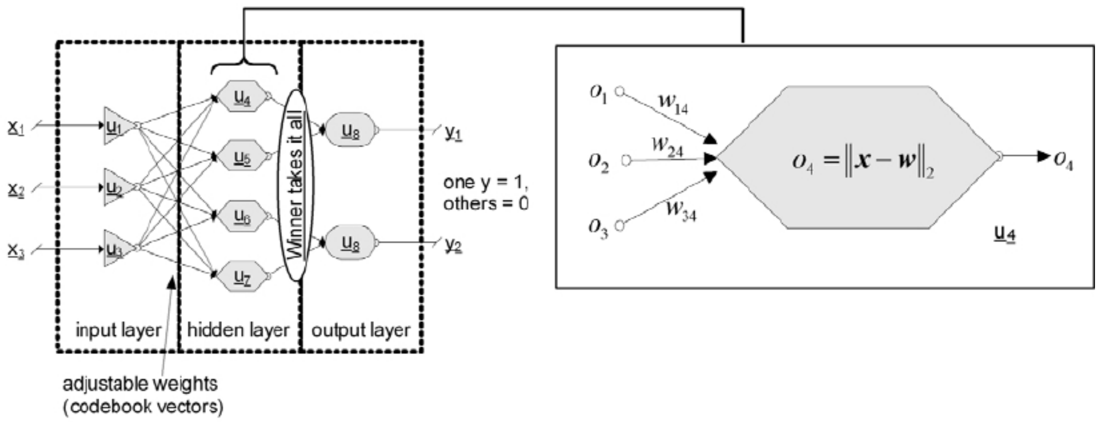
Chaque neurone de la couche d'entrée est relié à tous les neurones de la couche d'adaptation et un *poids* est associé à chacune de ces liaisons.

L'ensemble de ces liaisons constitue la première couche d'interconnexions, dite *couche plastique*.

Tous les neurones de la couche d'adaptation sont des neurones de *type distance* qui calculent la distance entre le vecteur d'entrée et le vecteur de poids qui leur est associé, puis appliquent à cette distance une *fonction d'activation*, généralement l'*identité*.

Modèle de Kohonen

Ceci peut être représenté de la façon suivante :



La deuxième couche d'interconnexions, plus difficile à représenter graphiquement, joue le rôle d'un *réseau compétitif* (également appelé *réseau à inhibitions latérales récursives*) destiné à renforcer *sélectivement* l'activité de la carte de Kohonen.

Modèle de Kohonen

Plus précisément, les *sorties* du réseau sont obtenues de la façon suivante :

$$\begin{aligned} s_j &= 1 \text{ si } ||w_j - x|| = \min_i ||w_i - x|| \\ s_j &= 0 \text{ sinon} \end{aligned}$$

On définit ainsi le (les) neurone(s) "gagnant(s)" ou "actif(s)" comme celui (ceux) dont le vecteur de poids est *le plus proche* des données présentées en entrée.

Modèle de Kohonen

Supposons à présent qu'on dispose d'une *base d'apprentissage* $\{x^p : p = 1, \ldots, N\}$ et qu'on présente les données x^p successivement (ou aléatoirement) en entrée du réseau de Kohonen, alors,

à condition de modifier convenablement les poids après la présentation de chaque nouveau vecteur de données, on pourra aboutir à une situation où *chaque neurone ne sera plus déclaré gagnant que pour un sous-ensemble précis de la base d'apprentissage*.

Modèle de Kohonen

On pourra alors considérer que le vecteur de poids correspondant représente une *classe* d'éléments de la base d'apprentissage.

L'ensemble des vecteurs de poids fournit de ce fait une *classification* des données et le treillis de neurones une *visualisation* de celles-ci en faible dimension.

Algorithme de Kohonen

De nombreux algorithmes ont été proposés pour l'*adaptation* des poids après présentation de chaque nouvel élément de la base d'apprentissage au réseau.

Parmi ceux-ci, le plus ancien et le plus utilisé – mais pas le plus facile à étudier – est l'*algorithme on-line de Kohonen*.

Algorithme de Kohonen

La formule de mise à jour correspondante s'écrit

$$\begin{aligned} w_j &\leftarrow w_j + \eta(x - w_j) \text{ si } j \in \mathcal{V}_g \\ w_j &\leftarrow w_j \text{ sinon} \end{aligned}$$

Algorithme de Kohonen

Une version plus détaillée de celle-ci, faisant apparaître l'ordre de l'itération, n, s'écrit

$$\begin{aligned} w_j(n+1) &\leftarrow w_j(n) + \eta(n)(x^{p(n)} - w_j(n)) \\ &\qquad \text{si } j \in \mathcal{V}_g(n) \\ w_j(n+1) &\leftarrow w_j(n) \text{ sinon} \end{aligned}$$

Algorithme de Kohonen

Les poids sont initialisés de façon aléatoire. $x^{p(n)}$ désigne l'élément de la base d'apprentissage présenté en entrée du réseau à l'itération n.

Celui-ci est en général obtenu en parcourant plusieurs fois la base d'apprentissage de *façon systématique* ou par *échantillonnage aléatoire avec remise* dans celle-ci.

Algorithme de Kohonen

$\eta(n)$ est appelé *taux*, *pas*, ou *gain d'apprentissage*.

Une spécification *typique* de celui-ci est que $\eta(n)$ décroît linéairement de 1 à 0.04 sur les 1000 premières entrées, puis linéairement de 0.04 à 0 sur les 1000 suivantes.

Sous SAS, la décroissance est linéaire de 0.9 à 0.02 sur les 1000 premiers exemples.

Algorithme de Kohonen

$\mathcal{V}_g(n)$ désigne le *voisinage d'adaptation* sur la carte topologique du neurone gagnant à l'itération n.

Une spécification typique de celui-ci consiste à le définir comme l'ensemble des neurones dont la distance *sur la carte* est inférieure ou égale à $\delta(n)$, où $\delta(n)$ décroît linéairement de façon à ce que $\mathcal{V}_g(n)$ recouvre la moitié de la carte (du treillis) en $n = 0$ et ne recouvre plus que le neurone gagnant après la millième itération.

$\mathcal{V}_g(n)$ varie donc au cours du temps et permet de contrôler le nombre de neurones autour du neurone gagnant dont les poids sont modifiés.

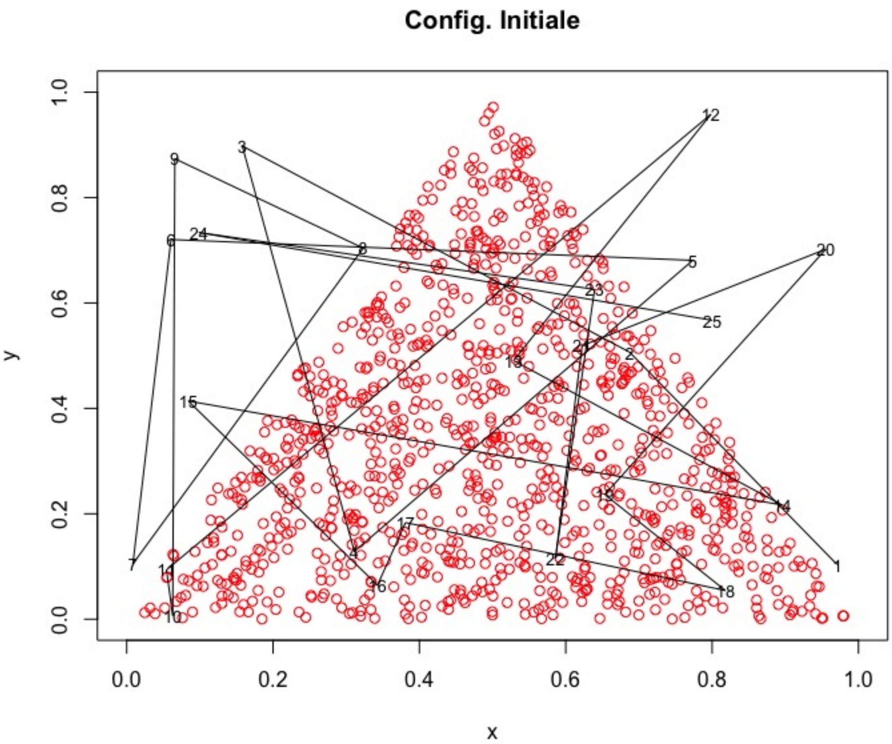
Algorithme de Kohonen

Le moyen le plus simple de percevoir l'effet de l'algorithme de Kohonen consiste à *se représenter* les neurones dans *l'espace des entrées*.

En effet, le vecteur de poids associé à chaque neurone étant de même dimension que le vecteur des entrées, il peut être représenté par un point dans l'espace des entrées, chaque point ainsi obtenu correspondant à un neurone de la carte topologique.

Les vecteurs de poids étant initialisés *aléatoirement*, on part d'une configuration *désordonnée* des neurones de la carte topologique dans l'espace des entrées.

Algorithme de Kohonen

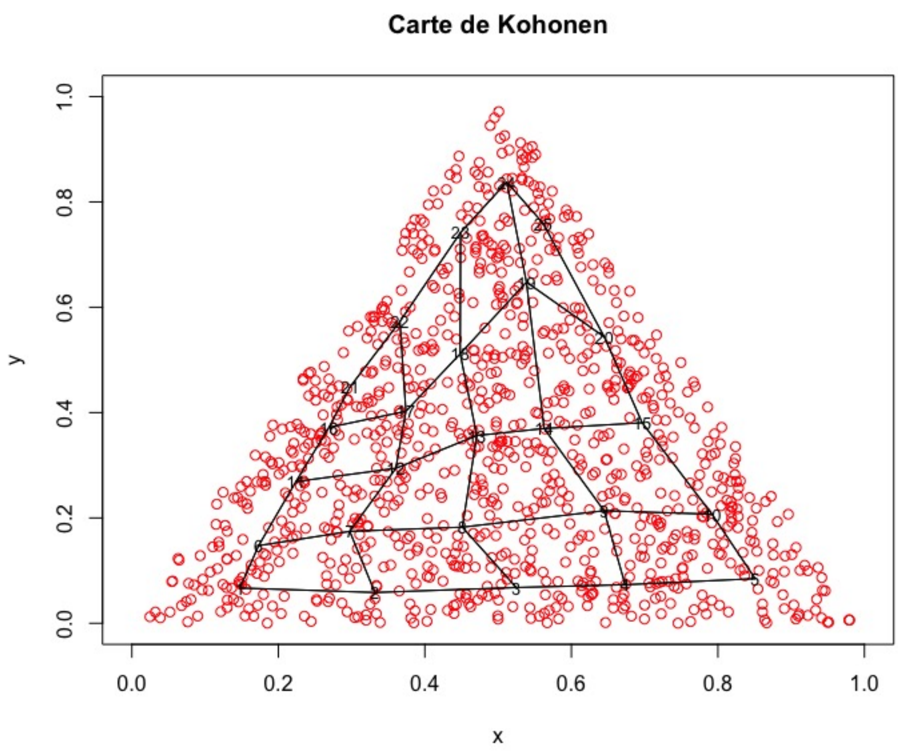


Algorithme de Kohonen

On observe alors, au cours du processus d'*apprentissage*, l'émergence d'une *organisation* du réseau qui produit une *quantification vectorielle* (une *discrétisation*) de l'espace des entrées respectant les relations de proximité sur celui-ci :

deux vecteurs de poids dont les neurones associés sont voisins sur la carte topologiques correspondent alors (sont activés par) des points de l'espace des entrées proches les uns des autres.

Algorithme de Kohonen



Algorithme de Kohonen

On dispose ainsi à l'issue de la phase d'apprentissage d'une "cartographie", sous la forme d'un treillis, de l'espace des entrées, chaque vecteur de poids pouvant être assimilé à *un centre* des vecteurs d'entrée dont il est le plus proche.

On considère en général que l'apprentissage est terminé après un *nombre fixé d'itérations*, généralement quelques milliers.

Corrélations entre domaines

Les cartes auto-organisatrices de Kohonen sont une façon de modéliser la façon dont la structure géométrique d'un domaine peut être créée à partir d'information au niveau sub-conceptuel.

Il y a un autre type de procédé inductif qui est plus typique, toutefois; c'est le procédé consistant à établir des *corrélations* entre les propriétés de *différents* domaines (Question 4 de l'Induction : *Quelles connections peut-on établir entre différents domaines ?*).

Corrélations entre domaines

A niveau symbolique, le résultat de généralisations inductives est typiquement un énoncé universel du type "Tous les F sont des G", où F et G sont des propriétés appartenant à des domaines différents (ou, alternativement, un énoncé probabiliste disant qu'une certaine proportion des F sont des G).

Si on s'intéresse à la faculté qu'a l'être humain de détecter des agrégats de propriétés, on note qu'il est peu doué pour effectuer des évaluations de corrélations *abstraites*.

Corrélations entre domaines

Toutefois, les travaux de Billman (1983) et Billman & Knutson (1996) indiquent que l'être humain est apte à détecter des corrélations qui agrègent *plusieurs dimensions*, en dépit de son inaptitude à détecter des corrélations isolées entre variables.

Une explication plausible de ce phénomène est que nos perceptions des objets "naturels" présentent des corrélations selon plusieurs dimensions et que, comme résultat de la sélection naturelle, nous avons développé une compétence à détecter de telles corrélations agrégées.

En bref, nous sommes plus aptes à apprendre des concepts qu'à apprendre des corrélations isolées entre propriétés.

Corrélations entre domaines

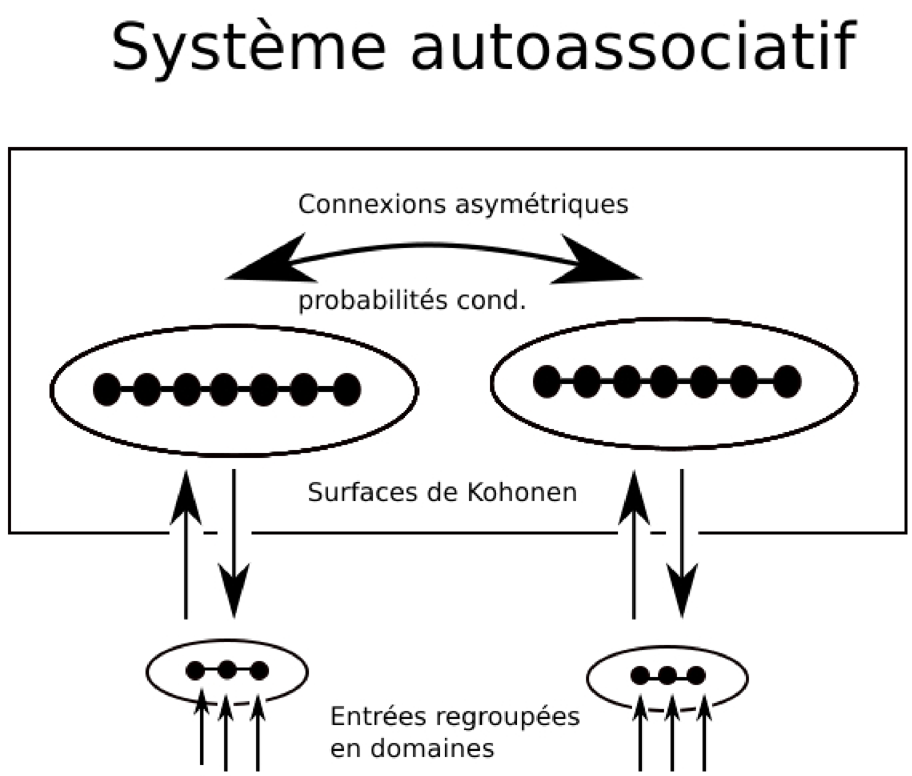
Il existe plusieurs façons de modéliser mathématiquement le procédé d'apprentissage des corrélations entre différents domaines.

Considérons un exemple dû à Balkenius & Gtardenförs (1994, 2000).

Corrélations entre domaines

Le système est basé sur des RNA implémentant les cartes auto-organisatrices de Kohonen (1988).

L'architecture globale du système est décrite dans la figure ci-dessous.



Corrélations entre domaines

Les récepteurs d'entrée sont divisés en un petit nombre de sous-ensembles (il y en a deux sur la figure).

Le but de cette répartition est de grouper ensemble les récepteurs qui contiennent de l'information à propos du même domaine, par exemple : les récepteurs visuels appartiennent à un groupe, les récepteurs auditifs à un autre, etc.

Corrélations entre domaines

Lorsque le réseau est activé, la décision sur la répartition des récepteurs dans les différents groupes doit être faite par l'utilisateur.

Les vecteurs d'entrée sont appliqués sur une carte de Kohonen chacun.

Sur la figure, ces cartes sont uni-dimensionnelles, mais elles peuvent être bi- ou tri-dimensionnelles. De même, il peut y avoir plus de deux cartes.

Corrélations entre domaines

L'une des cartes peut être un pur espace de classification, représentant les "noms" des catégories identifiées par le réseau.

Les cartes de Kohonen sont alors mises en interaction par des connexions asymétriques entre les neurones des deux cartes. Les connections sont totales en ce sens que chaque neurone d'une carte est connecté à tous les neurones de l'autre carte.

La règle d'apprentissage de ces connexions fonctionne de façon à ce que la force de la connexion (c'est-à-dire le poids w_{ij} entre le neurone x_i d'une carte et le neurone y_j de l'autre) reflète la probabilité conditionnelle (estimée à partir des exemples d'apprentissage) pour que y_j soit activé sachant que x_i est activé.

Corrélations entre domaines

Les connexions varient entre -1 et +1, ces valeurs extrêmes n'étant obtenues que lorsque x_i et y_j ne sont, respectivement, jamais ou toujours activés simultanément.

Une fois la phase d'apprentissage terminée, il est possible d'utiliser le réseau pour classifier de nouveaux objets.

En soumettant au système un vecteur d'entrées *partiel* – par exemple les valeurs pour l'un des domaines – le réseau peut calculer les localisations espérées dans toutes les cartes de Kohonen.

Corrélations entre domaines

De cette façon, le réseau "devine" les propriétés inconnues de l'objet sur lequel il n'a qu'une information partielle.

Ainsi, le réseau *généralise* à partir d'expériences et effectue des inférences inductives grâce aux connexions entre les différentes cartes de Kohonen.

Le système permet donc d'illustrer comment il est possible d'effectuer des inférences à partir d'information imitée sur un objet.

Conclusion

L'induction peut être abordée à chacun des trois niveaux de la représentation des connaissances : symbolique, conceptuel et sub-conceptuel.

Les trois perspectives obtenues sur l'induction sont complémentaires plutôt qu'en conflit.

Le niveau conceptuel a toutefois un statut particulier pour plusieurs raisons.

Conclusion

D'abord, les théories scientifiques sont essentiellement formulées sous l'hypothèse d'un espace conceptuel sous-jacent avec un ensemble spécifique de dimensions.

Ensuite, le niveau conceptuel forunit la *sémantique* pour les expressions du niveau symbolique (on l'a vu lors du premier exposé).

Finalement, l'information exprimée à ce niveau peut servir de résumé complet de tout ce qui parvient aux différents types de capteurs au niveau sub-conceptuel.

Bibliographie

Billman, D. O. and Knutson, J. (1996) Unsupervised Concept Learning and Value Systematicity: A complex whole aids learning the parts. Journal of Experimental Psychology: Learning, Memory and Cognition 22: 458-475.

Gärdenfors, P. Conceptual Spaces : The Geometry of Thought. MIT Press, 2000.

Gärdenfors, P. The geometry of meaning : Semantics based on conceptual spaces. MIT Press, 2014.

Genesereth & Nilsson (1987). Logical foundations of Artifical Intelligence. Morgan Kaufmann.

Kohonen (1988). Self-organization and Associative Memory, 2nd edition. Springer.

Kohonen (1995). Self-Organizing Maps. Springer.

Kuhn (1970). The structure of Scientific Revolutions, 2nd edition. University of Chicago Press.

Shapere (1982). The concept of observation in Science and Philosophy. Philosophy of Science 49: 485-525.

Shepard (1962a). The analysis of proximities: multidimensional scaling with an unknown distance function I. Psychometrika 27: 125-140.

Shepard (1962b). The analysis of proximities: multidimensional scaling with an unknown distance function II. Psychometrika 27: 219-246.

Zenker, F. & Gärdenfors, P. Applications of Conceptual Spaces : The case for Geometric Knowledge Representation. Springer, 2015.