A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH and WALTER H. PITTS

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

INTRODUCTION

THEORETICAL neurophysiology rests on certain cardinal assumptions. The nervous system is a net of neurons, each having a soma and an axon. Their adjunctions, or synapses, are always between the axon of one neuron and the soma of another. At any instant a neuron has some threshold, which excitation must exceed to initiate an impulse. This, except for the fact and the time of its occurrence, is determined by the neuron, not by the excitation. From the point of excitation the impulse is propagated to all parts of the neuron. The velocity along the axon varies directly with its diameter, from less than one meter per second in thin axons, which are usually short, to more than 150 meters per second in thick axons, which are usually long. The time for axonal conduction is consequently of little importance in determining the time

Scanned by CamScanner

of arrival of impulses at points unequally remote from the same source. Excitation across synapses occurs predominantly from axonal terminations to somata. It is still a moot point whether this depends upon irreciprocity of individual synapses or merely upon prevalent anatomical configurations. To suppose the latter requires no hypothesis *ad hoc* and explains known exceptions, but any assumption as to cause is compatible with the calculus to come. No case is known in which excitation through a single synapse has elicited a nervous impulse in any neuron, whereas any neuron may be excited by impulses arriving at a sufficient number of neighboring synapses within the period of latent addition, which lasts less than one quarter of a millisecond. Observed temporal summation of impulses at greater intervals is impossible for single neurons and empirically depends upon structural properties of the net. Between the arrival of impulses upon a neuron and its own propagated impulse there is a synaptic delay of more than half a millisecond. During the first part of the nervous impulse the neuron is absolutely refractory to any stimulation. Thereafter its excitability returns rapidly, in some cases reaching a value above normal from which it sinks again to a subnormal value, whence it returns slowly to normal. Frequent activity augments this subnormality. Such specificity as is possessed by nervous impulses depends solely upon their time and place and not on any other specificity of nervous energies. Of late only inhibition has been seriously adduced to contravene this thesis. Inhibition is the termination or prevention of the activity of one group of neurons by concurrent or antecedent activity of a second group. Until recently this could be explained on the supposition that previous activity of neurons of the second group might so raise the thresholds of internuncial neurons that they could no longer be excited by neurons of the first group, whereas the impulses of the first group must sum with the impulses of these internuncials to excite the now inhibited neurons. Today, some inhibitions have been shown to consume less than one millisecond. This excludes internuncials and requires synapses through which impulses inhibit that neuron which is being stimulated by impulses through other synapses. As yet experiment has not shown whether the refractoriness is relative or absolute. We will assume the latter and demonstrate

Scanned by CamScanner

that the difference is immaterial to our argument. Either variety of refractoriness can be accounted for in either of two ways. The "inhibitory synapse" may be of such a kind as to produce a substance which raises the threshold of the neuron, or it may be so placed that the local disturbance produced by its excitation opposes the alteration induced by the otherwise excitatory synapses. Inasmuch as position is already known to have such effects in the case of electrical stimulation, the first hypothesis is to be excluded unless and until it be substantiated, for the second involves no new hypothesis. We have, then, two explanations of inhibition based on the same general premises, differing only in the assumed nervous nets and, consequently, in the time required for inhibition. Hereafter we shall refer to such nervous nets as *equivalent in the extended sense*. Since we are concerned with properties of nets which are invariant under equivalence, we may make the physical assumptions which are most convenient for the calculus.

Many years ago one of us, by considerations impertinent to this argument, was led to conceive of the response of any neuron as factually equivalent to a proposition which proposed its adequate stimulus. He therefore attempted to record the behavior of complicated nets in the notation of the symbolic logic of propositions. The "all-or-none" law of nervous activity is sufficient to insure that the activity of any neuron may be represented as a proposition. Physiological relations existing among nervous activities correspond, of course, to relations among the propositions; and the utility of the representation depends upon the identity of these relations with those of the logic of propositions. To each reaction of any neuron there is a corresponding assertion of a simple proposition. This, in turn, implies either some other simple proposition or the disjunction or the conjunction, with or without negation, of similar propositions, according to the configuration of the synapses upon and the threshold of the neuron in question. Two difficulties appeared. The first concerns facilitation and extinction, in which antecedent activity temporarily alters responsiveness to subsequent stimulation of one and the same part of the net. The second concerns learning, in which activities concurrent at some previous time have altered the net permanently, so that a stimulus which would previously have been inadequate is now

Scanned by CamScanner

adequate. But for nets undergoing both alterations, we can substitute equivalent fictitious nets composed of neurons whose connections and thresholds are unaltered. But one point must be made clear: neither of us conceives the formal equivalence to be a factual explanation. *Per contra!*—we regard facilitation and extinction as dependent upon continuous changes in threshold related to electrical and chemical variables, such as after-potentials and ionic concentrations; and learning as an enduring change which can survive sleep, anaesthesia, convulsions and coma. The importance of the formal equivalence lies in this: that the alterations actually underlying facilitation, extinction and learning in no way affect the conclusions which follow from the formal treatment of the activity of nervous nets, and the relations of the corresponding propositions remain those of the logic of propositions.

The nervous system contains many circular paths, whose activity so regenerates the excitation of any participant neuron that reference to time past becomes indefinite, although it still implies that afferent activity has realized one of a certain class of configurations over time. Precise specification of these implications by means of recursive functions, and determination of those that can be embodied in the activity of nervous nets, completes the theory.

THE THEORY: NETS WITHOUT CIRCLES

We shall make the following physical assumptions for our calculus.

1. The activity of the neuron is an "all-or-none" process.

2. A certain fixed number of synapses must be excited within the period of latent addition in order to excite a neuron at any time, and this number is independent of previous activity and position on the neuron.

3. The only significant delay within the nervous system is synaptic delay.

4. The activity of any inhibitory synapse absolutely prevents excitation of the neuron at that time.

5. The structure of the net does not change with time.

Scanned by CamScanner

Université Paris Descartes
IUT, Département Informatique
Communication

Semaine du 11/03/2019

Questions de réflexion sur le chapitre* : Ce qui manque encore à l'intelligence artificielle

* **Dehaene, S.** (2018) « Ce qui manque encore à l'intelligence artificielle » (chap. 2), dans *Apprendre ! Les talents du cerveau, le défi des machines*. Odile Jacob, p 68 – 76.

En vous appuyant sur le texte, répondez aux question suivantes :

1) Donner trois caractéristiques de nos opérations mentales qui demeurent pour l'instant propres à l'homme (il ne s'agit pas de citer des fonctions cognitives comme l'apprentissage de concepts abstraits mais de préciser la nature de nos opérations mentales).

2) Les réseaux **profonds** ne sont pas tout à fait bien nommés, si on considère la nature de leur apprentissage. Expliquez.
Indice *si vous en avez besoin* : reconnaissance d'un même concept sous des *apparences* très variables.

3) L'homme apprend de ses congénères. Quel caractère de *l'apprentissage* des réseaux de neurones rend difficile la communication de ce savoir machine ?

4) L'homme, contrairement à la machine n'aurait pas un unique système d'apprentissage mais deux. Le premier niveau filtre les entrées (en attribuant des poids) et apprends les meilleures combinaisons. Que fait le niveau de hiérarchie supérieur ?

Scanned by CamScanner

CHAPITRE 2

Pourquoi notre cerveau apprend mieux que les machines

Les succès récents de l'intelligence artificielle peuvent laisser croire que nous avons enfin compris comment imiter l'apprentissage et l'intelligence de l'espèce humaine dans des machines – au point que, selon certains prophètes autoproclamés, les machines seraient sur le point de nous dépasser. Rien n'est plus faux. En fait, la plupart des chercheurs en sciences cognitives, même s'ils admirent les progrès récents des réseaux de neurones artificiels, savent très bien que ces machines demeurent limitées. Dans un article récent, j'ai argumenté que les réseaux de neurones conventionnels correspondent étroitement aux opérations que notre cerveau réalise inconsciemment, en deux dixièmes de seconde, lorsqu'il perçoit une image : il la reconnaît, la catégorise et accède à son sens[13]. Cependant, notre cerveau, lui, va beaucoup plus loin : il est capable de l'explorer consciemment, avec attention, point par point, pendant plusieurs secondes. Il formule des représentations symboliques, des théories explicites du monde que nous pouvons partager avec d'autres personnes par le biais du langage.

Scanned by CamScanner

bonques, demeurent pour l'instant l'apanage de notre espèce. Les réseaux de neurones actuels les modélisent mal, même si, chaque année, des progrès sont faits dans la compréhension du langage, la traduction automatique ou le raisonnement logique. C'est un reproche que l'on fait souvent aux réseaux de neurones artificiels : ils essaient de tout apprendre au même niveau, comme si chaque problème revenait à une question de classification automatique. Pour celui qui ne possède qu'un marteau, tout ressemble à un clou ! Notre cerveau, lui, est bien plus flexible. Il parvient très vite à hiérarchiser les informations et, lorsque c'est possible, à en extraire des principes généraux, logiques, explicites.

Ce qui manque encore à l'intelligence artificielle

Il est intéressant d'essayer de préciser ce qui manque encore à l'intelligence artificielle, car c'est une manière de définir, le plus précisément possible, ce qu'il y a d'unique dans notre propre capacité d'apprentissage. Voici une petite liste, sans doute incomplète, de fonctions que même un très jeune enfant possède, et qui font échouer lamentablement la plupart des réseaux actuels :

• *L'apprentissage de concepts abstraits.* La plupart des réseaux de neurones contemporains ne modélisent convenablement que la toute première passe de traitement de l'information pendant laquelle, en moins d'un cinquième de seconde, les aires visuelles analysent une image. Ces algorithmes connexionnistes sont loin d'être aussi profonds qu'on le dit. Selon l'un de leurs inventeurs, Joshua Bengio, en réalité, « les réseaux dits "profonds" ont tendance à apprendre des régularités statistiques superficielles dans

niveau ». Pour reconnaître la présence d'un objet, ils s'appuient sur des éléments anecdotiques de l'image, tels que la couleur ou la forme. Changez ces détails, et leurs performances s'effondrent : ils sont absolument incapables de reconnaître ce qui fait l'essence d'un objet et de concevoir qu'une chaise reste une chaise même si elle est faite de verre, d'un seul pied de métal plié ou de plastique gonflable. Cette propension à ne faire attention qu'à la surface des choses les rend susceptibles d'erreurs massives. Il existe toute une littérature sur la manière de tromper un réseau de neurones : prenez une banane, modifiez-lui quelques pixels ou collez-lui un autocollant bien particulier, et le réseau de neurones la prendra pour un grille-pain !

Il est vrai que, si l'on flashe une image à une personne pendant une fraction de seconde, elle fait parfois le même genre d'erreur que la machine, et peut confondre par exemple un chien et un chat[15]. Cependant, dès qu'on lui laisse un peu plus de temps, le cerveau humain ne s'y trompe pas. Contrairement à la machine, il s'interroge, réanalyse, porte son attention sur tel ou tel aspect de l'image qui ne correspond pas à sa première impression. Cette seconde analyse, consciente, intelligente, fait appel à nos capacités de raisonnement et d'abstraction. Les réseaux de neurones négligent un point essentiel : apprendre, c'est se former un modèle abstrait du monde, pas juste un filtre de reconnaissance de formes. En apprenant à lire, par exemple, nous avons acquis un concept abstrait de chaque lettre de l'alphabet, qui nous permet de la tracer aussi bien que de la reconnaître sous tous ses déguisements :

Douglas Hofstadter, informaticien et cogniticien, a dit un jour que le vrai challenge pour l'intelligence artificielle

Scanned by CamScanner

consistait à reconnaître la lettre A… Boutade, certes, mais boutade profonde. L'intelligence abstraite que les humains déploient même dans cette situation triviale est à l'origine d'un amusant objet de la vie quotidienne : le CAPTCHA, cette petite chaîne de lettres que certains sites Internet vous demandent de reconnaître pour prouver que vous êtes un être humain, et non une machine.

Pendant des années, les CAPTCHA ont résisté aux machines. Mais l'informatique évolue vite : en 2017, un système artificiel est parvenu à reconnaître les CAPTCHA à peu près aussi bien qu'un humain[16]. Sans surprise, cet algorithme imite bon nombre d'aspects de notre cerveau. Véritable tour de force, il extrait le graphe de chaque lettre, l'essence d'un A, et il utilise toutes les ressources du raisonnement statistique pour vérifier à quel point cette idée abstraite s'applique à l'image actuelle. Cependant, cet algorithme informatique sophistiqué ne s'applique qu'aux CAPTCHA. Notre cerveau, lui, applique cette faculté d'abstraction à tous les aspects de notre vie quotidienne.

• *La vitesse d'apprentissage.* Tout le monde s'accorde à dire que les réseaux de neurones actuels apprennent bien trop lentement : il leur faut des milliers, des millions, voire des milliards de données pour acquérir l'intuition d'un domaine. Cette lenteur, nous en avons des preuves expérimentales. Il faut pas moins de 900 heures de jeu pour que le réseau de neurones conçu par DeepMind atteigne un niveau raisonnable sur une console Atari – alors qu'un être humain atteint le même niveau en 2 heures[17] !

Autre exemple : l'apprentissage du langage. Le psycholinguiste Emmanuel Dupoux estime que, dans la plupart des familles françaises, un enfant entend environ 500 à 1 000 heures de parole par an, ce qui lui suffit à apprendre sa langue maternelle. Encore s'agit-il certainement d'une

60 heures de parole par an, ce qui ne les empêche pas de devenir d'excellents locuteurs de la langue chimane. En comparaison, les meilleurs systèmes informatiques actuels d'Apple, de Baidu ou de Google nécessitent entre vingt et mille fois plus de données. Dans le domaine de l'apprentissage, l'efficacité du cerveau humain reste inégalée : « *Machines are data hungry, but humans are data efficient.* » L'apprentissage, dans notre espèce, sait tirer le meilleur parti de la moindre donnée.

• *L'apprentissage social.* Notre espèce est la seule à pratiquer le partage d'informations : nous apprenons énormément d'informations des autres êtres humains, par imitation ou par le biais du langage. Cette capacité est, pour l'instant, hors de portée des réseaux de neurones actuels. Chez eux, la connaissance est cryptée, diluée dans les valeurs de centaines de millions de poids synaptiques, où elle demeure implicite. Il est impossible de l'extraire pour la communiquer à d'autres. L'extraordinaire efficacité avec laquelle nous parvenons, en quelques mots, à partager notre savoir avec d'autres (« Pour la boulangerie, prenez à droite dans la petite rue derrière l'église ») reste inégalée dans le monde animal comme en informatique.

• *L'apprentissage en un essai.* Version extrême de cette efficacité : il nous arrive d'apprendre en un seul essai. Si j'utilise un nouveau verbe, disons « daxer », ne fût-ce qu'une seule fois, cela vous suffit à le connaître. Entendons-nous bien : certains réseaux de neurones aussi sont capables de stocker un épisode spécifique. Mais ce que les machines ne font pas encore bien, et que le cerveau humain réussit à merveille, c'est d'intégrer cette information au sein d'un réseau de connaissances. Instantanément, vous parvenez non seulement à mémoriser le verbe « daxer », mais également à le conjuguer et à l'insérer dans d'autres phrases (Vous

Scanned by CamScanner

d'entendre une série de sons (bip bip bip boup), sans en théoriser immédiatement la structure abstraite (trois sons identiques suivis d'un son différent). Placés dans la même situation, les singes détectent trois sons, entendent que le dernier est différent, mais ne semblent pas intégrer ces connaissances dans une formule unique[22]. Il faut des dizaines de milliers d'essais pour qu'un singe apprenne à renverser l'ordre d'une séquence (passer de ABCD à DCBA), alors que cinq essais suffisent à n'importe quel gamin de 4 ans[23]. Même un bébé de quelques mois est capable d'inférer des règles abstraites et systématiques – une capacité qui échappe totalement à la fois aux réseaux de neurones conventionnels et aux autres espèces de primates.

• *La composition des connaissances.* Une fois que j'ai appris, disons, à additionner deux chiffres, cette faculté fait partie de mon répertoire de talents : elle devient immédiatement disponible pour l'ensemble de mes facultés mentales. Je peux l'utiliser dans des dizaines de contextes distincts, par exemple au restaurant ou sur ma feuille d'impôt. Surtout, je peux la combiner avec d'autres facultés – je n'ai aucune difficulté, par exemple, à suivre un algorithme qui me demande de prendre un nombre, de lui ajouter 2 et de décider si le résultat est plus grand ou plus petit que 5. Le cerveau humain semble disposer d'une sorte d'ordinateur interne, une véritable machine de Turing capable d'enchaîner les opérations dans un ordre arbitraire[24].

Il est étonnant de voir que les réseaux de neurones actuels n'ont pas encore cette flexibilité. Ce qu'ils apprennent reste confiné dans des connexions cachées, inaccessibles, dispersées et très difficiles à réutiliser pour d'autres tâches. La capacité de *composer* les connaissances, c'est-à-dire de les recombiner pour résoudre des problèmes nouveaux, leur échappe. L'intelligence artificielle actuelle ne résout que

borné, incapable de généraliser ses talents à tout autre jeu un tant soit peu différent (y compris le jeu de go sur un échiquier 15 × 15 plutôt que 19 × 19). Pour notre cerveau, par contre, apprendre, c'est expliciter les connaissances en sorte que l'on puisse les recombiner avec d'autres.

Là encore, nous avons affaire à un aspect singulier du cerveau humain, lié au langage et difficile à reproduire dans une machine. René Descartes l'avait constaté dès 1637 dans le célèbre *Discours de la méthode* :

> « S'il y avait [des machines] qui eussent la ressemblance de nos corps, et imitassent autant nos actions que moralement il serait possible, [...] nous aurions toujours deux moyens très certains pour reconnaître qu'elles ne seraient point pour cela de vrais hommes. Le premier est que jamais elles ne pourraient user de paroles ou d'autres signes en les composant, comme nous faisons pour déclarer aux autres nos pensées. Car on peut bien concevoir qu'une machine soit tellement faite qu'elle profère des paroles [...], mais non pas qu'elle les arrange diversement pour répondre au sens de tout ce qui se dira en sa présence, ainsi que les hommes les plus hébétés peuvent faire. Et le second est que, bien qu'elles fissent plusieurs choses aussi bien ou peut-être mieux qu'aucun de nous, elles manqueraient infailliblement en quelques autres, par lesquelles on découvrirait qu'elles n'agiraient pas par connaissance, mais seulement par la disposition de leurs organes. Car, au lieu que la raison est un instrument universel qui peut servir en toutes sortes de rencontres, ces organes ont besoin de quelque particulière disposition pour chaque action particulière. »

La raison, instrument universel... Les capacités que recense Descartes pointent vers un deuxième système d'apprentissage, hiérarchiquement plus élevé que le précédent, et qui repose sur des règles et des symboles. Dans ses premières étapes, notre système visuel ressemble vaguement aux réseaux de neurones actuels : il filtre ses entrées, apprend les combinaisons fréquentes, et cela lui suffit à

Scanned by CamScanner

radicalement : l'apprentissage se met à ressembler à un raisonnement, une inférence logique qui tente de capturer les règles d'un domaine. Parvenir à créer des machines qui atteignent ce second niveau d'intelligence est le grand défi de la recherche contemporaine.

Apprendre, c'est inférer la grammaire d'un domaine

C'est une caractéristique de l'espèce humaine : nous essayons en permanence de tirer, d'une situation particulière, des conclusions de haut niveau, qu'en retour nous mettons à l'épreuve sur de nouvelles observations. Hiérarchiser ainsi ses connaissances, en tentant de formuler des lois abstraites qui rendent compte de nos observations, est une méthode d'apprentissage extraordinairement efficace, car les lois les plus abstraites sont précisément celles qui s'appliquent au plus grand nombre d'observations. Trouver la bonne loi, la règle logique qui rend compte de toutes les données, c'est accélérer massivement l'apprentissage.

Prenons un exemple : imaginons que je vous présente une dizaine d'urnes remplies de billes de différentes couleurs. Je prends une urne au hasard, dans laquelle je n'ai encore jamais puisé, j'y plonge la main, et j'en sors une bille verte. Pouvez-vous en déduire quoi que ce soit sur le contenu de cette urne ? Quelle sera la couleur du prochain tirage ?

La première réponse qui vous vient sans doute à l'esprit est : « Je n'en ai pas la moindre idée – vous ne m'avez donné pratiquement aucune information, comment pourrais-je connaître la couleur des autres billes ? » Oui mais... imagi-

règle suivante : dans une urne donnée, toutes les billes sont de la même couleur. Dans ce cas, le problème devient trivial. Même si vous me montrez une urne pour la première fois, il me suffit d'en extraire une seule boule verte pour en déduire que toutes les autres boules seront de cette couleur. Muni de cette règle générale, un seul tirage me suffit. Ainsi, une connaissance de haut niveau, que l'on appelle le niveau « méta », peut guider tout un ensemble d'observations de plus bas niveau. La métarègle qui dit que toutes les billes d'une urne sont de la même couleur, une fois apprise, accélère massivement l'apprentissage. Bien sûr, elle peut s'avérer fausse. Vous serez vigoureusement surpris (je devrais dire « métasurpris ») si la dixième urne que vous explorez contient des billes de toutes les couleurs. Dans ce cas, vous devrez changer de modèle et remettre en question l'hypothèse que toutes les urnes sont semblables. Peut-être hasarderez-vous une hypothèse d'encore plus haut niveau, une méta-méta-hypothèse selon laquelle les urnes sont de deux sortes : unicolores ou multicolores – auquel cas il vous faudra au moins deux tirages par urne avant de conclure quoi que ce soit. Dans tous les cas, le fait de hiérarchiser vos connaissances et de formuler des hypothèses de haut niveau vous aura fait gagner un temps précieux.

Apprendre efficacement, c'est donc hiérarchiser les informations : se forger, dès que possible, des règles générales, qui résument toute une série d'observations. Dès l'enfance, notre cerveau applique à bon escient ce principe de hiérarchie. Prenez un petit enfant de 2 ou 3 ans qui se promène dans un jardin et à qui ses parents apprennent un mot nouveau, disons « papillon ». Souvent, il lui suffit d'entendre le mot une fois ou deux, et c'est terminé : son sens est mémorisé. Cette vitesse d'apprentissage est stupéfiante. Elle dépasse tout ce que l'intelligence artificielle

Scanned by CamScanner

Université Paris Descartes
IUT, Département Informatique
Communication

Semaine du 11/03/2019

Interrogation sur le cours : Introduction à l'apprentissage profond

Cours 2 : Le perceptron

Complétez le texte ci-dessous avec les mots qui conviennent.

1) Le perceptron apprend à [classer] des données dans une catégorie en s'entraînant avec des [exemples] Pour s'entrainer, chaque donnée lui est fournie sous forme numérique en tant que vecteur. Les n-1 premières coordonnées codent les [propriété]d'une données. La dernière coordonnée étant sa [classe].

2) Dans l'exemple de classification de mail versus spam vu en cours, pour s'entraîner l'algorithme suit les étapes suivantes :

a) En ne considérant que les n-1 coordonnées de chaque vecteur mail, l'algorithme calcule la [moyenne] [pondérée] de chacun. Il s'agit d'une manière filtrer les données en accordant un facteur multiplicatif (plus ou moins important) appelé poids à chaque [propriété]. Il ajoute ensuite un terme b, nommé biais au résultat de son calcul.

b) Le perceptron calcule ensuite, pour chaque vecteur, la probabilité que ce vecteur appartienne à la classe spam grâce à la fonction [sigmoïde].

c) L'algorithme calcule ensuite l'erreur de classement qui est représentée par la distance mathématique entre la classe [estimée] du vecteur et sa classe [réelle]. Le perceptron dispose donc d'une erreur de classement pour chaque mail.

d) L'algorithme calcule ensuite l'erreur globale de classement des mails en effectuant la somme des carrés des

[erreurs] des vecteurs mail.

e) Une fois l'erreur globale calculée, une fonction mathématique le Gradient Descendant fournit, à partir de calculs qui tiennent compte de l'erreur globale, un nouveau vecteur de [poids] (chaque coordonnée de ce vecteur est un nouveau facteur multiplicatif pour

Scanned by CamScanner

la [propriété] correspondant à cette coordonnée) ainsi qu'une nouvelle valeur pour le biais.

f) Les étapes a) à e) sont réitérées un grand nombre de fois et lors de ces multiples itération, grâce à la fonction Gradient Descendant, l'erreur globale va [diminuer]. Elle finira par atteindre une valeur stable et [minimale].

Les dernières valeurs des poids et du biais sont alors utilisées pour [classer] de nouveaux mails dans la bonne catégorie. Ainsi, le Perceptron aura appris à partir d'un ensemble d'exemples les meilleures coefficients à utiliser dans la fonction [sigmoïde] pour que cette dernière prenne en entrée un vecteur mail et fournisse en sortie une [classe/prédiction] proche de [1] si le mail est un vrai mail, et proche de [0] si le mail est un spam.

Scanned by CamScanner

- Idée de Réseaux de neurone artificiel.
↳ traitement des propositions logiques.

A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH and WALTER H. PITTS

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

INTRODUCTION

THEORETICAL neurophysiology rests on certain cardinal assumptions. The nervous system is a net of neurons, each having a soma and an axon. Their adjunctions, or synapses, are always between the axon of one neuron and the soma of another. At any instant a neuron has some threshold, which excitation must exceed to initiate an impulse. This, except for the fact and the time of its occurrence, is determined by the neuron, not by the excitation. From the point of excitation the impulse is propagated to all parts of the neuron. The velocity along the axon varies directly with its diameter, from less than one meter per second in thin axons, which are usually short, to more than 150 meters per second in thick axons, which are usually long. The time for axonal conduction is consequently of little importance in determining the time

Scanned by CamScanner

of arrival of impulses at points unequally remote from the same source. Excitation across synapses occurs predominantly from axonal terminations to somata. It is still a moot point whether this depends upon irreciprocity of individual synapses or merely upon prevalent anatomical configurations. To suppose the latter requires no hypothesis *ad hoc* and explains known exceptions, but any assumption as to cause is compatible with the calculus to come. No case is known in which excitation through a single synapse has elicited a nervous impulse in any neuron, whereas any neuron may be excited by impulses arriving at a sufficient number of neighboring synapses within the period of latent addition, which lasts less than one quarter of a millisecond. Observed temporal summation of impulses at greater intervals is impossible for single neurons and empirically depends upon structural properties of the net. Between the arrival of impulses upon a neuron and its own propagated impulse there is a synaptic delay of more than half a millisecond. During the first part of the nervous impulse the neuron is absolutely refractory to any stimulation. Thereafter its excitability returns rapidly, in some cases reaching a value above normal from which it sinks again to a subnormal value, whence it returns slowly to normal. Frequent activity augments this subnormality. Such specificity as is possessed by nervous impulses depends solely upon their time and place and not on any other specificity of nervous energies. Of late only inhibition has been seriously adduced to contravene this thesis. Inhibition is the termination or prevention of the activity of one group of neurons by concurrent or antecedent activity of a second group. Until recently this could be explained on the supposition that previous activity of neurons of the second group might so raise the thresholds of internuncial neurons that they could no longer be excited by neurons of the first group, whereas the impulses of the first group must sum with the impulses of these internuncials to excite the now inhibited neurons. Today, some inhibitions have been shown to consume less than one millisecond. This excludes internuncials and requires synapses through which impulses inhibit that neuron which is being stimulated by impulses through other synapses. As yet experiment has not shown whether the refractoriness is relative or absolute. We will assume the latter and demonstrate

Scanned by CamScanner

that the difference is immaterial to our argument. Either variety of refractoriness can be accounted for in either of two ways. The "inhibitory synapse" may be of such a kind as to produce a substance which raises the threshold of the neuron, or it may be so placed that the local disturbance produced by its excitation opposes the alteration induced by the otherwise excitatory synapses. Inasmuch as position is already known to have such effects in the case of electrical stimulation, the first hypothesis is to be excluded unless and until it be substantiated, for the second involves no new hypothesis. We have, then, two explanations of inhibition based on the same general premises, differing only in the assumed nervous nets and, consequently, in the time required for inhibition. Hereafter we shall refer to such nervous nets as *equivalent in the extended sense*. Since we are concerned with properties of nets which are invariant under equivalence, we may make the physical assumptions which are most convenient for the calculus.

Many years ago one of us, by considerations impertinent to this argument, was led to conceive of the response of any neuron as factually equivalent to a proposition which proposed its adequate stimulus. He therefore attempted to record the behavior of complicated nets in the notation of the symbolic logic of propositions. The "all-or-none" law of nervous activity is sufficient to insure that the activity of any neuron may be represented as a proposition. Physiological relations existing among nervous activities correspond, of course, to relations among the propositions; and the utility of the representation depends upon the identity of these relations with those of the logic of propositions. To each reaction of any neuron there is a corresponding assertion of a simple proposition. This, in turn, implies either some other simple proposition or the disjunction or the conjunction, with or without negation, of similar propositions, according to the configuration of the synapses upon and the threshold of the neuron in question. Two difficulties appeared. The first concerns facilitation and extinction, in which antecedent activity temporarily alters responsiveness to subsequent stimulation of one and the same part of the net. The second concerns learning, in which activities concurrent at some previous time have altered the net permanently, so that a stimulus which would previously have been inadequate is now

Scanned by CamScanner

adequate. But for nets undergoing both alterations, we can substitute equivalent fictitious nets composed of neurons whose connections and thresholds are unaltered. But one point must be made clear: neither of us conceives the formal equivalence to be a factual explanation. *Per contra!*—we regard facilitation and extinction as dependent upon continuous changes in threshold related to electrical and chemical variables, such as after-potentials and ionic concentrations; and learning as an enduring change which can survive sleep, anaesthesia, convulsions and coma. The importance of the formal equivalence lies in this: that the alterations actually underlying facilitation, extinction and learning in no way affect the conclusions which follow from the formal treatment of the activity of nervous nets, and the relations of the corresponding propositions remain those of the logic of propositions.

The nervous system contains many circular paths, whose activity so regenerates the excitation of any participant neuron that reference to time past becomes indefinite, although it still implies that afferent activity has realized one of a certain class of configurations over time. Precise specification of these implications by means of recursive functions, and determination of those that can be embodied in the activity of nervous nets, completes the theory.

THE THEORY: NETS WITHOUT CIRCLES

We shall make the following physical assumptions for our calculus.

1. The activity of the neuron is an "all-or-none" process.
2. A certain fixed number of synapses must be excited within the period of latent addition in order to excite a neuron at any time, and this number is independent of previous activity and position on the neuron.
3. The only significant delay within the nervous system is synaptic delay.
4. The activity of any inhibitory synapse absolutely prevents excitation of the neuron at that time.
5. The structure of the net does not change with time.

Scanned by CamScanner

animaux, a tendance à adapter l'environnement à elle-même. Comme Grock, le célèbre clown suisse, qui attirait le piano à lui au lieu d'avancer son tabouret.

— À l'échelle de l'évolution, notre génome et notre corps s'adaptent très lentement. Mais cela est compensé par la grande capacité d'apprentissage de notre cerveau. Il nous est donc plus facile de modifier notre environnement en créant des outils que notre cerveau apprend facilement à utiliser. L'usage de l'outil, c'est peut-être *la* caractéristique de l'intelligence humaine. Quand on étudie l'émergence du genre *Homo*, les principaux indices dont on dispose, ce sont des squelettes et des outils. L'usage d'outils complexes fait pratiquement partie de la définition de notre espèce.

— *Nous avons vu que d'autres êtres vivants, comme les corvidés, utilisent et même fabriquent des outils sommaires. Mais ils sont beaucoup moins dépendants d'eux pour survivre.*

— Dans notre cas, on peut effectivement se demander ce que serait notre intelligence sans tous les outils qui nous entourent. Notamment le papier et le crayon. L'invention de l'écriture, dont nous avons beaucoup parlé, ne conduit pas seulement à une transformation mentale. C'est un outil qui étend nos compétences en modifiant notre environnement. L'écriture nous permet de projeter nos pensées, de les stocker dans une mémoire externe, et de les retrouver après des années

ou même des siècles. C'est une invention qui augmente le potentiel des circuits du cerveau, parce qu'elle externalise en quelque sorte une partie de notre intelligence.

— *C'est aussi une interaction avec le passé de notre espèce.*

— Absolument. Mais il faut quand même apprendre à le « ré-internaliser ». On ne se branche pas simplement sur une source externe d'informations. Des années d'éducation sont nécessaires avant de parvenir à reconstituer dans notre tête le réseau de pensées des écrivains qui nous ont précédés. C'est Newton qui aurait dit : « J'ai vu loin parce que je suis monté sur les épaules de géants. » Peut-être aurait-il dû dire : « Je suis monté sur une pyramide de nains. » Nous sommes tous des nains, mais l'effet cumulatif de la culture humaine nous porte vers le haut. Et certains peuvent monter très haut.

— *Et la pyramide continue de monter.*

— Absolument. Je suis fasciné par l'intelligence des ordinateurs qui complète la nôtre. Aujourd'hui, plus aucun être humain ne peut se passer de cet environnement numérique. Pourtant, il s'agit d'une invention extrêmement récente.

— *En même temps, nous perdons certaines de nos capacités. Fini le marchand de légumes, sur le marché, avec un crayon sur l'oreille et un bloc, capable d'effectuer,*

Scanned by CamScanner

à une vitesse impressionnante, des additions sans faute! Sans calculette, nous sommes de plus en plus lents.

— C'est vrai. Le psychologue Alfred Binet, qui a écrit un livre sur Inaudi, un calculateur prodige, était allé voir au Bon Marché où, à l'époque, il y avait des calculateurs humains – principalement des femmes, en fait – qui y étaient employés. Elles faisaient très peu d'erreurs, bien que calculant extrêmement vite. Binet a montré que cela n'avait rien de surhumain : nous avons tous ce potentiel. Il s'agit seulement de personnes qui se sont entraînées très tôt. Leurs cerveaux ne sont pas différents : elles appliquent des stratégies et y ont passé des milliers d'heures – comme pour apprendre à jouer du piano.

— *Ce que je veux dire, c'est qu'aujourd'hui, notre intelligence augmentée par les machines est devenue tellement dépendante de ces machines qu'elle est devenue vulnérable. Après un cataclysme qui détruirait tous ses outils, que deviendrait* Homo sapiens *?*

— Il se retrouverait à un niveau pré-Cro-Magnon. Il aurait à reconstituer trois cent mille ans de culture en ayant perdu, à mon avis, énormément d'intelligence, liée, effectivement, à tous les outils qui nous entourent. Par contre, *Homo sapiens* reste l'espèce la plus flexible de la planète. C'est la seule qui a réussi à envahir des biotopes complètement différents, de la forêt à la banquise en passant par le désert, et jusqu'à

l'espace, jusqu'à la lune, pourtant dépourvue d'atmosphère! C'est aussi l'une des caractéristiques principales de notre intelligence que de faire feu de tout bois, de s'adapter à toutes les circonstances. C'est pour cela que je reste optimiste. Il est vrai que nous avons dévasté la planète en l'adaptant à nos besoins présents. Elle change vite et va encore changer profondément. C'est inquiétant, bien sûr. Mais la flexibilité humaine devrait permettre de surmonter ces difficultés.

Alien est déjà là

— *Le développement spectaculaire de l'intelligence humaine ne bloque-t-il pas celui des autres intelligences?*

— La réponse est évidemment oui. Nous avons fait disparaître un nombre considérable d'espèces qui n'étaient pas idiotes, mais dont l'intelligence était moins avancée que la nôtre. Il y a seulement quelques dizaines de milliers d'années, la planète Terre comprenait au moins quatre espèces d'êtres humains : la nôtre, mais aussi les hommes de Neandertal, de Flores et de Denisova. Tous ont disparu… probablement de notre faute.

— *L'intelligence des espèces qui, pour diverses raisons, sont plus ou moins protégées, peut-elle encore progresser?*

Scanned by CamScanner

Bien sûr. L'évolution ne s'est pas arrêtée. Je dirais même : au contraire ! Nous avons parlé des corbeaux avec le crochet. C'est nous qui leur fournissons un environnement nouveau dans lequel ils évoluent. Des expériences ont montré que les oiseaux des villes adaptent la couleur de leur plumage et leurs comportements. Je ne sous-estimerais pas l'intelligence animale ! Si les humains disparaissaient, une intelligence supérieure apparaîtrait peut-être chez les corbeaux, en quelques millions d'années, ou bien chez les céphalopodes qui font déjà preuve d'une intelligence très vive. Il n'est peut-être pas nécessaire d'aller chercher Alien ailleurs que sur terre… ou sous la mer !

— *Et notre cerveau d'*Homo sapiens *peut-il encore évoluer ?*

— Il y a un blocage : l'accouchement. Jean-Jacques Hublin, du Collège de France, l'a bien étudié : les contraintes sont très fortes sur le passage du bébé dans les voies pelviennes au moment de l'accouchement. C'est sans doute pour ça que nous naissons prématurés par rapport aux autres espèces.

— *La taille de notre tête empêche toute évolution future ?*

— Rien n'est impossible, mais ce genre d'évolution est très lent et difficile. Notre cerveau consomme déjà vingt pour cent de l'énergie totale de notre corps au repos. Une grande partie de l'évolution humaine a consisté à trouver des sources nouvelles d'énergie pour faire face aux besoins du cerveau, en devenant omnivore, carnivore, en cuisant les aliments… Sauf à trouver de nouvelles manières d'alimenter ce cerveau incroyablement dispendieux, je doute fort qu'il y ait une évolution rapide de nos capacités intellectuelles.

Graphic design

— *Alors c'est vraiment la fin de notre histoire ?*

— Non, mais la prochaine évolution passera par les machines. Les interfaces cerveau/machine seront beaucoup plus efficaces. Notre espèce a déjà accompli un chemin considérable. Un écran d'ordinateur qui affiche de l'écriture, des graphiques, des photographies et des graphes mathématiques augmente considérablement la bande passante avec laquelle notre cerveau peut interagir avec la machine. Je pense que l'on sous-estime, dans les inventions récentes de l'humanité, l'importance des diagrammes comme ceux qui, dans le journal, nous montrent l'évolution du chômage ou du coût de la vie. C'est une remarquable utilisation de notre système visuel pour faire passer des milliers de données. D'un simple coup d'œil, on écarte les données erronées, on saisit les tendances, on comprend dans quel sens elles évoluent, et on parvient à extrapoler au futur. Si l'on

Scanned by CamScanner

nous donnait ces données sous la forme d'un tableau, on n'y arriverait jamais.

— *Et c'est un progrès de notre intelligence ?*

— Bien sûr. Nous l'avons augmentée de façon considérable en laissant l'ordinateur faire les calculs et afficher les données sous une forme imagée, qui permet à la formidable « carte graphique » de notre cerveau de les interpréter à une vitesse remarquable.

— *Ça a débuté avec le « camembert »…*

— … qui est considéré, aujourd'hui, comme l'un des pires graphiques qui soient ! Mais il existe actuellement toute une mouvance, le *graphic design*, qui réfléchit à la meilleure manière d'utiliser les points, les courbes, la couleur, le mouvement, les curseurs… afin de fournir au cerveau un maximum d'informations utiles.

— *La vision est un moyen privilégié de communication ?*

— Comparé aux interfaces cérébrales, certainement. Le débit d'informations est beaucoup plus grand par la vision, et même par la parole, que par les puces électroniques qui tentent d'enregistrer l'activité cérébrale et d'échanger les informations de façon bidirectionnelle entre le cerveau et un ordinateur. Néanmoins, à terme, je pense que la perspective va

s'inverser. On commence à le voir, chez l'animal. Il est devenu possible de disposer quelques milliers d'électrodes dans le cerveau et de visualiser ainsi l'activité de milliers de neurones. Surtout, il devient possible d'y réinjecter du courant, donc d'envoyer vers le cerveau des informations venues d'une machine. De belles expériences ont été faites par Thierry Bal au CNRS de Gif-sur-Yvette. Elles consistent à mettre une électrode dans un neurone et à injecter dans cette électrode un courant variable, savamment calculé, nécessaire pour simuler la présence d'autres neurones. On parvient ainsi à faire croire au neurone qu'il est entouré d'autres circuits neuronaux qui lui envoient des informations. Le neurone réel devient intégré dans un circuit virtuel.

— *Impressionnant.*

— Cela se fait depuis près de vingt ans en neurosciences, sur l'animal. Lorsque ça se fera chez l'homme… Déjà, certains patients tétraplégiques commandent un robot avec une puce implantée dans leur cortex moteur. Le système se fonde sur l'activité de quelques dizaines de neurones qui sont enregistrés en même temps. C'est très peu. Un jour, on pourra passer à des systèmes massivement parallèles qui communiqueront avec dix mille, cent mille, voire un million de neurones. Certaines méthodes optiques permettent de visualiser une grande population de neurones, et le microscope est si petit qu'il peut être implanté dans le cortex. Le débit d'informations va devenir tellement

Scanned by CamScanner

...portant qu'on aura un sentiment d'immédiateté encore plus grand que celui qu'on a vis-à-vis d'un smartphone par exemple. On n'aura pas l'impression de faire un effort pour poser une question et obtenir une réponse de la machine ou un mouvement du robot.

— *La machine deviendra une extension du cerveau?*

— C'est l'impression que nous aurons : une extension naturelle de notre cerveau. Les neurotechnologies évoluent si rapidement… J'espère que je le verrai.

— *Vous en parlez comme s'il s'agissait des performances du prochain smartphone!*

— Non, il s'agit bien sûr d'un progrès considérable, et qui pose de nombreux problèmes éthiques. Je crois d'ailleurs que c'est là que se situeront les blocages : s'autorisera-t-on à expérimenter sur le cerveau humain, avec tous les risques que cela présente?

— *Mais tout ce qui se passe aujourd'hui ne va-t-il pas poser de graves questions? Des problèmes éthiques, mais aussi, plus concrètement, une réorganisation totale du fonctionnement de nos sociétés?*

— Sans aucun doute. Prenez un outil familier comme la voiture. Tous les problèmes techniques liés à son fonctionnement autonome sont en passe d'être

résolus. Il y aura alors beaucoup moins d'accidents, à tel point que j'estime probable que, dans une quinzaine d'années peut-être, les humains n'auront plus le droit de conduire, sauf sur des circuits, pour s'amuser. Ils sont bien trop dangereux au volant. Il n'y a pas longtemps, j'ai vu au MIT un système de croisement pour véhicules autonomes qu'ils ont déjà inventé. Il faut avouer que le feu rouge est une invention assez stupide. On s'arrête parfois durant une minute alors qu'on est seul! Dans le carrefour intelligent du MIT, vous avez des voitures autonomes qui savent toutes où elles se trouvent les unes par rapport aux autres, et ralentissent imperceptiblement pour arriver au carrefour où elles se croisent sans se toucher. Plus besoin de s'arrêter. Dans une telle société, un humain devient un fardeau. Son cerveau ne peut pas gérer tant d'informations aussi vite : cela se joue à la fraction de seconde.

Homo sapiens *est une transition*

— *Des aiguilleurs du sol numériques… Les machines finiront-elles par devenir plus intelligentes que nous?*

— Peut-être pas, mais certainement mieux adaptées, plus efficaces dans des domaines particuliers. D'ailleurs, c'est déjà le cas au jeu d'échecs, au jeu de go… Même un domaine comme le mien, celui de la science, que l'on considère comme noble parce qu'il nécessite de l'intelligence humaine, sera touché. Plus

Scanned by CamScanner

j'y réfléchis, plus je pense que notre cerveau est médiocre. Il me faut à peu près une heure pour lire un article scientifique dont je ne me souviens même pas des détails… contre moins d'une seconde à Google! Aujourd'hui, les systèmes artificiels qui lisent la littérature scientifique ont encore des problèmes de compréhension. Ils se contentent de compiler sans vraiment comprendre. Mais je pense que très vite nous aurons des assistants de recherche artificiels qui nous aideront et rempliront un certain nombre de tâches bien plus vite que nous. Dans mon labo, c'est déjà le cas. Lorsque j'enregistre des signaux cérébraux, une IRM me renvoie, toutes les secondes, les données de deux cent mille voxels (des pixels en 3D) qui couvrent tout le cerveau. Il faut les analyser un par un en fonction du temps. Cela constitue un volume de données extraordinaire. D'ores et déjà, ce sont des algorithmes automatisés qui vont fouiller ces données, et déterminer où se trouvent les régions importantes, quels signaux ont bougé, en fonction de quels paramètres… Aujourd'hui, au premier abord, ce n'est plus un humain qui regarde ces données, c'est une machine.

— *Cela ne vous inquiète pas?*

— J'aime assez l'idée qu'*Homo sapiens* représente une espèce de transition, une sorte de précurseur de la prochaine intelligence. L'intelligence humaine sera profondément augmentée par celle de la machine. On voit maintenant qu'il nous est possible de devenir

beaucoup plus intelligents grâce à nos inventions. Mais c'est nous qui créons ces machines! Nous constituons une étape de l'évolution. C'est notre destin et je lui trouve une certaine grandeur.

Scanned by CamScanner

Université Paris Descartes
IUT, Département Informatique
DUT2 - Communication

Semaine du 18/02/2019

Questions de réflexion sur l'ouvrage

L'intelligence, des origines aux neurones artificiels : vers une nouvelle étape de l'évolution

--Yann Le Cun et Stanislas Dehaene--

Entretien avec Jacques Girardon

→ Lien

1) Pour commencer, pourriez-vous rappeler la notion de sélection naturelle ? plus précisément, qu'entend-on par le fait qu'un caractère génétique soit sélectionné par l'environnement ? Un caractère génétique est sélectionné par l'environnement lorsqu'il donne un avantage à l'individu.

2) Qu'est-ce qui caractérise le rapport qu'entretien l'homme avec l'outil si on le compare au rapport qu'entretiennent les animaux non humains avec l'outil ? Notre intelligence et notre savoir sont dépendant de la disponibilité de nos outils

3) Selon les scientifiques, l'évolution de notre espèce, sur le plan intellectuel, aurait une limite. Expliquez pourquoi ? Notre tête est déjà trop grosse et galererai à passer dans le bassin de la maman.

4) Quel organe sensoriel est devenu significativement plus puissant grâce aux ordinateurs ? Expliquez. Il s'agit des yeux. En effet, les ordinateurs peuvent compiler d'énormes quantités de données dans la forme graphique qui nous permettent de nous rendre compte d'un phénomène d'un simple coup d'œil.

Scanned by CamScanner

Université Paris Descartes
IUT, Département Informatique
Communication

Semaine du 25/02/2019

Un réseau de neurones simple : la régression logistique

L'Apprentissage Machine supervisé comprend un type d'apprentissage appelé régression logistique. En dépit de son nom, **la régression logistique n'est pas un algorithme de régression (prédire** une valeur) mais de classification (classer une valeur dans une catégorie).

La régression logistique, introduite initialement par Cox en 1958 est interessante à étudier ici car il **s'agit d'un réseau** mono-neuronal appelé aussi perceptron. **C'est un** exemple relativement simple qui décrit les principes essentiels de **l'apprentissage machine** et de **l'apprentissage profond**.

Principe de fonctionnement :

Dans l'algorithme du perceptron, on dispose d'un ensemble d'exemples, c'est-à-dire des données ainsi que la catégorie à laquelle elles appartiennent. L'algorithme s'entraîne à l'aide de ces exemples, c'est-à-dire qu'il apprend la meilleure procédure pour classer chacune de ses données dans sa catégorie. **Pendant la phase de test qui suit l'entraînement, l'algorithme classe de nouvelles données** jamais vues **pendant l'entraînement.**

Un perceptron ne traite toute information mais des données spécifiques, une image ou un son, etc.

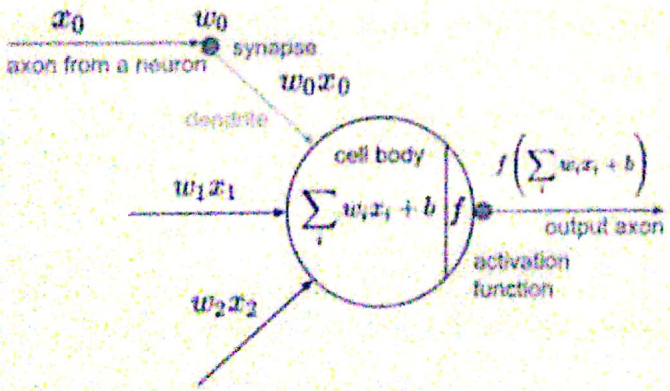

Figure 1. Schémas d'une régression logistique (Perceptron).

Le perceptron représenté ci-dessus code des données que l'on représente sous forme de vecteur à trois coordonnées. On ne considère que trois coordonnées dans cet exemple pour des raisons de simplification. Ces coordonnées sont les propriétés caractéristiques de l'image, du son etc. Pour chaque donnée traitée, les trois propriétés caractéristiques parviennent au perceptron grâce à trois neurones auxquels il est connecté.

Scanned by CamScanner

Quels sont les composants d'une régression logistique ?

On doit tout **d'abord** représenter les données et la classe à laquelle elles appartiennent. Pour cela on caractérise les données par des propriétés. Dans le cadre de ce cours, ces propriétés seront traduites sous forme numérique. Par exemple, si on veut classer des arbres dans une catégorie (bouleau, pin, chêne..), on peut choisir de représenter un arbre sous forme de vecteur à trois coordonnées : x1, sa hauteur, x2, son diamètre et x3, son type de feuille (codé numériquement). Ainsi, on représente un arbre donné, comme un vecteur x = (x1, x2, x3).

Prenons un autre exemple, supposons que notre algorithme perceptron soit un classificateur de spam. On peut décider de représenter un mail par des propriétés qui caractérisent sa nature : mail normal, ou spam. Par exemple, on le représentera de manière très simplifiée pour le contexte de ce cours sous forme de vecteur à trois coordonnées : x1, le pourcentage de mots mal orthographiés, x2, le pourcentage d'**occurrences du mot** clé 'achat'. La dernière coordonnée du vecteur représente la classe du mail. Ainsi, un mail sera représenté par un vecteur m = (x1, x2, x3).

Un cas simple de classification : mail versus spam

Reprenons le classificateur de mails. **Pour s'entrainer,** le perceptron reçoit des vecteurs ayant trois coordonnées. Les deux premières seront les propriétés caractéristiques du mail, la dernière coordonnée représentera sa classe **: 1 si c'est** un spam, 0 **si c'est un** mail.

Supposons que nous ayons trois exemples de vecteurs mail **pour l'entraînement** : m1 = (0.5, 1.1, 1), m2 = (0, 2.3, 0), m3 = (0.02, 0.7, 0).

Pour vérifier que vous avez bien compris, complétez le tableau suivant de représentation numérique **des trois exemples qui seront fourni à l'algorithme.**

Nom du mail	% fautes	% mots clé 'achat'	classe du mail exemple (0 ou 1)
m1	0.5	1.1	1
m2	0	2.3	0
m3	0.02	0.7	0

L'ensemble d'entraînement contient trois vecteurs dont les deux premières coordonnées sont les propriétés du mail et la dernière coordonnée, qui se distingue des deux autres, est la classe du mail.

Scanned by CamScanner

A partir de ces trois exemples, le perceptron doit apprendre à décider, si un nouveau mail appartient ou non à la classe spam. Le perceptron prendra en entrée un mail et fournira en sortie une réponse. En principe, la réponse sera 1 si le mail est un spam, 0 si le mail est un mail normal. En réalité, le perceptron ne fournira pas une réponse binaire mais une sorte de **probabilité d'appartenance à la classe spam. Plus cette probabilité sera proche de 1, plus le mail aura de chances d'être un spam, plus elle se** rapprochera de 0, plus **le mail aura de chances d'être un mail normal.**

On va introduire plusieurs fonctions et outils mathématiques qui implémentent **l'une des manières d'apprendre** à réaliser ce type de décision : la somme pondérée, la fonction sigmoïde et la formule de **l'erreur** que nous expliciterons mathématiquement ci-**dessous. Avant d'entrer dans le détail technique** voyons la procédure générale et le rôle de ces objets mathématiques.

Procédure

L'apprentissage est réalisé par l'algorithme en faisant plusieurs itérations (répétitions) : estimer la classe de chaque mail à partir de la fonction sigmoïde, calculer **l'écart** entre la classe estimée du mail et la classe **réelle (l'erreur), puis, en fonction de l'erreur, refaire une estimation, calculer à nouveau l'erreur** liée à cette nouvelle estimation **et ainsi de suite jusqu'à stabilisation de l'erreur à une valeur minimum.**

Intuitivement, ce que le perceptron va apprendre à partir des exemples, **c'est l'importance qu'il** accordera à chacune des propriétés caractéristiques du mail (ici, ces propriétés sont les fautes et les mots clés) de manière à bien classer ces exemples.

Mathématiquement, il va apprendre (1) un vecteur de poids et (2) un terme appelé biais.

Pour le moment, nous ne nous occuperons pas de la signification du terme b de la somme pondérée.

Que représentent les poids ?

A chaque propriété va être associé un poids, c'est un nombre par lequel on multiplie la valeur de la propriété et qui contrôle **l'importance accordée à cette propriété**. Un poids important « amplifie » la propriété, un poids faible la « réduit ».

Pour un vecteur de données comme un mail noté m = (x1,x2,x3,...xn) ou les xi représentent ici les propriétés du mail (il y en a n, donc la n+1ème sera la classe du mail), chaque coordonnée xi du vecteur m va être associée à un poids wi.

On calcule la somme pondérée de m dont la formule générale vaut :

Scanned by CamScanner

$$z = \left(\sum_{i=1}^{n} w_i x_i\right) + b$$

Dans l'exemple précédent des mails, n=2 car un vecteur mail n'a que deux propriétés et sa somme pondérée vaut ainsi

$(w_1 \times pFautes + w_2 \times pAchat) + b$

Pour apprendre, à partir des **exemples qu'on lui a** fournis, les meilleures valeurs des poids pour effectuer le classement, le perceptron va utiliser les deux objets mathématiques suivants :

1) La fonction sigmoïde σ, **qui** prend en entrée la somme pondérée et fournit en sortie la classe du **vecteur. La classe du vecteur ne va pas être fournie précisément mais sous la forme d'une** probabilité.

La fonction sigmoïde a pour forme générale :

$$y = \sigma(z) = \frac{1}{1 + e^{-z}}$$

où z est la somme pondérée.

Dans l'exemple avec les mails, la fonction sigmoïde vaut donc :

$$y = \sigma(z) = \frac{1}{1 + e^{-(w_1 \times pFautes + w_2 \times pAchat) + b}}$$

2) L'erreur *E* qui calcule **l'écart** entre la classe estimée sigmoïde et la classe réelle du mail.

L'algorithme « travaille sur » les exemples et apprends à établir une correspondance entre les propriétés **des vecteurs exemples qu'on lui donne en entrée et la classe à laquelle ils appartiennent.** Il ajuste la correspondance en modifiant les poids.

Scanned by CamScanner

Prenons un exemple calculatoire avec l'algorithme de classement des mails pour mieux comprendre.

Nous aurons besoin de valeurs initiales pour les poids w1 et w2 et pour le bais b.

Pour simplifier, choisissons ces valeurs initiales de manière aléatoire, par exemple w1=0.3, w2=0.5 et b=0.7. Autrement dit, le vecteur poids w = (0.3, 0.5).

Rappelons que nos trois exemples de mails codés numériquement sont m1 = (0.5, 1.1, 1), m2 = (0, 2.3, 0), m3=(0.02, 0.7, 0).

Etape 1-Calculons la valeur estimée de la classe de chaque vecteur avec cette pondération (initiale).

$y1 = \sigma(w1 \times 0.5 + w2 \times 1.1 + 0.7) = \sigma(0,3 \times 0,5 + 0,5 \times 1,1 + 0,7) = \sigma(1,4)$

$y2 = \sigma\left(\frac{1}{1+e^{-1,4}}\right) = \frac{1}{0,2466} = 0,864$

$y3 = \sigma(w1 \times 0.02 + w2 \times 0.7 + 0.7) = 0,7419$

Récapitulons les résultats :

Exemple d'entraînement	Classe estimée par σ	Clase réelle
m1	0,80218	1
m2	0,864427	0
m3	0,741925	0

Il nous faut maintenant évaluer pour chaque mail, si la valeur de la classe estimée (notée y ici) est proche de la valeur de la classe réelle (que l'on notera t –true- et qui vaut 0 ou 1) ou non.

Il s'agit de calculer une distance mathématique globale pour les trois mails, entre la classe réelle et la classe estimée, cela représentera l'erreur.

La formule générale de l'erreur est :

$$E = \frac{1}{2}\sum_{n}(t^{(n)} - y^{(n)})^2$$

où n est le nombre d'exemples.

Attention, $t^{(n)}$ est la classe réelle du nième exemple (nème mail dans notre cas).

Scanned by CamScanner

Etape 2 Calculons pour chaque mail l'erreur (distance entre la classe estimée et la classe réelle) et faisons la somme des carrés des erreurs des trois mails.

$E = \frac{1}{2}[(1 - 0,802148)^2 + (0 - 0,864277)^2 + (0 - 0,741954)^2]$

$= 0,668$

Jusqu'ici, nous avons

1. Estimé la classe de chaque mail avec la fonction sigmoïde
2. Comparé cette estimation à la classe réelle.
3. Calculé l'**erreur entre ces deux valeurs pour chaque** mail et fait la somme des carrés des erreurs des mails.

Maintenant, nous allons mettre à jour les poids grâce à une fonction Gradient Descendant que nous ne détaillerons pas ici.

La fonction Gradient **fournit de nouveaux poids de façon à réduire l'erreur**.

Supposons que nous ayons eu recours à cette fonction grâce à laquelle nous obtenons les poids et le biais suivants : w1 = 0.9, w2=0.1, b=0.7.

Nous allons répéter les étapes 1 et 2 avec cette nouvelle pondération.

Etape 1-Calculons la valeur estimée de la classe de chaque vecteur avec la deuxième pondération.

$y1 = \sigma(w1x0.5 + w2x1.1 + 0.7) = \sigma(0,5 \times 0,5 + 0,1 \times 1,1 + 0,7)$

$= \sigma(0,25 + 0,11 + 0,7) = \sigma(1,06)$

$= 0,779$

$y2 = \sigma(0 \times 0,5 + 2,3 \times 0,1 + 0,7)$

$= \sigma(0,23 + 0,7) = \sigma(0,93)$

$= 0,717$

$y3 = \sigma(0,02 \times 0,5 + 0,7 \times 0,1 + 0,7)$

$= \sigma(0,01 + 0,07 + 0,7) = \sigma(0,78)$

$= 0,687$

Récapitulons les résultats :

Scanned by CamScanner

Exemple d'entrainement	Classe estimée par σ	Classe réelle
m1	0,779	1
m2	0,717	0
m3	0,687	0

Etape 2 Calculons pour chaque mail l'erreur (distance entre la classe estimée et la classe réelle) et faisons la somme des carrés des erreurs des trois mails.

E =

≈ 0,5178

Comparez l'erreur de l'itération 1 avec l'erreur de l'itération 2. Que remarquez-vous ?

Nous pouvons continuer ces itérations un grand nombre de fois **et l'erreur va** décroitre jusqu'à atteindre une valeur minimum stable.

Nous venons de voir les principes essentiels de la régression logistique. C'est la valeur des poids stabilisés qui va représenter les paramètres de la sigmoïde, laquelle sera utilisée (de la même manière que pour les exemples) pour classer de nouvelles valeurs de mails.

Scanned by CamScanner