

(<https://ristretto.weave.eu>)

PLUS FORT QUE LE BIG DATA : LE SMALL DATA !



Pirmin Lemberger (<https://ristretto.weave.eu/author/pirmin-leMBERGER/>)

■ Digital & Technology (<https://ristretto.weave.eu/category/digital-technology/>) et Exploration & Innovation (<https://ristretto.weave.eu/category/exploration-innovation/>)

🕒 12h



AI (<https://ristretto.weave.eu/tag/ai/>), Data Science (<https://ristretto.weave.eu/tag/data-science/>), deep learning (<https://ristretto.weave.eu/tag/deep-learning/>), machine learning (<https://ristretto.weave.eu/tag/machine-learning/>)



Téléchargez le livre blanc : « IA, où en sommes nous ? »

(<https://ia.weave.eu/>)

Résumé de l'article

Les méthodes d'apprentissage statistiques actuelles, quoique très performantes, conduisent parfois à des prédictions absurdes. Ces techniques requièrent par ailleurs d'énormes volumes de données d'apprentissage pour être capable d'apprendre des concepts aussi élémentaires qu'un panneau de signalisation ou un caractère manuscrit, là où humain se contenterait d'un ou deux exemples. Cet article présente l'une des pistes poursuivies par les chercheurs en IA pour dépasser ces difficultés. Ces avancées reposent sur le concept de **modèle génératif causal** et sur un nouveau type de langage informatique : la **programmation probabiliste**. Ces recherches dessinent une stratégie à long terme visant à doter les machines d'une **intuition physique** et d'une forme de **bon sens social** qui les rendraient extrêmement utiles. Nous illustrerons ces idées au moyen d'un mécanisme de reconnaissance de caractères ultra-efficace développé récemment au MIT et nous comparerons ses performances à celles des humains et à celles des systèmes basés sur le Deep Learning. A plus court terme, la programmation probabiliste offre potentiellement une **alternative intuitive à la construction de modèles prédictifs** qui ne nécessite aucune connaissance approfondie en statistiques ou en machine learning.

*Les sections notées *** pourront être omises pour une lecture rapide, elles présupposent une certaine familiarité avec le calcul des probabilités et la programmation.*

1. **Des pannes de bon sens**
2. **Une piste – les modèles génératifs causals**
3. **Une premier succès – la reconnaissance « one-shot » de caractères**
4. **Un nouvel outil – la programmation probabiliste*****
5. **Conclusion et problèmes ouverts**

« *Le bon sens est la chose la mieux partagée car chacun pense en être si bien pourvu, que même ceux qui sont les plus difficiles à contenter en toute autre chose, n'ont point coutume d'en désirer plus qu'ils en ont.* »

Descartes – 1637

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des

cookies.

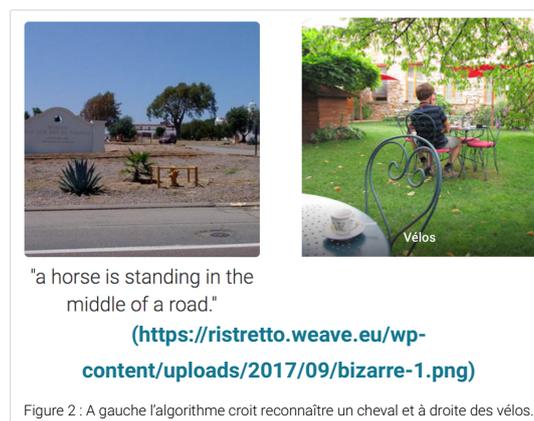
1. Des pannes de bon sens



(<https://ristretto.weave.eu/wp-content/uploads/2017/09/failure-icon-6.png>) Depuis quelques années existent des algorithmes capables d'analyser une scène représentée dans une photo pour en proposer une légende crédible sous forme d'une courte phrase. Ce sont les systèmes de **légendage automatiques**. En voici un exemple plutôt probant :



Il faut bien réaliser ici la performance de ces systèmes car ils ne se contentent pas de sélectionner une légende parmi une liste préétablie mais la construisent à partir des éléments figurant dans l'image. Pourtant, dans certains cas, ces algorithmes font des **erreurs étranges** qu'aucun humain ne commettrait, en voici deux exemples :



Dans l'image de gauche l'algorithme croit reconnaître un cheval qu'un œil humain aurait bien du mal à déceler et dans celle de droite il voit des vélos. Dans le premier cas, on peut imaginer que l'algorithme a été induit en erreur par le paysage en arrière-plan qui évoque un ranch, un endroit qui d'ordinaire contient des chevaux. Dans le second, les arabesques des chaises ont visiblement été prises pour des roues. Les fameux **Adversarial Examples** (<https://ristretto.weave.eu/2017/08/28/nouveau-chantier-securite-it-machine-learning/>) sont du même acabit. Toutes ces erreurs bizarres attirent notre attention sur un problème fondamental : en dépit de l'illusion qu'ils créent, **ces algorithmes** de Deep Learning **ne comprennent pas les scènes qu'on leur montre**, du moins ne les comprennent-ils pas au sens humain du terme. En réalité ils ne font que des associations, très complexes certes, entre les pixels d'une image et une structure grammaticale (la légende de l'image) ou une catégorie. A l'aide de millions d'exemples ils ont laborieusement « appris » à reconnaître des formes de chevaux, de frisbee et de vélos mais n'ont pas de « compréhension » authentique de ces concepts, ce qui explique leurs erreurs pathétiques.

Pour nombres d'applications les algorithmes existants suffisent car ils créent des **illusions utiles**, pensons par exemple aux **chatbots** (<https://ristretto.weave.eu/2017/04/10/chatbots-entre-reves-realite/>). Pour d'autres applications en revanche, comme l'aptitude à soutenir une conversation cohérente, même dans un domaine de connaissances limité, les algorithmes actuels ne suffiront jamais. Mais voyons cela d'un peu plus près.

2. Une piste – les modèles génératifs causals

2.1. L'intelligence ne se réduit pas à la reconnaissance de formes



(https://ristretto.weave.eu/wp-content/uploads/2017/09/pattern_reco.png) Les progrès récents du Deep Learning⁽¹⁾ ont permis, entre autres, de résoudre les problèmes de reconnaissance de formes associés à la **vision artificielle** (hormis les **bizareries**, que nous venons d'évoquer) et à la

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des

cookies.

reconnaissance vocale. Bien qu'il s'agisse là d'avancées historiques de l'IA, on conçoit bien cependant que l'intelligence, quelle que soit la définition que l'on en donne, ne saurait se réduire à ces seules aptitudes sensorielles.

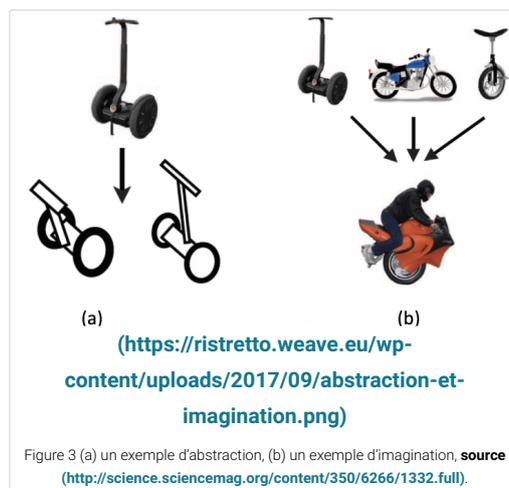
De manière générale on peut affirmer que les IA contemporaines basées sur le Machine Learning statistique ont une intelligence assimilable à un ensemble de réflexes dans des domaines cognitifs très limités. Les tâches qui requièrent un raisonnement, une planification à long terme ou l'élaboration d'une théorie leur échappent complètement.

Dans le cas de l'interprétation d'une scène, notre propre **compréhension humaine** repose en réalité sur un vaste corpus d'idées préalables concernant les causes de ce que nous observons. Pour l'essentiel nous faisons appel à notre **intuition physique** (comment se comporte un frisbee) et à notre **intuition sociale** (pourquoi des individus se réunissent-ils pour jouer au frisbee). Ces idées préalables, ces primitives de connaissance, innées ou acquises par l'expérience sensorielle et par l'éducation, nous dispensent d'avoir à examiner 50'000 frisbee avant d'en comprendre la raison d'être.

Se pose alors la question de savoir comment conférer aux machines un minimum de bon sens physique et social puisque ces aptitudes, même mal définies, semblent caractériser l'essence même de l'intelligence.

Notre cortex cérébral incorpore des modèles de toute sortes de phénomènes physiques et sociaux, combinables à l'infini, si bien que ses performances sont sans commune mesure avec ce que permet aujourd'hui le Machine Learning statistique (ML). Nous pouvons en effet :

- **Comprendre** un concept ou le rôle d'un objet **à partir de quelques exemples** (one-shot learning)
- **Abstraire** un concept pour concevoir des variantes d'un l'objet disposant des mêmes fonctions que l'original (figure 3 (a))
- **Imaginer** des objets que nous n'avons jamais rencontré et combinant des parties de concepts connus (figure 3 (b))
- **Généraliser** une connaissance pour la transférer d'un domaine à un autre
- **Expliquer** une décision, une perception ou un raisonnement
- **Réagir** dans des **situations incertaines** ou sous une **contrainte de temps**



Notons que beaucoup de ces capacités correspondent à des problèmes ouverts tels que nous les avons décrit au chapitre 2 de notre **livre blanc consacré à l'IA** (<https://weave.eu/intelligence-artificielle-le-livre-blanc/>).

Au cœur des remarques précédentes se niche l'idée que, pour construire des systèmes un peu plus intelligents, on ne pourra pas faire l'économie de modèles qui explicitent les causes des phénomènes qu'une IA doit analyser. On parle à ce titre de **modèles génératifs causals** (MGC). Pour rendre compte de l'incertitude qui entache le plus souvent les relations entre les causes et les observations, ces modèles seront **probabilistes**. En d'autres termes, ils ne prédiront pas de certitudes mais les chances selon lesquelles certains phénomènes sont susceptibles de se produire. Ces modèles devront par ailleurs pouvoir être progressivement affiné pour prendre en compte de nouvelles observations et ces améliorations successives correspondront alors à leur capacité d'apprentissage.

Un MGC permet donc d'une part une **prédiction vers l'avant** pour répondre à une question du type « Quels sont les effets probables de telle ou telle cause ? »

Ce qui fait cependant l'intérêt des MGC, comme nous le verrons, c'est qu'ils permettent également de répondre à la question inverse : « Etant donnés certains effets, quels en sont les causes probables ? » Une fois identifiées ces causes probables on pourra alors examiner leurs conséquences au moyen d'une prédiction vers l'avant.

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des

cookies.

Les mathématiques qui permettent de donner corps à ce mécanisme de déduction probabiliste existent depuis belle lurette, il s'agit de la **théorie des probabilités** et, plus précisément, de l'approche bayésienne du machine learning. Cette **inférence bayésienne** formalise une mode de raisonnement rationnel en présence d'incertitudes et, à ce titre, **généralise** donc **la logique ordinaire** qui ne s'applique qu'à des prémisses certaines.

Le ML statistique « classique », qui cherche à découvrir des associations entre de variables prédictives et une cible (plutôt qu'à exploiter des modèles causals), peut être formulé avec ces mêmes outils. Le paragraphe 2.3 clarifiera cette distinction entre les ces deux formes de ML.

Dans le paragraphe 2.2 nous nous contenterons de motiver l'intérêt de ces MGC en montrant comment ils pourraient permettre d'inculquer **une physique** et **une psychologie intuitive** aux machines. Pour rendre ces idées encore un peu plus concrètes, la section 3 présentera de manière plus détaillée une mise en œuvre réussie des idées de MGC dans le cadre d'un système de reconnaissance de caractères one-shot. La section 4 présente la programmation probabiliste comme l'outil idéal pour ce type de modélisation.

L'approche de l'IA esquissée ici est développée au MIT par une équipe pluridisciplinaire constituées d'experts en sciences cognitives, de spécialistes de l'IA et d'informaticiens sous la houlette du **prof. Josh Tenenbaum** (<http://web.mit.edu/cocosci/josh.html>). Leur ouvrage **Probabilistic Models of Cognition** (<http://probmods.org/>) propose une présentation détaillée et accessible de cette approche passionnante de l'IA.

2.2. La quête du bon sens

L'intelligence nécessaire à la compréhension d'une scène de la vie courante, ce qu'on appelle communément le **bon sens**, pourrait être définie comme l'aptitude d'une intelligence à **identifier les causes les plus vraisemblables de ce qui est observé** puis, à partir de là, à faire des prédictions à propos de d'autres évènements. Insistons sur le fait qu'à ce jour il n'existe **aucun exemple opérationnel** de ce genre de système intelligent. Il s'agit là de travaux de recherche en cours.

2.2.1 Les étapes de l'inférence bayésienne***

Avant de passer aux deux exemples d'intuition physique et psychologique, précisons les étapes du mécanisme d'inférence qui leur est commun.

1. On formule plusieurs hypothèses sur les causes possibles de ce que l'on observe. Sur la base d'expériences passés on estime que chaque cause C_i possède une certaine **probabilité à priori** $P(C_i)$.
2. **Prédiction vers l'avant** : on calcule la vraisemblance $P(O|C_i)$ selon laquelle chaque cause C_i pourrait être responsable de l'observation O . C'est précisément là qu'on utilise le modèle génératif.
3. **Inférence bayésienne** : on calcule la **probabilité à postériori** $P(C_i|O)$ de chaque cause C_i sachant qu'on a observé O au moyen de la célèbre formule de Bayes $P(C_i|O) \propto P(O|C_i) P(C_i)$.
4. On peut ensuite chercher la cause la plus probable C_{\max} de l'observation O en maximisant $P(C_i|O)$ parmi tous les C_i .
5. On peut aussi exploiter la connaissance des probabilités des causes $P(C_i|O)$ pour faire d'autres prédictions $P(A|O)$ concernant un phénomène A à condition bien sûr d'avoir un modèle pour $P(A|C_i)$.

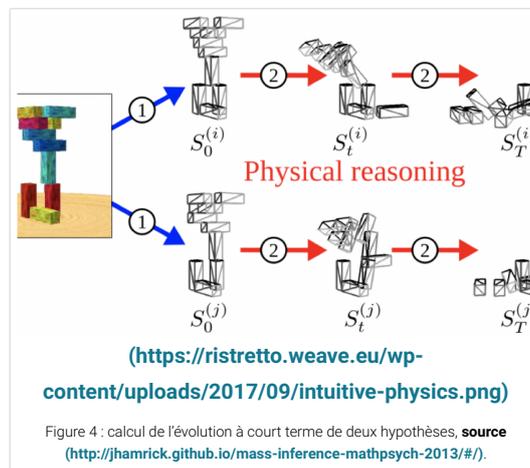
La mise en pratique de cette démarche est délicate pour deux raisons. D'une part les **modèles génératifs** $P(O|C_i)$ sont souvent **complexes** et couteux en calculs pour des modèles réalistes de phénomènes physiques ou de comportement humain. D'autre part, le **nombre de causes** C_i à examiner est en général **très grand** et il faudra recourir à des heuristiques pour éviter d'avoir à les considérer toutes.

2.2.2. Elaboration d'intuitions physiques

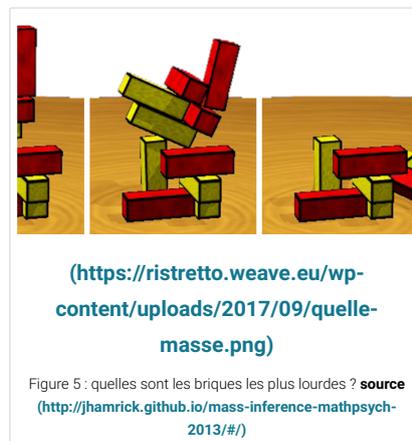
En réalité les machines maîtrisent fort bien les lois de la physique, du moins celles qui s'expriment au moyen des belles équations de Newton ou de Maxwell. En témoigne notre capacité à envoyer des sondes spatiales explorer le système solaire avec une précision inouïe ou à simuler une réaction de fusion thermonucléaire.

En revanche, elles ne maîtrisent pas la **physique intuitive**, celle qui nous permet, sans effectuer de longs calculs, de nous mouvoir en sûreté dans notre environnement, d'anticiper qu'une pile de livres est sur le point de tomber, que telle surface est vraisemblablement glissante ou que tel objet risque bien d'être trop lourd pour être soulevé.

Cette physique intuitive que les humains acquièrent au début de leur vie selon un processus encore mystérieux fait aujourd'hui l'objet de nombreuses recherches avec l'objectif d'en doter les machines. Une part de l'idée usuelle de bon sens consiste en effet à anticiper rapidement quels mouvements sont possibles pour une collection objets.



Les techniques de reconnaissance visuelle ne peuvent ici suffire. Le développement d'une intuition physique demande en fait de combiner l'inférence bayésienne (pour remonter des observations aux causes) avec la **simulation de l'évolution à court terme** du système pour juger de ce qui va advenir dans les instants qui suivent l'observation. Les moteurs physiques développés pour les **jeux vidéo** se sont avérés d'une grande utilité pour cette tâche car ils calculent aussi bien l'évolution d'une scène que l'image associée. L'hypothèse est que notre cerveau, et plus généralement celui des animaux, intègre dans son fonctionnement l'équivalent d'un tel moteur. L'objectif à terme de ces recherches est de comprendre comment les humains déduisent les propriétés physiques des objets qu'ils observent, comme leur masse par exemple, et comment ils le font si rapidement.

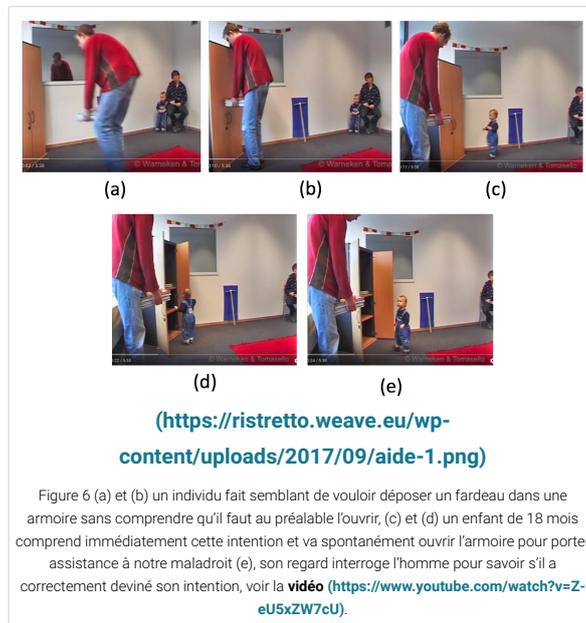


2.2.3. Elaboration d'intuitions psychologiques

Développer une intuition psychologique c'est se faire une idée des intentions, des souhaits ou des préférences probables d'un individu conçus comme les causes premières de son comportement. Dans ce contexte, l'identification des causes à partir des observations d'un comportement s'appelle **Inverse Reinforcement Learning**⁽²⁾, comme nous l'avions évoqué dans notre **livre blanc** (<https://weave.eu/intelligence-artificielle-le-livre-blanc/>).

L'analogue du moteur de jeu vidéo qui permettait de propager dans le futur une situation physique est en l'occurrence un **moteur de planification** qui, à partir d'une intention, planifiera un enchaînement d'actions permettant de la concrétiser. En d'autres termes, un agent doté d'intuition psychologique est un agent capable de faire des hypothèses explicatives quant au comportement d'un humain qu'il observe pour en déduire ensuite ses motivations probables par inférence bayésienne. Basés sur ces causes probables le système pourra alors anticiper d'autres comportements vraisemblables dans le futur.

Des machines qui seraient capables de détecter les intentions des humains et, simultanément, de faciliter leur exécution auraient alors saisi l'essence de la **notion d'aide** et on conçoit aisément à quel point elles pourraient être utiles. Le rêve de tout concepteur d'IA serait de parvenir à réaliser la rétro-ingénierie du comportement illustré dans la figure 6 où l'on voit un enfant de 18 mois faire preuve d'un altruisme spontané réjouissant ! Aujourd'hui nous en sommes loin, mais tel est l'objectif de ces recherches.

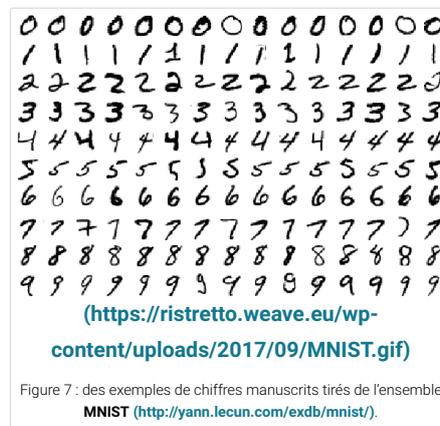


Plus ambitieuse encore serait la réalisation d'une IA qui, non contente de deviner les intentions des humains pour les aider, devinerait par surcroît ce que ces mêmes humains anticipent de son propre comportement. Voilà donc une IA qui anticiperait les conséquences de son comportement sur les hommes et que l'on pourrait qualifier d'**IA sociable**.

2.3. Comparaisons avec le ML supervisé classique

Pour tenter de clarifier notre propos nous décrivons dans cette section ce qui différencie sur un plan conceptuel l'approche par **MGC** de l'**approche classique du ML supervisé** (https://ristretto.weave.eu/2017/08/28/nouveau-chantier-securite-it-machine-learning#sect_2) tel que le pratiquent les data scientists qui utilisent des API comme *scikit-learn* ou *TensorFlow*.

Il est utile pour mieux comprendre les idées présentées d'avoir présent à l'esprit un exemple concret. Rappelons donc brièvement le problème de la **reconnaissance de caractères manuscrits** qui est l'un des benchmarks régulièrement utilisés pour évaluer les performances d'un algorithme de reconnaissance d'images. La section suivante présentera en détail comment appliquer les idées des MGC à ce problème précis.



Le but est de construire un système prédictif qui, une fois entraîné sur N d'exemples $\{(x_1, y_1), \dots, (x_N, y_N)\}$ correctement identifiés, sera capable d'associer un nouveau caractère x à un chiffre y parmi $\{0, 1, \dots, 9\}$ (ici x_i représente la matrice de pixels de l'image et y_i le digit qu'elle représente). Le cas $N = 1$ est celui qui nous intéressera dans la section 3.

Le tableau suivant résume les **principales différences entre les deux approches**. Elles sont décrites plus en détail dans les paragraphes 2.3.1, 2.3.2 et 2.3.3.

Principales différences entre le ML supervisé et l'approche par MGC

	ML supervisé classique	Approche par modèle génératif causal
Ce qui est modélisé	La distribution de proba $P(y \mathbf{x})$ des valeurs d'une cible y à prédire en fonction de variables prédictives \mathbf{x}	Les mécanismes causaux C susceptibles d'engendrer les observations \mathbf{x} dans un certain domaine

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies.

Apprendre	Trouver une bonne estimation $y_{\text{estimé}}(\mathbf{x})$ de y en fonction de \mathbf{x} qui conduit à faire peu d'erreurs sur une échantillon d'entraînement	Identifier un ensemble de mécanismes causals pour les phénomènes pour lesquels on souhaite faire des prédictions
Faire une prédiction	Trouver la cible y la plus probable si on a observé \mathbf{x}	Trouver le mécanisme causal C qui explique au mieux l'observation \mathbf{x} puis interroger C sur tout ce qu'il peut expliquer
Type de généralisation possible	Locale : on ne peut prédire y qu'au voisinage immédiat de points \mathbf{x} déjà observés ⇒ prédiction pauvre	On peut prédire tout ce qui s'explique au moyen des mécanismes causals C ⇒ prédiction riche
Volume de données pour l'entraînement	Big Data	Small Data
Importance du travail de conception préalable	Faible	Important pour l'instant
Possibilité de générer des exemplaires d'un concept	Non	Oui

2.3.1. Nature de la modélisation***

Approche ML supervisé

Dans l'approche classique du **ML supervisé** on **modélise directement la dépendance** entre des variables prédictives \mathbf{x} et une cible de prédiction y . Dans notre exemple, on postule l'existence d'une distribution de probabilité $P(y|\mathbf{x}; \mathbf{w})$ qui, à chaque image \mathbf{x} , associe une probabilité (un score compris entre 0 et 1) pour que \mathbf{x} représente le digit $y \in \{0, 1, \dots, 9\}$, \mathbf{w} étant les paramètres ajustables du modèle. Un tel modèle, où \mathbf{x} sont des variables prédictives et où y est une variable aléatoire, s'appelle un **modèle discriminant** dans la littérature ML.

Avant d'observer un échantillon d'images correctement identifiées on ne sait pas grand-chose des paramètres \mathbf{w} , si ce n'est qu'ils ne peuvent être trop grands pour éviter le sur-apprentissage. On incorpore cette connaissance a priori au modèle sous forme d'une distribution $P_{\text{prior}}(\mathbf{w})$ qui pénalise les \mathbf{w} avec de grandes composantes.

Après observation d'un échantillon $S = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$ d'images correctement identifiées on pourra calculer une nouvelle distribution de probabilité $P_{\text{posteriori}}(\mathbf{w}|S)$ qui prendra en compte ces données expérimentales S . C'est l'**approche bayésienne** du ML⁽³⁾.

L'**apprentissage** n'est dès lors rien d'autre que le calcul de $P_{\text{posteriori}}(\mathbf{w}|S)$.

Les **prédictions** d'un digit y pour une nouvelle image \mathbf{x} s'obtiennent en combinant (par marginalisation sur les \mathbf{w}) les distributions $P_{\text{posteriori}}(\mathbf{w}|S)$ et $P(y|\mathbf{x}; \mathbf{w})$ pour obtenir la distribution à postériori $P_{\text{posteriori}}(y|\mathbf{x}; S)$ qui nous intéresse. Le digit y_{max} prédit par le modèle est celui qui maximise $P_{\text{posteriori}}(y|\mathbf{x}; S)$.

Approche par MGC

Restons dans le même contexte. Cette fois nous serons beaucoup plus ambitieux puisque nous chercherons à modéliser le **mécanisme causal** C qui donne naissance aux images \mathbf{x} de symboles manuscrits que l'on souhaite reconnaître, qu'il s'agisse de chiffres ou d'autres symboles. Ce mécanisme C est en l'occurrence constitué d'une succession des gestes d'écriture qui permettent de tracer un caractère comme un '7' ou un 'K'. Comme ce processus est manuel, chaque instance d'un '7' ou de 'K' est un peu différente de la précédente. On modélise ces petites variations avec distributions de probabilité sur les images $P_{\text{causal}}(\mathbf{x}|C='7')$ ou $P_{\text{causal}}(\mathbf{x}|C='K')$, c'est le modèle génératif.

Très schématiquement, reconnaître une image \mathbf{x} revient alors à identifier les mouvements C vraisemblablement à l'origine de l'image \mathbf{x} . Le point central de cette approche est dès lors le suivant :

La connaissance du mécanisme causal $P_{\text{causal}}(\mathbf{x}|C)$ de génération des images \mathbf{x} fournit une quantité d'informations qui n'étaient pas disponibles dans un modèle discriminant $P_{\text{posteriori}}(C|\mathbf{x})$. D'une part la reconnaissance d'un caractère sera plus précise et, d'autre part, il sera possible d'échantillonner $P_{\text{causal}}(\mathbf{x}|C)$ pour créer de nouveaux exemplaires d'un caractère.

2.3.2. Type de généralisation et volume des données d'entraînement

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies.

Approche du ML supervisé

Comme nous l'avions expliqué dans notre **survol du ML supervisé** (https://ristretto.weave.eu/2017/08/28/nouveau-chantier-securite-it-machine-learning/#sect_2), le ML présuppose que l'on dispose d'un ensemble de données représentatif assez dense du phénomène pour lequel on souhaite faire des prédictions. Pour un problème comme la reconnaissance de caractères il faudra donc disposer d'un grand réservoir de '7' et de '3' avec toutes sortes de variantes pour que le modèle soit capable de reconnaître ces symboles. C'est la raison d'être de jeux de données comme le **MNIST** (<http://yann.lecun.com/exdb/mnist/>) (**Big Data**).

Dans ce sens le ML supervisé n'est capable en réalité que de **généralisations très locales**. En d'autres termes il fait des prédictions à condition d'avoir « vu » beaucoup d'exemples très similaires à celui qu'on essaie de reconnaître.

Approche par MGC

Ces capacités limitées du ML sont à contraster avec celles d'un humain qui, ayant vu quelques exemplaires (**Small Data**) du symbole '0' sera probablement capable d'en reconnaître d'autres et même d'en dessiner de nouveaux exemplaires. Un humain est capable de comprendre le concept de '0' car il possède dans son esprit l'équivalent d'un modèle génératif de caractères qui lui permet d'imaginer comment ce '0' a été dessiné. Il **induit les causes à partir des conséquences**, il **crée une théorie** de la génération des '0' en général et, dans ce sens procède, à une **généralisation à grande échelle**.

Procéder à la rétro-ingénierie de ce processus cognitif d'induction des causes à partir des effets est précisément le but de l'approche par la MGC.

2.3.3. Travail de conception

Potentiellement les MGC offrent donc de meilleures performances de généralisation que le ML supervisé. Pour autant, celles-ci ne vont pas sans contreparties. En effet, il faut bien réaliser qu'on a troqué un important volume de données avec une **grande complexité du modèle causal** qui doit conçu à neuf pour chaque nouveau problème. Or, pour l'instant, ce travail de conception n'a rien d'automatique est demande une **importante expertise humaine** comme nous en verrons dans l'exemple de la section 3. Les capacités de généralisation d'un système basé sur un MGC sont à l'exacte mesure de la sophistication du modèle causal qu'il embarque.

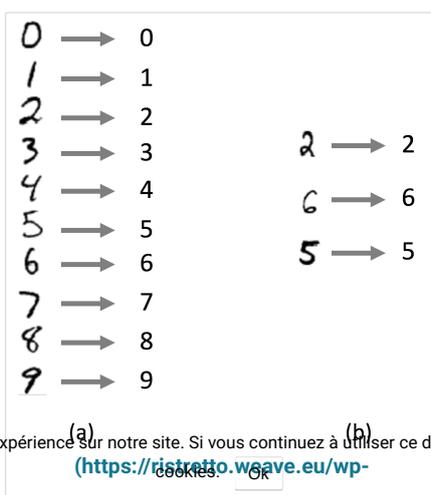
On peut spéculer qu'il manque pour l'instant à l'IA un principe général qui, partant de quelques primitives, permettrait de construire hiérarchiquement des modèles causals par encapsulation successives de modèles toujours plus complexes. Une esquisse de cette **idée de modèle hiérarchique** est toutefois concrétisée dans le travail sur la reconnaissance de caractères présenté dans la section 3.

3. Une premier succès – la reconnaissance « one-shot » de caractères

(<https://ristretto.weave.eu/wp-content/uploads/2017/09/caractère-chinois.png>) Dans cette section nous présentons les grandes

lignes d'un travail de recherche pionnier mené récemment au MIT [TEN1]. Il démontre quantitativement la pertinence d'un MGC pour résoudre un problème spécifique en comparant ses performances à celles des meilleurs algorithmes

de Deep Learning et à celles des humains. Le problème en question est celui de la **reconnaissance de caractères manuscrits** dite **one-shot**, où il s'agit d'apprendre à reconnaître les caractères d'un alphabet à partir d'un seul exemplaire par caractère.



Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies. ([https://ristretto.weave.eu/wp-](https://ristretto.weave.eu/wp-content/uploads/2017/09/caractère-chinois.png)

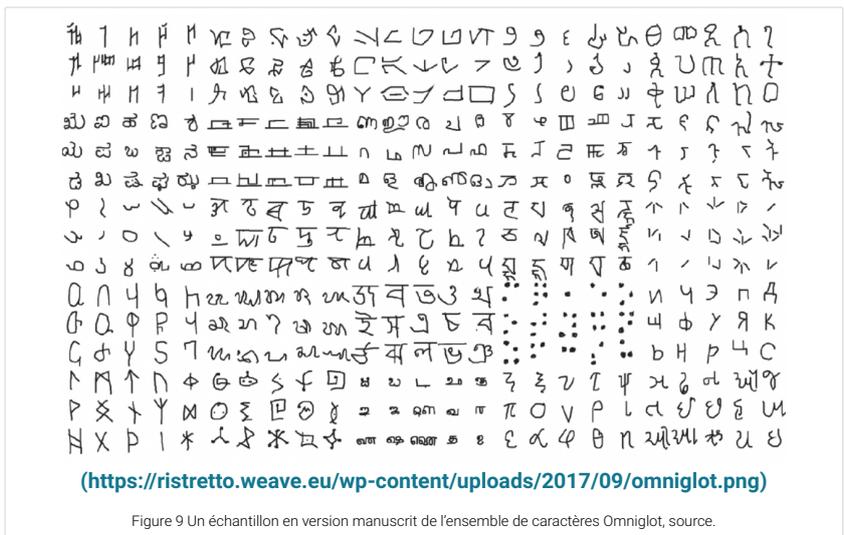
<content/uploads/2017/09/one-shot.png>

Figure 8 (a) apprentissage one-shot avec un exemplaire par caractère à reconnaître (b) reconnaissance de nouveaux caractères d'un ensemble de test

Les détails du modèle probabiliste sont assez techniques [TEN1], aussi ne présenterons-nous que les principales étapes de construction du MGC.

3.1. Créer le jeu de données

Dans l'approche classique de la reconnaissance de caractères par le ML, la première étape consiste à collecter un grand nombre d'exemplaires de chaque caractère à reconnaître, des centaines voire des milliers, pour créer un jeu de données d'entraînement tel que celui du MNIST de la figure 7. Un tel ensemble contient donc **beaucoup d'exemplaires de peu de concepts**. L'approche utilisée dans [TEN1] pour créer un MGC est radicalement différente. Elle exploite un jeu de données constitué de **quelques exemplaires de beaucoup de concepts différents**. Ces concepts sont en l'occurrence des caractères appartenant à 50 alphabètes différents, réels ou fantaisistes, répertoriés dans l'encyclopédie **Omniglot** (<https://www.omniglot.com/>). Une **version manuscrite** de chacun de ces caractères a ensuite été créée par des humains grâce au service de délégation de tâches **Amazon Mechanical Turk** (<https://www.mturk.com/mturk/help?helpPage=overview>).

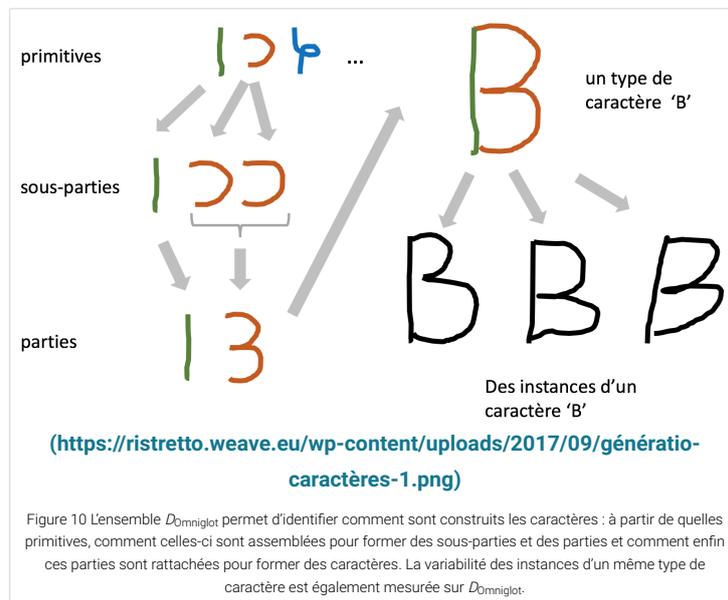


Chaque participant avait pour mission de dessiner quelques caractères avec la plus grande précision possible. Ses mouvements étaient enregistrés dans un petit film associé à chaque caractère à retranscrire. Au final, une liste de 1623 films associés aux images des caractères constitue les données utilisées pour construire le MGC. Dénotons cet ensemble par $D_{Omniglot}$.

3.2. Apprendre à apprendre

Les données $D_{Omniglot}$ ne permettent pas directement d'apprendre à reconnaître les caractères. Elles contiennent en revanche une grande quantité d'information statistique sur le **processus d'écriture** manuscrit lui-même. Ce processus causal d'écriture est modélisé (dans les grandes lignes) de la manière suivante et il est illustré dans la figure 10 :

- Chaque caractère manuscrit est décomposé⁽⁴⁾ en plusieurs **parties**, chaque partie étant d'un seul tenant, tracée sans lever le crayon.
- Chaque partie est rattachée à la précédente selon un certain **mode de rattachement** : « au début », « à la fin », « librement » ou « le long de la partie précédente ».
- Chaque partie est elle-même décomposée en **sous-parties**, chacune d'elles étant définie comme le tracé du crayon entre deux pauses successives du mouvement du crayon.
- Les sous-parties elles-mêmes sont des petites variations de **1250 primitives** identifiées à partir de $D_{Omniglot}$ ⁽⁴⁾.
- Un **type de caractère** est la donnée d'une suite de parties et d'un mode de rattachement entre elles.
- Une **instance de caractère** est fabriquées en ajoutant de petites fluctuations statistiques aux mouvements définis par un type de caractère, l'intensité de ces fluctuations étant elle-même déterminée empiriquement à partir de $D_{Omniglot}$.



Désignons par C une telle structure de parties, de sous-parties et de modes de rattachement qui définit un type de caractère. L'information statistique extraite de D_{omniglot} concernant la structure des caractères est alors encapsulée dans une distribution de probabilité empirique $P(C)$. En clair, si nous tirons au hasard un type C selon la loi $P(C)$ nous obtiendrons un type de caractère inédit, ne figurant pas dans D_{omniglot} , mais construit de manière similaire à ceux présents dans D_{omniglot} .

Désignons par θ une instance d'un type de caractère C . Les fluctuations des différentes réalisations $\theta_1, \dots, \theta_m$ d'un même type de caractère C sont alors décrites par une distribution de probabilité conditionnelle $P(\theta|C)$ dont on peut également évaluer tous les paramètres à partir de D_{omniglot} . Si nous tirons au hasard un exemplaire θ pour un type C correspondant à un caractère connu, un 'B' par exemple, nous obtiendrons une variation raisonnable d'un 'B'.

Le **modèle génératif complet** est donc la distribution conjointe $P(\theta, C) = P(\theta|C) P(C)$ sur les types et les instances.

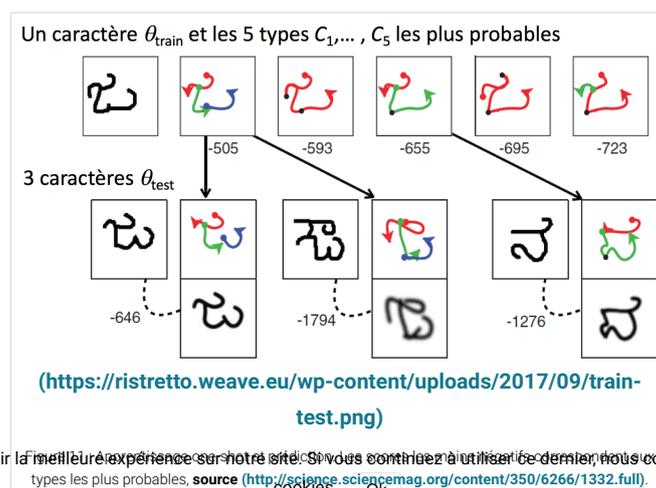
La distribution de probabilité empirique $P(\theta, C)$ obtenue à partir de D_{omniglot} encapsule tout la connaissance sur la manière dont les humains dessinent des caractères alphanumériques.

A ce stade, **le modèle a appris à apprendre**. Voyons maintenant comment apprendre.

3.3. Apprendre en un seul coup***

Apprendre en un seul coup consiste, à partir de l'observation d'un seul caractère θ_{train} , et à déterminer quels sont les types C_i qui ont le plus de chance d'expliquer l'instance θ_{train} , ceci parmi les types d'un alphabeté fixé $\{C_1, C_2, \dots, C_N\}$ (pour notre alphabeté $N = 26$). Schématiquement, nous souhaitons donc savoir quelles sont les types C_i pour lesquels la probabilité à postériori $P(C_i|\theta_{\text{train}})$ est importante. En théorie, notre distribution à priori $P(\theta, C)$ contient toute l'information nécessaire pour effectuer ce calcul par inférence bayésienne. En pratique, il faut cependant recourir à de nombreuses astuces pour que ce calcul puisse être mené approximativement en un temps raisonnable, c'est là l'une des challenges techniques relevé avec succès dans le travail [TEN1].

Les auteurs sélectionnent les 5 types C_1, \dots, C_5 pour lesquels $P(C_i|\theta_{\text{train}})$ est la plus grande comme l'illustre la figure 11.



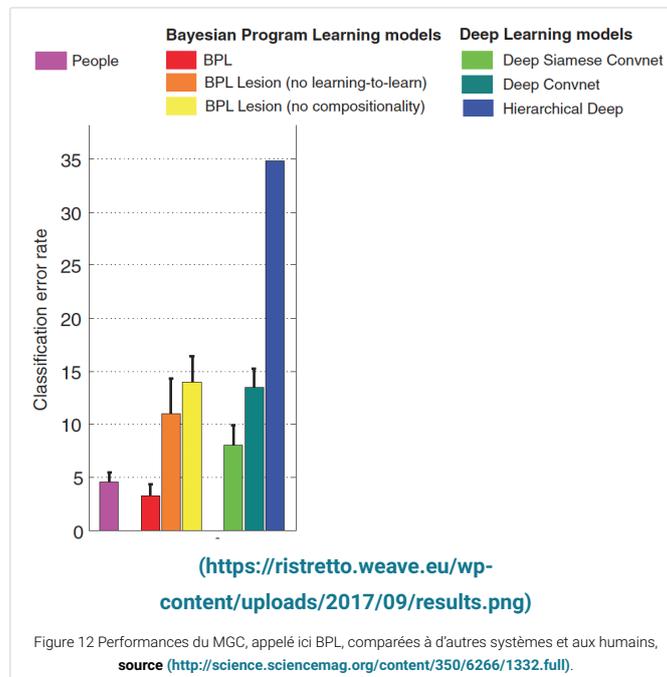
Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies. source (<http://science.sciencemag.org/content/350/6266/1332.full>). cookies. Ok

3.4. Prédire***

Une fois l'apprentissage effectué, on connaît pour chacun des N exemplaires θ_{train} de l'alphabète les 5 types $C_1(\theta_{\text{train}}), \dots, C_5(\theta_{\text{train}})$ qui ont le plus de chance de l'expliquer. En restant à un niveau très schématique un caractère θ_{test} d'un ensemble de test est alors identifié en évaluant les 5 scores $P(\theta_{\text{test}} | C_i(\theta_{\text{train}}))$ pour chacun des θ_{train} de l'alphabète et en retenant celui qui obtient les meilleurs scores.

3.5. Evaluer les performances

Les auteurs de [TEN1] comparent les performances de leur MGC avec les performances humaines et avec celles de plusieurs versions de modèles de Deep Learning spécialement optimisés pour la tâche de reconnaissance de caractères. La figure 12 compare le taux d'erreur de classification de différents systèmes. On constate que le modèle MGC (appelé ici BPL) bat non-seulement les meilleurs réseaux de neurones, mais même les humains !



Plus intéressant encore, les performances de **MGC affectés de différentes lésions** sont également évaluées pour tester l'importance de différents ingrédients du modèle :

- **Lésion de la faculté d'apprendre à apprendre** : la distribution de probabilité à priori $P(\theta, C)$ qui encapsule toute la connaissance du processus d'écriture humaine est remplacée par une distribution ne comportant aucune information, ni sur les primitives, ni sur l'agencement des différentes parties etc...
- **Lésion de la prise en compte de la décomposition en parties** : le modèle est remplacé par un modèle plus simple où chaque caractère est supposé constitué d'une seule partie.

Globalement ces résultats démontrent que les performances du MGC tiennent à deux ingrédients :

1. La **modélisation du processus causal de génération des caractères** au moyen d'une distribution à priori sur les types et les instances de caractères. C'est sur elle que s'appuie l'inférence bayésienne de reconnaissance.
2. Le **fait que le modèle soit hiérarchique** dans le sens où le modèle d'un caractère est lui-même une composition des modèles des parties qui le composent.

4. Un nouvel outil – la programmation probabiliste***



(<https://ristretto.weave.eu/wp-content/uploads/2017/09/dice-1.png>) Comme l'a bien illustré l'exemple précédent, la construction d'un MGC repose d'abord sur la **construction d'une distribution de probabilité à priori complexe** qui rend compte du processus causal à l'origine des observations pour lesquelles on souhaite faire des prédictions. Dans l'exemple de la section précédente il s'agissait de la distribution $P(\theta, C)$ où C représentait un type de caractère et θ une instance de

caractère. Rappelons que l'**apprentissage** et la **prédiction** impliquait de déterminer les types C les plus à même de générer chacun des N exemplaires uniques θ_{train} d'un alphabète. En d'autres termes il s'agissait d'estimer la distribution de probabilité à postériori $P(C | \theta_{\text{train}})$ sur les types C sachant que le caractère θ_{train} a été observé. La **génération** de nouveaux caractères à partir d'un exemplaire de type C impliquait quant à elle de savoir échantillonner θ selon la distribution $P(\theta | C)$.

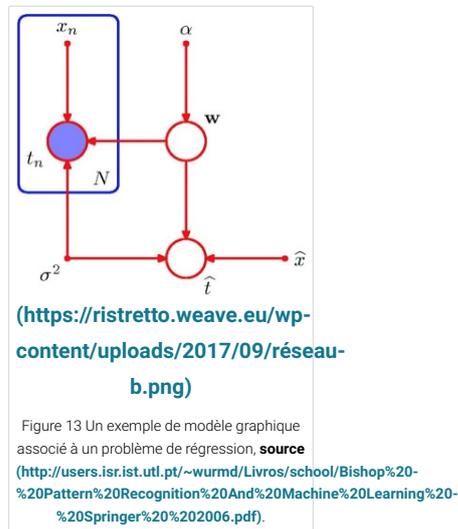
Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des

Cette situation est en fait générale :

cookies.

Pour tout MGC, l'apprentissage, le calcul de prédictions et l'échantillonnage impliquent de savoir calculer efficacement des distributions de probabilités complexes conditionnées de différentes manières sur ce qui est observé.

Les mathématiciens ont depuis longtemps inventé des outils pour mener efficacement ce genre de calculs. L'un des plus connus s'appelle les **réseaux bayésiens**⁽⁵⁾ qui font partie de la classe plus générale des **modèles graphiques**. Nous n'entrerons pas dans les détails de ces modèles. Disons simplement qu'ils permettent de représenter graphiquement la structure des dépendances entre un grand nombre de variables aléatoires et d'exprimer graphiquement des calculs de conditionnement complexes comme ceux qui nous intéressent dans les MGC.



La **programmation probabiliste** (PP) est un outil qui, d'une certaine manière, généralise les calculs que l'on peut effectuer à l'aide des réseaux bayésiens mais qui s'appuie sur un langage de programmation dédié plutôt que des méthodes graphiques. Nous les évoquons brièvement dans cette dernière section car ces langages sont aujourd'hui en plein essor dans l'IA et les sciences cognitives [TEN1, TEN2, EVA].

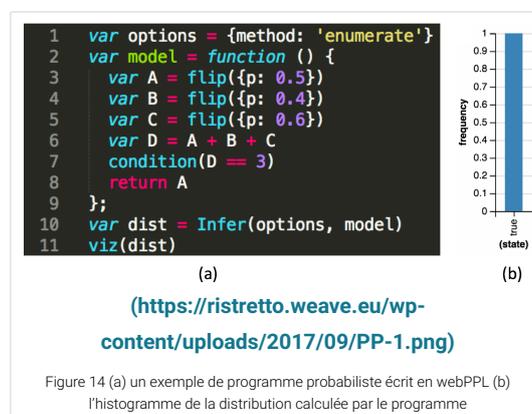
Une autre manière de mettre en perspective la PP est de l'envisager comme un moyen de formaliser un processus de déduction dans des situations où l'information est incertaine ou incomplète, là où la programmation ordinaire, basée sur logique, ne permet de faire que des déductions dans les situations certaines. Alors que la programmation ordinaire manipule des nombres, **la PP manipule donc des distributions de probabilité**.

La PP ajoute deux instructions aux constructions syntaxiques des langages de programmation usuels :

1. La possibilité de **construire des distributions de probabilités** $P(\mathbf{w})$, dépendant de paramètres \mathbf{w} , dont on pourra échantillonner des exemplaires $\mathbf{x} \sim P(\mathbf{w})$.
2. La possibilité conditionner ces distributions de toutes les manières possibles pour en créer de nouvelles : $\mathbf{x} \sim P(\mathbf{w}|C)$.

La possibilité de **conditionner ainsi arbitrairement** un modèle défini par PP permet de poser des questions d'une grande diversité. C'est ce qui en fait un modèle d'une **grande richesse prédictive**.

Voici un exemple de programme probabiliste écrit dans le langage *webPPL* dont la syntaxe étend celle du langage *JavaScript* :



- Ligne 1 : définit des options techniques sur l'implémentation utilisée pour faire le calcul d'inférence
- Ligne 2 : définit une distribution de probabilité sans paramètres appelée model
- Lignes 3-6 : définissent 3 variables aléatoires binaires A, B, C avec différentes probabilités prescrites par les paramètres passés en argument
- Ligne 6 : définit une variable D comme somme des 3 variables A, B et C
- Ligne 7 : conditionne sur la valeur observée de D qui doit valoir 3.

- Ligne 10 : calcule la distribution ainsi conditionnée
- Ligne 11 : visualise la distribution sous forme d'histogramme

Il existe aujourd'hui plusieurs **langages de PP** (<http://probabilistic-programming.org/wiki/Home>) en cours de développement. La principale difficulté à surmonter réside dans le fait qu'il n'existe en général aucune méthode exacte pour calculer une distribution spécifiée par un tel programme, une tâche qu'on appelle le calcul d'inférence. On utilise soit des techniques d'échantillonnage classiques comme **MCMC** (https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Monte-Carlo_par_cha%C3%AEnes_de_Markov) soit, lorsque c'est possible, une compilation du programme en un modèle graphique pour lequel on sait calculer l'inférence explicitement [GOR].

Un autre aspect prometteur de la PP à noter :

La PP offre peut-être une voie alternative et pragmatique à la construction de modèle prédictifs complexes pour des personnes sans connaissance approfondie du machine learning ou des statistiques.

Les applications potentielles de ce type de programmation sont très nombreuses : systèmes de recommandations, diagnostic médical, conception de stratégie de sécurité IT, NLP, élaboration de modèles cognitifs, système anti-spams etc...[GOR] Bien conscients du potentiel de la PP des organismes comme la **DARPA** (<https://www.darpa.mil/program/probabilistic-programming-for-advancing-machine-Learning>) et **Microsoft Research** (<https://www.microsoft.com/en-us/research/publication/probabilistic-programming/>) mènent aujourd'hui une recherche active dans ce domaine.

5. Conclusion et problèmes ouverts



(<https://ristretto.weave.eu/wp-content/uploads/2017/09/questions.png>) L'approche du machine learning par les MGC postule qu'il est extrêmement utile, pour accroître les capacités de généralisations d'un modèle prédictif, de modéliser les processus qui expliquent comment fonctionne le monde physique et social qui nous entoure.

Partant d'un modèle qui prédit les effets de certaines causes, on va chercher à inverser cette relation pour déterminer les causes responsables des effets observés : qu'il s'agisse de propriétés physiques de certains objets ou d'intentions, de croyances et de souhaits de certains individus. Ayant une connaissance de ces causes une IA pourra alors faire des **prédictions d'une richesse qui excèdent largement celles des systèmes prédictifs du ML actuels** qui ne modélisent que des associations entre des variables prédictives et une cible.

Les concepts mathématiques nécessaires pour formaliser cette approche existent depuis longtemps, il s'agit pour l'essentiel de l'**approche bayésienne au machine learning** [BSH].

Les outils informatiques pour la concrétiser sont actuellement en cours de développement, il s'agit des langages de **programmation probabiliste** dont *webPPL* est un exemple avec lequel il est aisé de se familiariser puisqu'il s'exécute dans un navigateur web [TEN2, EVA].

Proposer une liste d'applications de la PP s'avère bien hasardeux car la principale limite aux applications est l'imagination, mentionnons toutefois qu'il existe d'ores et déjà des applications dans des domaines aussi variés que le diagnostic médical, les systèmes de recommandation, l'évaluation de joueurs en ligne ou la sécurité IT [GOR]. **A court terme** la programmation probabiliste offre un nouveau moyen de modéliser les systèmes prédictifs sans qu'il soit nécessaire de connaître les subtilités du machine learning. **A long terme** elle permet d'envisager le développement d'agents intelligents dotés d'intuitions physiques et sociales qui leur permettront d'être très utiles car elles comprendront ce que signifie « aider les humains ».

L'exemple de la reconnaissance de caractères one-shot [TEN1] démontre de manière quantifiable la pertinence de ce nouveau paradigme dans un contexte fonctionnel limité même si, pour l'instant, la quantité de travail humain nécessaire pour construire le savoir a priori reste très importante.

Restent désormais à résoudre les problèmes vraiment difficiles :

- Existe-t-il des principes génériques pour construire de ces modèles génératifs ?
- Comment implémenter un **principe hiérarchique** qui permette de les construire par encapsulation successives de modèles toujours plus abstraits ?
- Quelles sont les **connaissances à priori** que l'évolution biologique a encré dans nos cerveaux et qui nous permettent de construire tout au long de la vie les modèles causals qui nous sont nécessaires pour évoluer dans le monde physique et social ?
- Peut-on faire un copier-coller de ces connaissances à priori ou faut-il, au contraire, **simuler le processus d'évolution** lui-même ?
- Comment initier et maintenir un processus d'**apprentissage continu** ?

Si l'idée de départ des MGC est en définitive très simple, l'ambition elle est grande : construire à long terme des machines qui apprennent comme des enfants.

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des

cookies.

Notes

1. Que l'on doit à la concomitance de trois facteurs : l'augmentation de la puissance de calcul des processeurs (GPU), la disponibilité de volumineux jeux de données et la découverte d'un certain nombre d'astuces techniques qui permettent un entraînement efficace de ces algorithmes.
2. Le Reinforcement Learning étant, rappelons-le, la partie du ML qui consiste à développer des agents qui apprennent un comportement sous l'effet d'un feedback de l'environnement dans lequel ils évoluent, typiquement en minimisant une certaine fonction de coût.
3. Cette distribution $P_{\text{post\u00e9riori}}$ possède un maximum pour certains paramètres w_{MAP} (MAP=maximum a posteriori probability) que le peut estimer au moyen d'une descente de gradient par exemple. La prise en compte de la distribution à priori $P_{\text{prior}}(w)$ dans ce calcul correspond en fait à rajouter un terme de régularisation au calcul du maximum de vraisemblance (voir p.ex. p217 [BSH]).
4. Selon un processus statistique complexe que nous ne pouvons décrire ici, voir [TEN1].
5. Consulter le chapitre 8 de [BSH] pour une excellente introduction à ce sujet.

Sources

- [TEN1] **Human-level concept learning through probabilistic program induction** (<http://science.sciencemag.org/content/350/6266/1332>), Brenden M. Lake, Ruslan Salakhutdinov, Joshua B. Tenenbaum, Science – 11 Dec 2015: Vol. 350, Issue 6266, pp. 1332-1338.
- [TEN2] **Probabilistic Models of Cognition 2nd edition** (<http://probmods.org/>), Noah D. Goodman & Joshua B. Tenenbaum.
- [EVA] **Modeling Agents with Probabilistic Programs** (<http://agentmodels.org/>), Owain Evans, Andreas Stuhlm\u00fcller, John Salvatier, et Daniel F\u00e9lan.
- [BSH] **Pattern Recognition and Machine Learning** (<http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop - Pattern Recognition And Machine Learning - Springer 2006.pdf>), Christopher Bishop, Springer 2006 * [BPB] Bayesian Programming, Pierre Bessière Juan-Manuel Ahuactzin Kamel Mekhnacha Emmanuel Mazer, CRC Press, 2013.
- [GOR] **Probabilistic Programming** (<https://www.microsoft.com/en-us/research/people/adg/>), Andy Gordon, Thomas A. Henzinger, Aditya Nori, Sriram Rajamani
- [SPP] **The State of Probabilistic Programming** (<https://moalquraishi.wordpress.com/2015/03/29/the-state-of-probabilistic-programming/>), Mohammed Al Quraishi.
- [LEM] **Un nouveau chantier pour la sécurité IT : le Machine Learning** (<https://ristretto.weave.eu/2017/08/28/nouveau-chantier-securite-it-machine-learning/>), P. Lemberger, Ristretto by Weave.

AI (<https://ristretto.weave.eu/tag/ai/>) Data Science (<https://ristretto.weave.eu/tag/data-science/>)

deep learning (<https://ristretto.weave.eu/tag/deep-learning/>) machine learning (<https://ristretto.weave.eu/tag/machine-learning/>)

Partager cette publication



J'aime 0

Tweeter

G+



PIRMIN LEMBERGER ([HTTPS://RISTRETTO.WEAVE.EU/AUTHOR/PIRMIN-LEMBERGER/](https://ristretto.weave.eu/author/pirmin-leMBERGER/))

Directeur scientifique

+33677262772

pirmin.leMBERGER@weave.eu



(<https://twitter.com/pirminleMBERGER>)



(<https://fr.linkedin.com/in/pirminleMBERGER>)

Newsletter

Recevez notre newsletter bi-mensuelle

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies.

J'accepte
 Refuser

Commenter

Connecté en tant que **Pirmin Lemberger**. [Se déconnecter ? \(https://ristretto.weave.eu/wp-login.php?action=logout&redirect_to=https%3A%2F%2Fristretto.weave.eu%2F%3Fp%3D23043&_wpnonce=ce4e4bc919\)](https://ristretto.weave.eu/wp-login.php?action=logout&redirect_to=https%3A%2F%2Fristretto.weave.eu%2F%3Fp%3D23043&_wpnonce=ce4e4bc919)

Répondre



28 août 2017

Digital & Technology Exploration & Innovation

UN NOUVEAU CHANTIER POUR LA SÉCURITÉ IT : LE MACHINE LEARNING [\(HTTPS://RISTRETTO.WEAVE.EU/2017/08/28/NOUVEAU-CHANTIER-SECURITE-IT-MACHINE-LEARNING/\)](https://ristretto.weave.eu/2017/08/28/nouveau-chantier-securite-it-machine-learning/)

Téléchargez le livre blanc : « IA, où en sommes nous ? » Résumé L'objectif de cet article est d'expliquer en termes non-techniques quelques-uns des enjeux de sécurité posé par le Machine Learning (ML) lorsqu'il est appliqué à des systèmes critiques. Certains de ces problèmes pourraient être des symptômes [...]

LIRE LA SUITE



Pirmin Lemberger (<https://ristretto.weave.eu/author/pirmin-lemberger/>) (<https://ristretto.weave.eu/2017/08/28/nouveau-chantier-securite-it-machine-learning/#comments>)

[\(https://ristretto.weave.eu/2017/04/10/chatbots-entre-reves-realite/\)](https://ristretto.weave.eu/2017/04/10/chatbots-entre-reves-realite/)



10 avril 2017

Digital & Technology Exploration & Innovation

OÙ EN SOMMES-NOUS?

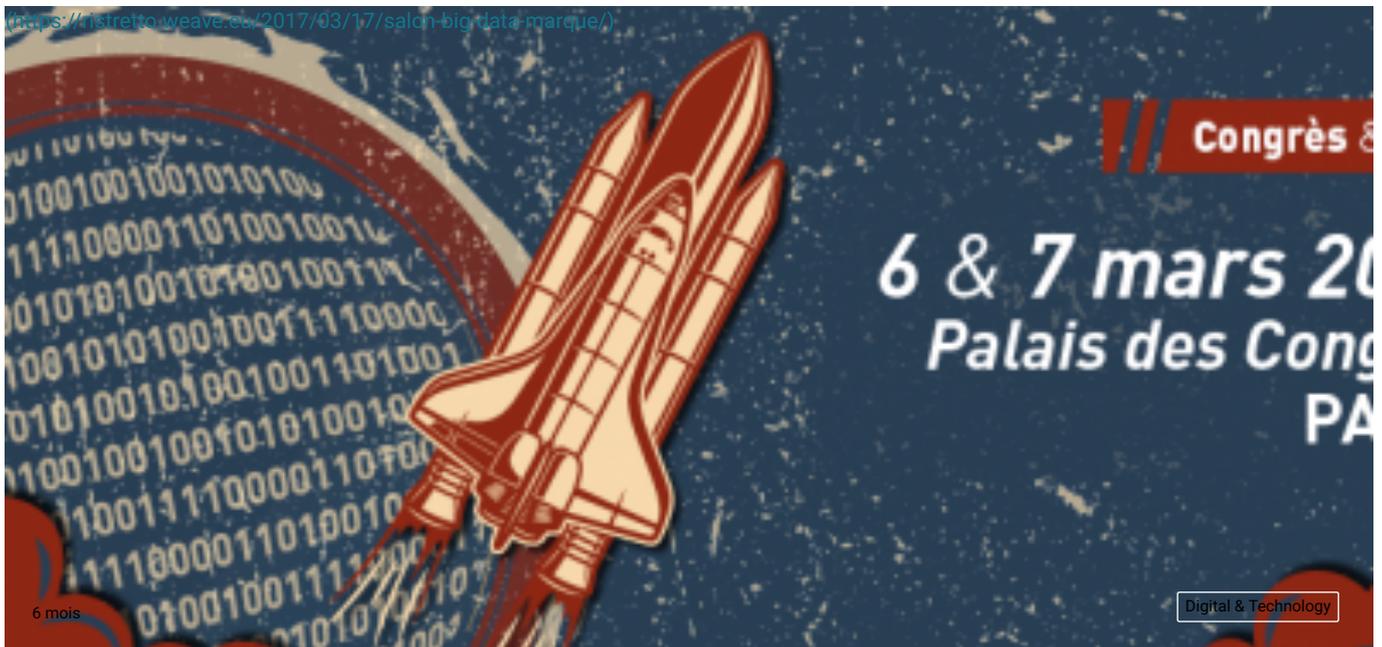
CHATBOTS – ENTRE RÊVES ET RÉALITÉ ([HTTPS://RISTRETTO.WEAVE.EU/2017/04/10/CHATBOTS-ENTRE-REVES-REALITE/](https://ristretto.weave.eu/2017/04/10/chatbots-entre-reves-realite/))

Téléchargez le livre blanc : « IA, où en sommes nous ? » Pourquoi aujourd’hui ? La question est récurrente à chaque apparition d’une nouvelle technologie : pourquoi un tel engouement aujourd’hui ? S’agissant des chatbots qui défraient la chronique depuis quelques mois, la question apparaît d’autant plus pertinente que le [...]

LIRE LA SUITE



Pirmin Lemberger (<https://ristretto.weave.eu/author/pirmin-lemberger/>) (<https://ristretto.weave.eu/2017/04/10/chatbots-entre-reves-realite/#comments>)



6 mois

Digital & Technology

SALON DU BIG DATA : CE QUI M'A MARQUÉ ([HTTPS://RISTRETTO.WEAVE.EU/2017/03/17/SALON-BIG-DATA-MARQUE/](https://ristretto.weave.eu/2017/03/17/salon-big-data-marque/))

La grand-messe des entreprises du Big Data se tenait la semaine dernière, les 7 et 8 Mars, au palais des congrès à Paris. En parallèle des exposants, des conférences et retours d’expérience autour du sujet avaient lieu. Je vous partage mon retour d’expérience du salon et mes convictions pour vous aider à mieux piloter les différentes [...]

LIRE LA SUITE



Jacques Peeters (<https://ristretto.weave.eu/author/jpeeters/>) (<https://ristretto.weave.eu/2017/03/17/salon-big-data-marque/#comments>)

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies.



IA OU EN SOMMES NOUS ? QUELLE DÉFINITION ? ([HTTPS://RISTRETTO.WEAVE.EU/2017/03/16/IA-SOMME-1/](https://ristretto.weave.eu/2017/03/16/ia-somme-1/))

Cet article est le premier chapitre d'un livre blanc consacré à faire le point sur l'intelligence artificielle. Voici la liste des sujets qui seront abordés : L'IA en quête de définition Pourquoi une définition de l'IA est délicate Quels objectifs pour l'IA ? Comment caractériser une IA ? Les progrès [...]

LIRE LA SUITE

 Pirmin Lemberger (<https://ristretto.weave.eu/author/pirmin-leMBERGER/>) (<https://ristretto.weave.eu/2017/03/16/ia-somme-1/#comments>)

VOIR PLUS

Articles similaires



RESTAURATION COLLECTIVE : OÙ EN-EST-ON ? ([HTTPS://RISTRETTO.WEAVE.EU/2017/09/20/RESTAURATION-COLLECTIVE-MUTATION-MARCHE/](https://ristretto.weave.eu/2017/09/20/restauration-collective-mutation-marche/))

La restauration collective se prépare à un changement de paradigme. Le consommateur évolue et les opportunités ouvertes par la digitalisation doivent être correctement exploitées pour permettre aux acteurs de se différencier et de faire évoluer leurs produits et services. « La pause cantine est tombée à moins de 25 minutes contre plus [...]

LIRE LA SUITE

 Agathe Cebron (<https://ristretto.weave.eu/author/agathe-cebron/>)

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies.



ENERGY OBSERVER : UNE MISSION POUR LE FUTUR ([HTTPS://RISTRETTO.WEAVE.EU/2017/09/19/ENERGY-OBSERVER-MISSION-FUTUR/](https://ristretto.weave.eu/2017/09/19/ENERGY-OBSERVER-MISSION-FUTUR/))

Il devient difficile d'échapper à la présentation du catamaran Energy Observer. Cet ONNI (Objet Naviguant Non Identifié) high-tech couvert de panneaux solaires étonne ses contemplateurs depuis qu'il a été inauguré au mois de juillet. Menée par Victorien Erussard et Jérôme Delafosse, cette mission est hors norme. Un projet hors-norme [...]

LIRE LA SUITE

ristretto.weave.eu/2017/09/19/ENERGY-OBSERVER-MISSION-FUTUR/ (<https://ristretto.weave.eu/author/benjamin-morvan/>)

(<https://ristretto.weave.eu/2017/09/15/LE-GLYPHOSATE-UN-NOUVEAU-SCANDALE-POUR-MONSANTO/>)



LE GLYPHOSATE, UN NOUVEAU SCANDALE POUR MONSANTO ? ([HTTPS://RISTRETTO.WEAVE.EU/2017/09/15/GLYPHOSATE-NOUVEAU-SCANDALE-MONSANTO/](https://ristretto.weave.eu/2017/09/15/GLYPHOSATE-NOUVEAU-SCANDALE-MONSANTO/))

Malgré des condamnations et des scandales à répétition (PCB, agent orange, dioxine, OGM, aspartame, hormones de croissance, ...), rien n'a arrêté jusqu'ici la croissance du géant américain des phytosanitaires (plus de 15 Mds d'euros de chiffre d'affaires) ! Aujourd'hui, le nouveau scandale porte sur le glyphosate ! Cette [...]

LIRE LA SUITE

ristretto.weave.eu/2017/09/15/LE-GLYPHOSATE-UN-NOUVEAU-SCANDALE-POUR-MONSANTO/ (https://ristretto.weave.eu/author/stephane_gregoire/)

VOIR PLUS

Nos tweets



News_Ristretto (http://www.twitter.com/News_Ristretto)

Un concentré d'actualités offert par les weavers, consultants de @weaveconseil #digital #agroalimentaire #RSE #innovation #RH #relationclient #consulting
Paris

4 mois

Les réseaux de femmes : un levier pour dépasser le « plafond de verre » ? #richesseshumaines <https://t.co/kUgi08UFPq> <https://t.co/k1yhd7WaOD>
Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des

[Retweet \(http://www.twitter.com/News_Ristretto/status/8.644100941628994175\)](http://www.twitter.com/News_Ristretto/status/8.644100941628994175)

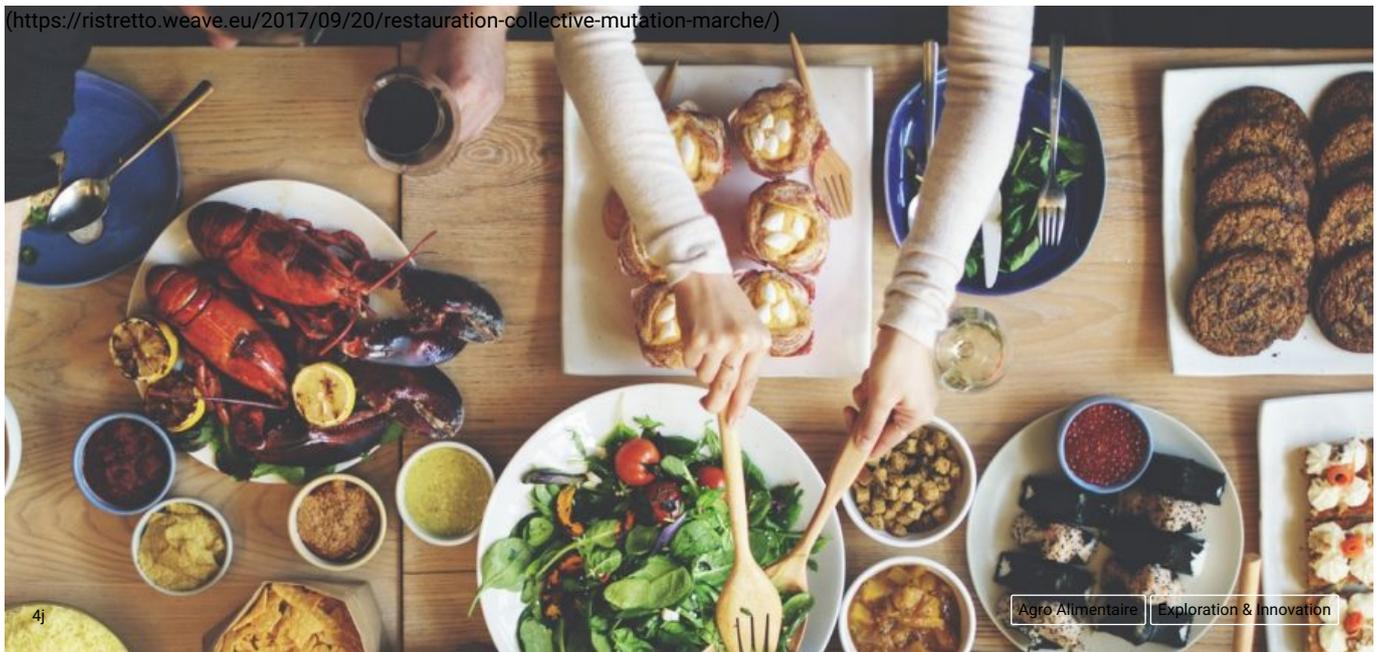
4 mois

Prélèvement à la source – Les assureurs au milieu du gué #banque&assurance <https://t.co/og6m4BkLok> <https://t.co/fCWmgBNXW4>↻ Retweet (http://www.twitter.com/News_Ristretto/status/8.6410786037194E+17)

5 mois

« Assur-News » : sport, Insurtechs, expérience client #banque&assurance <https://t.co/vOFeA50NZI> <https://t.co/IMQGOBDH2f>↻ Retweet (http://www.twitter.com/News_Ristretto/status/8.6293756642483E+17)

Dernier post



RESTAURATION COLLECTIVE : OÙ EN-EST-ON ? ([HTTPS://RISTRETTO.WEAVE.EU/2017/09/20/RESTAURATION-COLLECTIVE-MUTATION-MARCHE/](https://ristretto.weave.eu/2017/09/20/restauration-collective-mutation-marche/))

La restauration collective se prépare à un changement de paradigme. Le consommateur évolue et les opportunités ouvertes par la digitalisation doivent être correctement exploitées pour permettre aux acteurs de se différencier et de faire évoluer leurs produits et services.
« La pause cantine est tombée à moins de 25 minutes contre plus [...] »

LIRE LA SUITE



Agathe Cebron de Lisle (<https://ristretto.weave.eu/2017/09/20/restauration-collective-mutation-marche/#comments>)

Made by



Conseil en stratégie augmentée ®

Découvrez au quotidien un angle de vue décalé pour vous aider à penser différemment.

weave.eu (<https://weave.eu>)

Paris

37, rue du Rocher
75008 Paris

Nous utilisons des cookies pour vous garantir la meilleure expérience sur notre site. Si vous continuez à utiliser ce dernier, nous considérerons que vous acceptez l'utilisation des cookies.