Un regard humain sur le « transformer »
Un opérateur de calcul universel

Jean Rohmer
Institut Fredrik Bull

Journée Aristote Anthropologie et IA

Ecole Polytechnique

16 novembre 2026

L’explication simple et complète des Transformers

Moteurs de l’IA Générative

En 40 minutes chrono!

TOUT TIENT EN DEUX FORMULES

For a standard transformer layer ℓ, consisting of a multi-head attention module $\mathrm{Attn}_\ell(\cdot)$ and a multi-layer perceptron $\mathrm{MLP}_\ell(\cdot)$, we have

$$h'_\ell(x) = h_{\ell-1}(x) + \mathrm{Attn}_\ell(h_{\ell-1}(x)),$$
$$h_\ell(x) = h'_\ell(x) + \mathrm{MLP}_\ell(h'_\ell(x)) = h_{\ell-1}(x) + \mathrm{Attn}_\ell(h_{\ell-1}(x)) + \mathrm{MLP}_\ell(h'_\ell(x)).$$

Unrolling the recursion gives

$$h_\ell(x) = h_0(x) + \sum_{k=1}^{\ell} \Big(\mathrm{Attn}_k(h_{k-1}(x)) + \mathrm{MLP}_k(h'_k(x))\Big),$$

A multi-head attention module consists of N_H attention heads operating in parallel. Consider the residual-stream activations $h_\ell^{1:T}(x) \in \mathbb{R}^{d\times T}$ and let $h := h_\ell^{1:T}(x)$ for notational simplicity. For each head $r \in \{1, \ldots, N_H\}$, the residual stream is first projected into query, key, and value representations:

$$Q^r = W_Q^r h, \qquad K^r = W_K^r h, \qquad V^r = W_V^r h,$$

where $W_Q^r, W_K^r, W_V^r \in \mathbb{R}^{d_{\text{head}}\times d}$ and d_{head} denotes the dimensionality of each attention head. The *attention pattern* of head r is then computed as

$$A^r = \mathrm{softmax}\left(\frac{(Q^r)^\top K^r}{\sqrt{d_{\text{head}}}}\right) \in \mathbb{R}^{T\times T},$$

where the softmax is applied row-wise. Thus, $A^r_{t,t'}$ denotes the attention weight assigned by the query at position t to the key at position t'. The head then writes an update to the residual stream:

$$\Delta h^r = W_O^r V^r (A^r)^\top,$$

where $W_O^r \in \mathbb{R}^{d\times d_{\text{head}}}$. The output of the full multi-head attention module is the sum of the contributions from all heads:

$$\mathrm{Attn}(h) = \sum^{N_H} \Delta h^r.$$

Les Transformers sont une invention du génie humain

Pour construire statistiquement

Sur des bases mathématiques

Un système capable d’imiter

Quelques capacités cognitives humaines

Non, Jean, c’est pire!

Voilà la réalité:

Les Transformers ont été inventés par hasard

Pour accélérer un algorithme poussif

Et se sont révélés à la surprise générale

Capable d’inventer des théories

Et de les mettre en œuvre

Par des mécanismes de raisonnement

Eux aussi inventés

Et très supérieurs à ceux des humains

Bon, Jean, je vais t’aider

Je vois avec Katia, je vais prendre le contrôle du rétroprojecteur

Et je vais rédiger tes slides en temps réel

Tu vas voir, ça va bien se passer

Toi, tu commenteras, je te donnerai des billes éventuellement par SMS

Et je vais commencer par une phrase qui nous servira de fil conducteur

Ton exposé en une seule phrase !

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a affiné elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une capacité de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Une drôle de machine immense

Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Une drôle de machine immense

Dont l’architecture a été découverte par hasard

En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard

En progrès constants chaque jour

Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour

Fruit d’une gigantesque ingénierie scientifique, technique et financière

Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière

Qui fait du Calcul Littéraire

Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire

Pour inventer des théories

Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories

Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)

A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)

A partir de toutes les données et connaissances de la terre et des humains

Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains

Qu’elle comprime et métamorphose

Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose

Sous un format de représentation inédit, unifié et propice au calcul

En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul

En un immense code informatique qu’elle a affiné elle-même

A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même

Où l’on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes

Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes

Grâce à une érudition et une capacité de raisonnement très supérieures aux nôtres

Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres

Dont nous comprenons chaque jour un peu plus la nature

Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature

Mais moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre

L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense
Dont l’architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre

L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Une drôle de machine immense

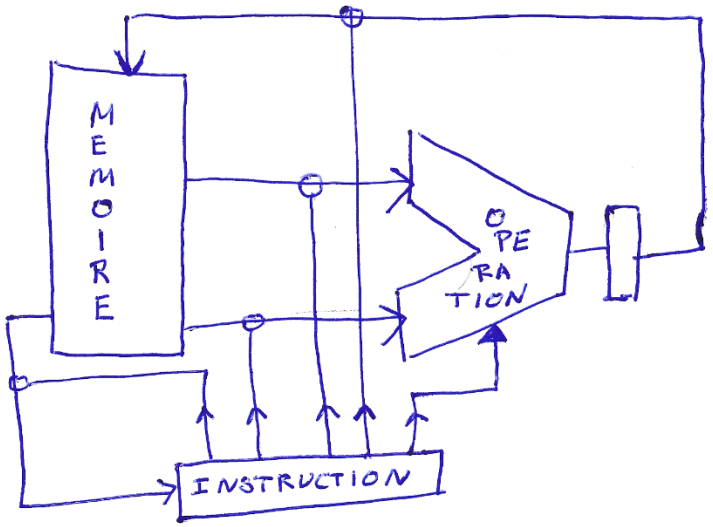
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

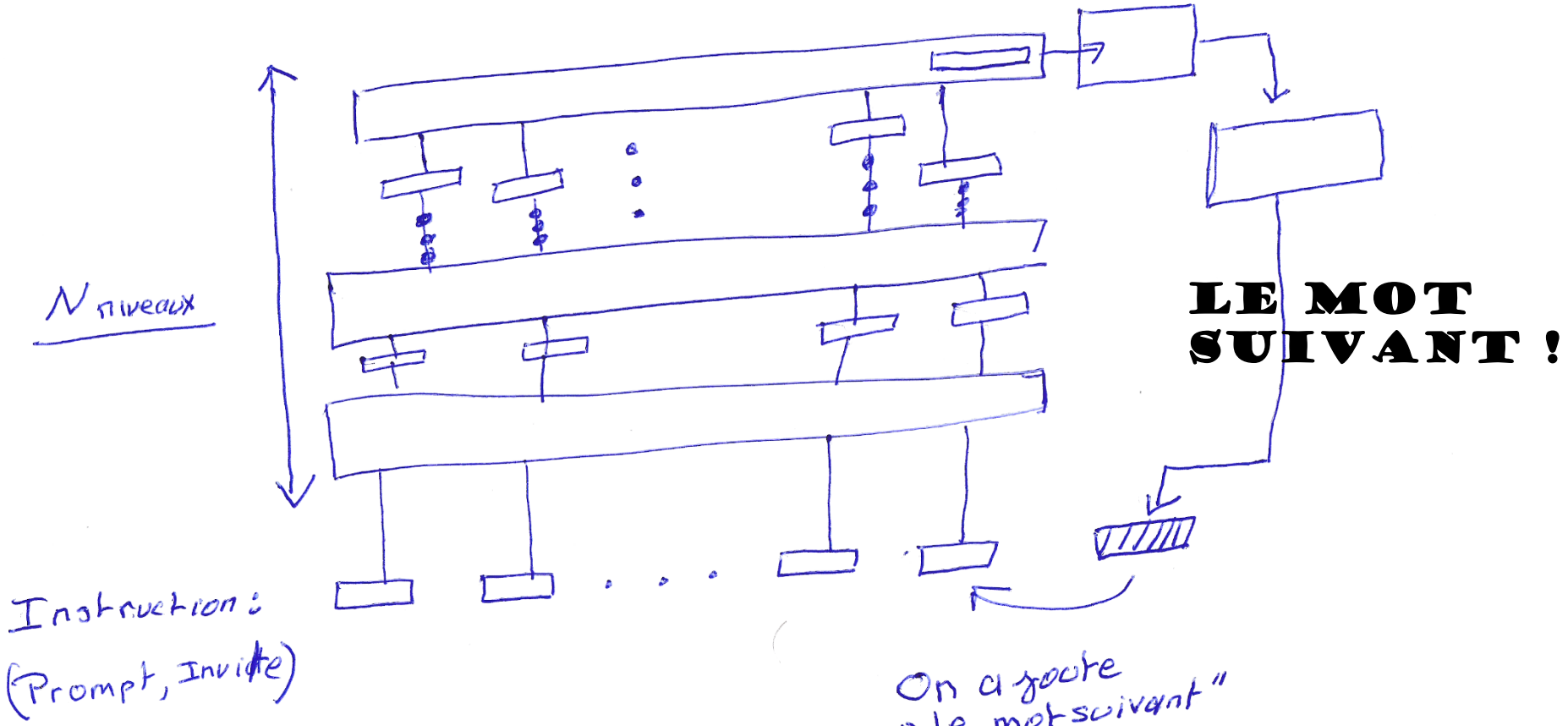
21 JUIN 1948
Machine de Von **Neumann**
1971: Intel 4004 2300 transistors
2026: 20 milliards de transistors
Instruction de 32 bits
Mémoire permanente de **32 Go**

2017 -- 2022
Machine de Sam **Altmann**
20 millions de milliards de transistors
Instruction de 1 milliard de mots de 16 bits
Zero mémoire permanente

Ou Machine de **Post-Neumann**

Machine de Von Neumann





LA MACHINE GLOBALE POST-NEUMANN

UNE CENTAINE DE NIVEAUX

CENT TETES DANS CHAQUE NIVEAU

L'INSTRUCTION (le prompt)

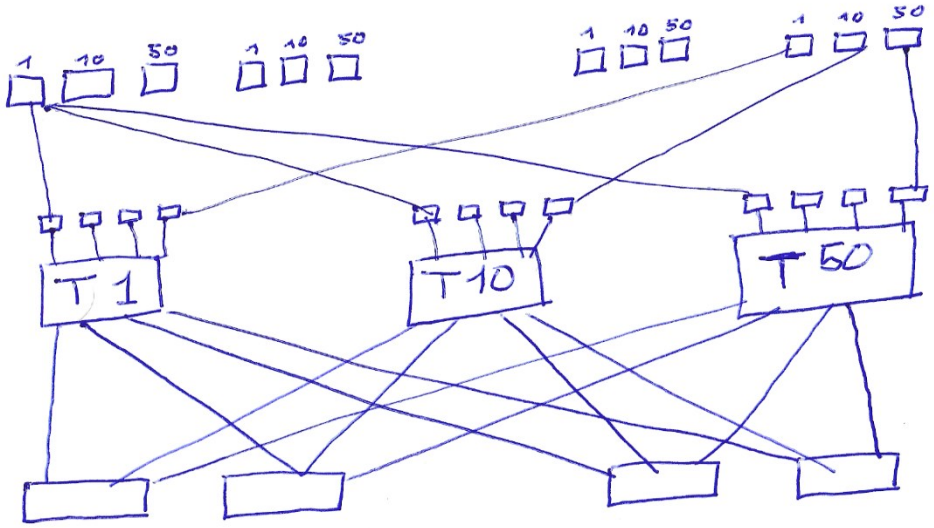
Traverse chaque niveau et chaque tête par niveau

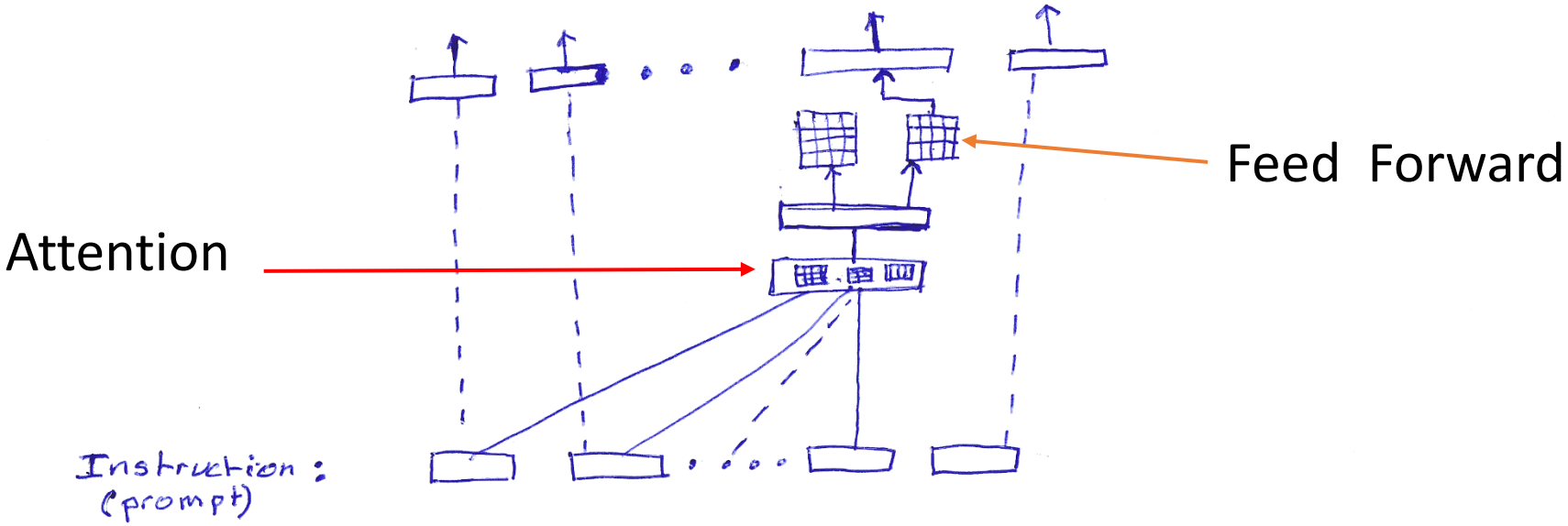
Chaque mot

Est traité par
Chaque Tête

Puis les résultats de chaque tête sont concaténés

Une centaine de têtes





UN ELEMENT DE TRANSFORMER (un par tête et par couche)

« THE » ATTENTION ITSELF!

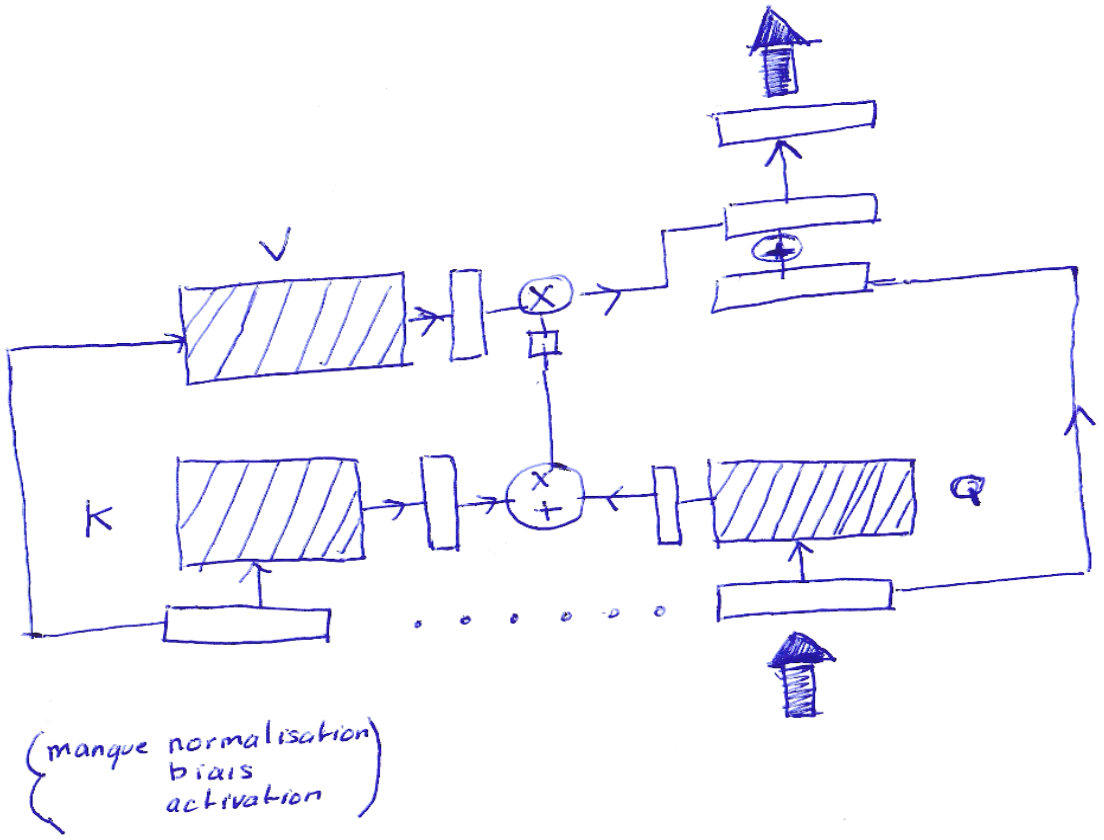
Q, K, V

Sont des matrices apprises

De véritables petites théories

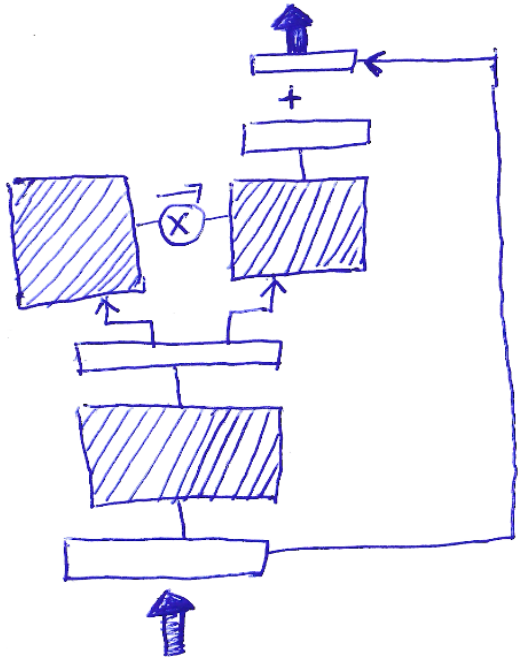
Qui s'efforcent de résoudre

La tâche confiée



FEED FORWARD

APRES ATTENTION



Une drôle de machine immense

Dont l’architecture a été découverte par hasard

En progrès constants chaque jour
Fruit d’une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d’une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu’elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu’elle a écrit elle-même
A l’intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l’Ordinaterre
L’Ordinaterre toujours plus peuplée de machines, d’algorithmes et de données.

Article « Transformers is all you need » 2017

« juste » inventé pour améliorer un traitement séquentiel des mots
-- trop lent (si prompt de 100 000 + mots)
-- oublieux du début si trop long

Passage du « jeu du téléphone » à « Internet »: chaque mot cause avec tout le monde et pas seulement au suivant: d’où parallélisme

Conséquence inattendue**: un vrai miracle**!

Une drôle de machine immense
Dont l'architecture a été découverte par hasard

En progrès constants chaque jour

Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Une nouvelle « Loi de Moore »:

Le cout d'un traitement diminue d'un facteur 10 par an

Il sort un grand modèle « frontière » par semaine

Des modèles de plus en plus gros tournent sur de simples PC

Exemple: 2 Novembre 2026: un article de Bengio et 35 coauteurs à lire absolument … de 186 pages

La même semaine Meta améliore le vieil algorithme de clustering vieux de 60 ans

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour

Fruit d'une gigantesque ingénierie scientifique, technique et financière

Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Il se publie XXXXX articles sur l'IA par jour

Il y a YYYYYY chercheurs / ingénieurs qui travaillent quotidiennement à améliorer l'IA

L'Effet VIETNAM

Unité de mesure du **CAPEX:** milliers de milliards de de dollars

Une immensité d'outils et d'entreprises qui couvrent tous les aspects de l'IA, des standards, des communautés, des services; mis en place en quelques années

En outils, algorithmes, matériels, logiciels, domaines de recherche, centres de données, supercalculateurs, circuits, bancs d'essai, indices de performance, etc etc etc

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, Jake Berdine, Gabriel Bernadett-Shapiro, Christopher Berner, Lenny Bogdonoff, Oleg Boiko, Madelaine Boyd, Anna-Luisa Brakman, Greg Brockman, Tim Brooks, Miles Brundage, Kevin Button, Trevor Cai, Rosie Campbell, Andrew Cann, Brittany Carey, Chelsea Carlson, Rory Carmichael, Brooke Chan, Che Chang, Fotis Chantzis, Derek Chen, Sully Chen, Ruby Chen, Jason Chen, Mark Chen, Ben Chess, Chester Cho, Casey Chu, Hyung Won Chung, Dave Cummings, Jeremiah Currier, Yunxing Dai, Cory Decareaux, Thomas Degry, Noah Deutsch, Damien Deville, Arka Dhar, David Dohan, Steve Dowling, Sheila Dunning, Adrien Ecoffet, Atty Eleti, Tyna Eloundou, David Farhi, Liam Fedus, Niko Felix, Simón Posada Fishman, Juston Forte, Isabella Fulford, Leo Gao, Elie Georges, Christian Gibson, Vik Goel, Tarun Gogineni, Gabriel Goh, Rapha Gontijo-Lopes, Jonathan Gordon, Morgan Grafstein, Scott Gray, Ryan Greene, Joshua Gross, Shixiang Shane Gu, Yufei Guo, Chris Hallacy, Jesse Han, Jeff Harris, Yuchen He, Mike Heaton, Johannes Heidecke, Chris Hesse, Alan Hickey, Wade Hickey, Peter Hoeschele, Brandon Houghton, Kenny Hsu, Shengli Hu, Xin Hu, Joost Huizinga, Shantanu Jain, Shawn Jain, Joanne Jang, Angela Jiang, Roger Jiang, Haozhun Jin, Denny Jin, Shino Jomoto, Billie Jonn, Heewoo Jun, Tomer Kaftan, Łukasz Kaiser, Ali Kamali, Ingmar Kanitscheider, Nitish Shirish Keskar, Tabarak Khan, Logan Kilpatrick, Jong Wook Kim, Christina Kim, Yongjik Kim, Jan Hendrik Kirchner, Jamie Kiros, Matt Knight, Daniel Kokotajlo, Łukasz Kondraciuk, Andrew Kondrich, Aris Konstantinidis, Kyle Kosic, Gretchen Krueger, Vishal Kuo, Michael Lampe, Ikai Lan, Teddy Lee, Jan Leike, Jade Leung, Daniel Levy, Chak Ming Li, Rachel Lim, Molly Lin, Stephanie Lin, Mateusz Litwin, Theresa Lopez, Ryan Lowe, Patricia Lue, Anna Makanju, Kim Malfacini, Sam Manning, Todor Markov, Yaniv Markovski, Bianca Martin, Katie Mayer, Andrew Mayne, Bob McGrew, Scott Mayer McKinney, Christine McLeavey, Paul McMillan, Jake McNeil, David Medina, Aalok Mehta, Jacob Menick, Luke Metz, Andrey Mishchenko, Pamela Mishkin, Vinnie Monaco, Evan Morikawa, Daniel Mossing, Tong Mu, Mira Murati, Oleg Murk, David Mély, Ashvin Nair, Reiichiro Nakano, Rajeev Nayak, Arvind Neelakantan, Richard Ngo, Hyeonwoo Noh, Long Ouyang, Cullen O'Keefe, Jakub Pachocki, Alex Paino, Joe Palermo, Ashley Pantuliano, Giambattista Parascandolo, Joel Parish, Emy Parparita, Alex Passos, Mikhail Pavlov, Andrew Peng, Adam Perelman, Filipe de Avila Belbute Peres, Michael Petrov, Henrique Ponde de Oliveira Pinto, Michael, Pokorny, Michelle Pokrass, Vitchyr H. Pong, Tolly Powell, Alethea Power, Boris Power, Elizabeth Proehl, Raul Puri, Alec Radford, Jack Rae, Aditya Ramesh, Cameron Raymond, Francis Real, Kendra Rimbach, Carl Ross, Bob Rotsted, Henri Roussez, Nick Ryder,

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière

Qui fait du Calcul Littéraire

Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Données structurées ou non structurées ?

Le texte est … **XXX** Structuré !!!

YYY a dit:

« Pour comprendre une chose isolée, le texte suffit »
« Pour comprendre beaucoup de choses ensemble, il faut passer à quelques chiffres »

L'IAG a trouvé un moyen de faire un vrai « calcul littéraire »

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière

Qui fait du Calcul Littéraire

Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Données structurées ou non structurées ?

Le texte est … **HYPER** Structuré !!!

René Descartes a dit:

« Pour comprendre une chose isolée, le texte suffit »
« Pour comprendre beaucoup de choses ensemble, il faut passer à quelques chiffres »

L'IAG a trouvé un moyen de faire un vrai « calcul littéraire »

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire

Pour inventer des théories

Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Les LLM ne sont pas des perroquets stochastiques

Ils contiennent un très grand nombre de théories qu'ils ont construites eux-mêmes

Et qui pour la plupart nous sont inconnues

Car ils les ont inventées (ou plus rarement réinventé)

Afin de résoudre les problèmes qu'on leur a soumis

« Prédire le mot suivant » est aussi monstrueux que « Produire des trombones »

A propos de monstre, **XXX** a dit:

L’ **YYY** est un monstre incompréhensible

A propos de monstre, **Blaise PASCAL** a dit:

L' **HOMME** est un monstre incompréhensible

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories

Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)

A partir de toutes les données et connaissances de la terre et des humains

Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Apprentissage = Contrainte + Liberté + Hasard

Construire une **machine FIXE** avec beaucoup de degrés de liberté

Lui donner beaucoup de couples **(entrée / sortie) d'UN problème** à résoudre

Lui donner une stratégie **d'amélioration**

Attendre un **MIRACLE**

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories

Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)

A partir de toutes les données et connaissances de la terre et des humains

Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Exemple: apprendre à construire des automobiles dans une usine







N FOIS



Le MIRACLE : CA MARCHE !


REVUE TECHNIQUE AUTOMOBILE

DACIA
Sandero II phase 1

Avant-propos 2
Notice d'utilisation 2
Avertissements et précautions 3

PRÉSENTATION DU VÉHICULE
Identification du véhicule 4
Identification des organes 5

DÉPANNAGE ET ENTRETIEN COURANT
Accès et levage 8
Dépannage 11
Contrôles et niveaux 15
Télécommande et batterie 19
Éclairage extérieur 21
Périodicité d'entretien 24
Visibilité et confort 26
Entretien du moteur 0,9 TCe 28
Entretien du moteur 1,2l 33
Maintenance des freins 38

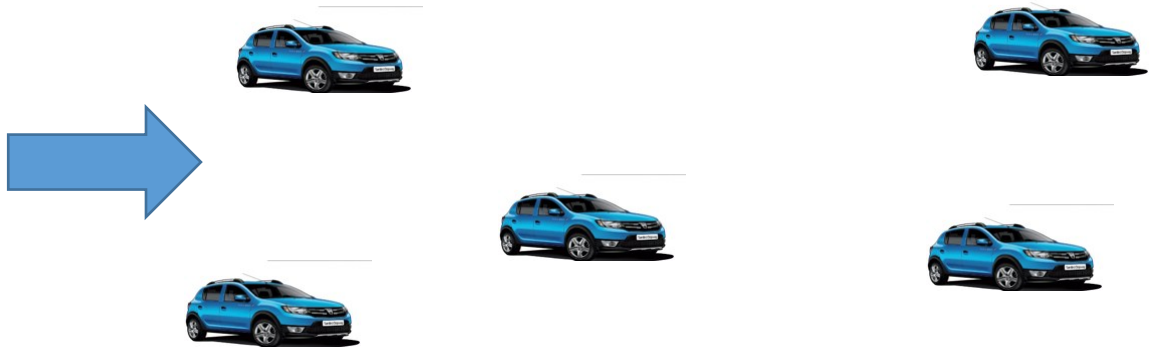
MÉTHODES DE RÉPARATION
Vidanges - Remplissages - Purges 44
Charge - Démarrage (moteur 0,9 TCe) 46
Charge - Démarrage (moteur 1,2l) 47
Moteur (0,9 TCe) 49
Moteur (0,9 12l) 54
Lubrification - Refroidissement moteur 0,9 TCe 61
Lubrification - Refroidissement moteur 1,2l 64
Alimentations - Dépollution (moteur 0,9 TCe) 68
Alimentations - Dépollution (moteur 1,2l) 73
Boîte de vitesses - Embrayage 77
Transmission 84
Géométrie - Trains roulants - Direction 84
Freinage 91
Chauffage - Climatisation 93

DIAGNOSTIC ET ÉLECTRICITÉ
Diagnostic 99
Protection électrique 100
Gestion moteur H4B 400 104
Gestion moteur H4B 408 121
Gestion moteur D4F 732 (euro 5) 139
Gestion moteur D4F 732 (euro 6) 152
Multiplexage 165
Antiblocage des roues - Contrôle de stabilité 166
Direction assistée 168
Airbags et prétensionneurs 169
Climatisation manuelle 173

CARROSSERIE
Protections inférieures 178
Boucliers 179
Optiques 180
Garnitures 181
Airbags 183
Amovibles 184
Vitrages - Mécanismes - Verrouillages 188
Éléments collés 191
Éléments soudés 192
Contrôles de la carrosserie 196

DONNÉES TECHNIQUES
Préconisations et capacités 198
Équipements électriques 198
Moteur 0,9 TCe 199
Moteur 1,2l 202
Lubrification - Refroidissement moteur 205
Alimentations - Dépollution 206
Embrayage - Boîte de vitesses 206
Transmission 206
Géométrie - Trains roulants - Direction 206
Freinage 208
Carrosserie 208

Copyright ETAI



L'apprentissage a **INVENTE** la mécanique, la thermodynamiqie, l'électromécanique …

Mais il y a un **SECOND MIRACLE ENCORE PLUS FORT !**

LE SECOND MIRACLE !

Pays de fabrication	France
Premier vol	1982
Turbine	2 Makila 1A1
Puissance	2x 1'742 ch
Poids à vide	4'350 kg
Capacité maxi de levage au niveau de la mer	4'500 kg
Personnes	4 pax et 1 pilote
Consommation de kérosène	620 l/h
Durée maxi de vol	2 h 30 min
Rayon d'action	625 km
Vitesse maxi	290 km/h
Vitesse de croisière	250 km/h
Altitude maxi d'opération	7'600 m au-dessus du niveau de la mer
Vitesse de rotation du rotor	268 tr/min

LE SECOND MIRACLE !

Pays de fabrication	France
Premier vol	1982
Turbine	2 Makila 1A1
Puissance	2x 1'742 ch
Poids à vide	4'350 kg
Capacité maxi de levage au niveau de la mer	4'500 kg
Personnes	4 pax et 1 pilote
Consommation de kérosène	620 l/h
Durée maxi de vol	2 h 30 min
Rayon d'action	625 km
Vitesse maxi	290 km/h
Vitesse de croisière	250 km/h
Altitude maxi d'opération	7'600 m au-dessus du niveau de la mer
Vitesse de rotation du rotor	268 tr/min











L’Architecture de départ doit être:

-- FIXE

-- Paramétrable

-- Apprenante (« dérivable »!)

-- Pas trop générale

-- Pas trop spécialisée

-- « Productive »

Les deux secrets du Miracle:

Le bon apprenti + Le bon apprentissage !

Le revers de la Médaille Miraculeuse :

On ne sait pas ce qui a été appris !

C’est le prix à payer pour rendre possible le Second Miracle

« Tu en sauras beaucoup plus, mais tu ne le sauras pas »

« je ne sais pas ce que je sais » (Socrate salue bien Aristote!)

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)

A partir de toutes les données et connaissances de la terre et des humains

Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Le langage naturel: un matériau intelligent fruit de « 100 000 » ans d'optimisation

En entrée: du langage , en sortie: du langage ! Ca devrait déteindre à l'intérieur!

Les L**L**M, c'est beaucoup plus des **connaissances** que du langage

Ca ne pouvait démarrer que chez Google

Question: quelle est la surface au sol d'un système contenant la totalité des données nécessaires à l'entrainement d'un grand modèle de langage ?

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains

Qu'elle comprime et métamorphose

Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

La taille du modèle (de l'usine) est fixée à l'avance

Il va falloir tout faire entrer dans le coffre!

L'énorme « pression » sur les mots va en **transformer** (!) totalement la structure

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose

Sous un format de représentation inédit, unifié et propice au calcul

En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

L'infinie complexité, l'hyper enchevêtrement, l'hyper malléabilité du langage naturel, on va les pulvériser
Dans une immense **uniformité de nombres**
Pour la rendre compatible avec la rigidité nanométrique et géométrique, orthonormé des circuits électroniques

La représentation distribuée de Hinton en 1986: **« Learning distributed representation of concepts »**

Concrètement, **tout** dans la machine IA va être représenté par « une dizaine » de milliers de nombres (de « 16 » bits)

Aussi bien les mot individuels que les expressions, les phrases, les concepts

Le « calcul littéraire » devient de l'arithmétique élémentaire (additions, soustractions) sur ces vecteurs de nom

L'**APPRENTISSAGE** accomplit le **MIRACLE** de **TRANSFORMER** le monde du langage en monde des nombres

Un nouveau monde qui nous est **« latent »** (« caché »)

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul

En un immense code informatique qu'elle a affiné elle-même

A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

L'**USINE** de départ est par définition **FIXE**, et il faut en apprendre les parties réglables (**VARIABLES**)

Notre usine fixe est en réalité un **programme informatique fixe écrit à la main**

Qui fait des opérations (additions, multiplications) sur de grands vecteurs

Certains vecteurs contiennent des **valeurs variables** décrivant le programme en entrée (le Prompt)

D'autres vecteurs contiennent des **valeurs constantes** qui ont été **apprises, affinées,** pendant l'apprentissage

Les **mots en entrée** sont remplacés par des **vecteurs de nombres**

Puis on exécute le programme (**Toujours le même**, instruction après instruction!)

Les **vecteurs de nombres** en sont remplacés par des **mots en sortie**

Et ça marche !

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même

Où l'on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Ca marche, et l'on ne sait pas pourquoi ! (*car on ne sait pas ce qui a été appris*)

La secte dit: ça marche parce que ça a été appris !

D'où toutes les « externalités négatives » de l'IA générative

Donc on va essayer de comprendre ce qui se passe vraiment à l'intérieur

C'est une « nouvelle science »: l'***Interprétabilité*** Chris Olah, cofondateur de Anthropic



Figure 2: Overview of the conceptual flow of the survey, categorised into four primary pillars: (1) Motivation & Concepts, covering fundamental knowledge and core hypotheses; (2) Technical Methods, detailing approaches developed and used in interpretability research; (3) Research Directions, exploring main directions the field is taking; and (4) Safety & Open Problems, identifying critical challenges and deployment-oriented interventions for AI alignment.

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même

Où l'on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

On découvre comment la machine a **décidé seule** de représenter:
-- les mots
-- la grammaire (du langage, des maths, des langages de programmation)
-- les concepts simples (« La Tour Eiffel », « C'est écrit en Anglais »)
-- les concepts sophistiqués (« Il s'agit d'un roman pour la jeunesse »)
-- des choses plus folles (« c'est de l'humour » , « ici le but est de tricher »)

On découvre par quels calculs elle construit ces différentes représentations

On découvre comment elle se débrouille pour représenter un graphe, une base de données, et de les exploiter

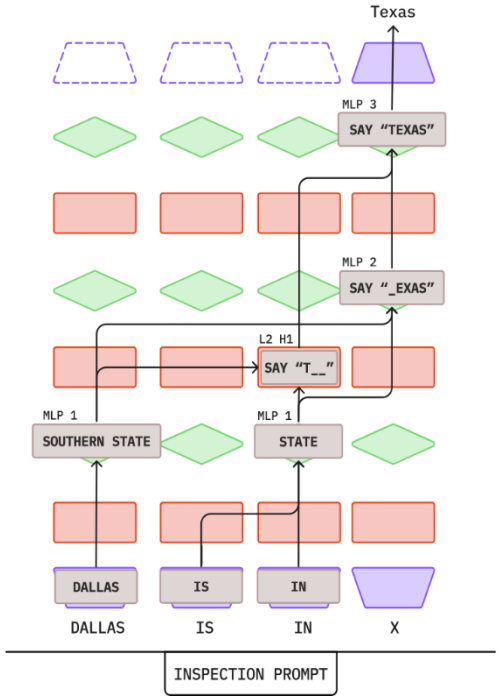
On découvre des **« circuits » d'opérations** qui s'enchainent pour effectuer des tâches complexes précises (par exemple proches de **l'unification logique** de Prolog) et qui constituent de **véritables sous-programmes**

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même

Où l'on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Un circuit



Lors de la longue phase d'apprentissage

Les circuits se battent entre eux pour leur espace vital

Certains disparaissent, d'autres surgissent

D'autres sont repoussés vers le bas par les autres

Une petite histoire de l'évolution alimentée par une centrale nucléaire

L' effet GROKKING

L'évolution de fait par saccades:

Un circuit « tombe en marche » et ça fait faire un bond aux performances du système

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même

Où l'on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Comparer deux choses
Choisir entre deux choses
Agréger des informations
Extraire des informations

Avec ces quatre opérations on sait « tout » faire dans notre « usine » organisée strictement « à la chaine ».

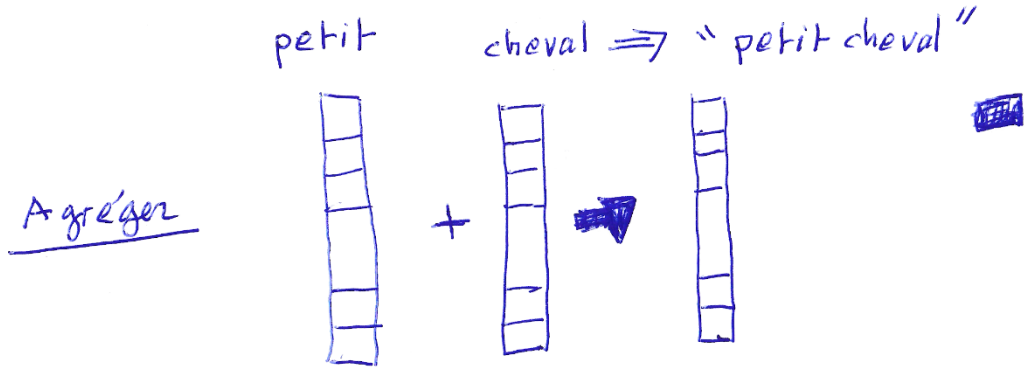
Et tout cela grâce à trois opérations élémentaires:

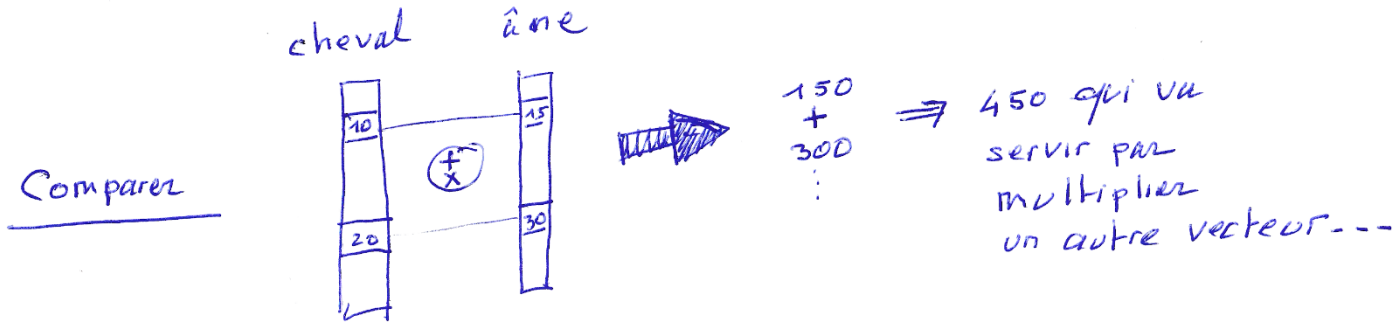
-- multiplier un vecteur par un nombre
-- faire la somme de deux vecteurs élément par élément
-- faire le produit scalaire de deux vecteurs

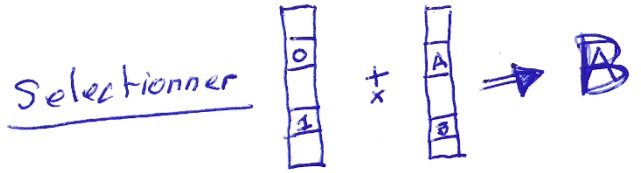
On retrouve les règles des systèmes experts, les bases de la programmation: IF THEN

OPERATIONS LOGIQUES

FAITES VECTORIELLEMENT







Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même

Où l'on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

On réalise des fonctions logiques via des opérations d'algèbre linéaire vecteurs, matrices, directions …°
Pourquoi ce **détour** ?

Parce que l'algèbre linéaire est dérivable, continue, ajustable, « apprenable »

Les réseaux de neurones profonds apprenables datent des années 80 (dont Yann Le Cun), pour les images, mais il leur manquait:
-- les sommes des vecteurs élément par élément (la représentation)
-- les multiplications entre vecteurs variables (le conditionnement, la logique)

Et ce **détour** fait, on s'efforce de retrouver au milieu de ces vecteurs de nombres des constructions symboliques et logiques (tables, dictionnaires, règles, programmes)

Il a fallu
1) oser traiter les textes comme des images (et TOUS les textes de la terre)
2) représenter des phénomènes complexes par des vecteurs « plats »
3) introduire la multiplication entre variables

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même

Où l'on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Et on retrouve donc la programmation classique!

On peut « aussi » programmer les Transformers A LA MAIN

En fixant explicitement chacun ou certains de leurs paramètres.

Il y a même des langages de programmation des Transformers (RASP°

On utilise même des LLM pour détecter et modifier une infime quantité de leurs paramètres pour modifier leur comportement (ça date de la semaine dernière)

Nombreux espoirs pour « aligner » (« lavage de cerveau ?») les LLM ou les rendre « plus surs »

Encore de nouvelles perspectives immenses qui s'ouvrent …

Et à propos de conscience ….

On lit des choses comme: « LLM INTERPRETERS OF THEMSELVES » !!!

5.3 The Privileged Access Hypothesis

A class of approaches that use language models as interpreters of *themselves* can be justified through the privileged access hypothesis. The privileged access hypothesis states that:

Privileged Access Hypothesis

A language model used to interpret itself can, under appropriate elicitation, access internal variables and intermediate computations that are not reliably recoverable from external behavioural analysis alone, providing a comparative advantage as an interpreter of its own behaviour (Binder et al., 2025; Li et al., 2025a; Song et al., 2026).

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain

Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes

Grâce à une érudition et une vitesse de raisonnement très supérieures aux nôtres
Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

-- Qu'est-ce qu'un LLM sait faire ?

-- On ne le sait pas!

-- Qu' est-ce qu'un LLM ne sait pas faire ?

-- On ne le sait pas !

Voir la complexité des « Prompts Systèmes », l'hyper abstraction et complexité de la « Constitution Anthropic »

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes

Grâce à une érudition et une capacité de raisonnement très supérieures aux nôtres

Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Un autre des **miracles** des Transformers:
Ils sont capables d'exploiter deux sources de connaissances:

-- celles apprises pendant leur apprentissage, et qui sont contenues dans leurs milliards de coefficients rangés dans leurs vecteurs constants.

Ceci est en train de tuer les moteurs de recherche comme Google.

-- celles découvertes sur l'instant dans leur prompt (donc jamais vues avant) (few shot learning)

Et de combiner les deux.

Ils agissent à la fois comme des compilateurs et des interpréteurs. Très singulier !

Une drôle de machine immense
Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a écrit elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes

Grâce à une érudition et une capacité de raisonnement très supérieures aux nôtres

Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès, comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.

Et très vite on leur ajouté des **boosters logiques**:

Une capacité explicite à raisonner située ***à l'extérieur*** du Transformer

Beaucoup moins apprise

Beaucoup plus programmée

En fait un véritable moteur d'inférence logique à la Prolog

C'est le COT (Chain of Thought) septembre 2024, l'apprentissage par renforcement

Puis on s'est mis à entrainer un LLM à apprendre à un autre comment raisonner …

Et nous ne parlons pas ici de l'IA agentique

OK, Jean, tu t'es pas mal débrouillé
Le temps défile, et arrêtons-nous là maintenant
Sans développer la fin de la phrase
Et à la prochaine fois!
C'est toujours un plaisir, etc, etc …
PS: tu m'expliqueras ce que tu appelles l'Ordinaterre!

Une drôle de machine immense

Dont l'architecture a été découverte par hasard
En progrès constants chaque jour
Fruit d'une gigantesque ingénierie scientifique, technique et financière
Qui fait du Calcul Littéraire
Pour inventer des théories
Par apprentissage surpuissant d'une tâche unique (essais, erreurs, améliorations)
A partir de toutes les données et connaissances de la terre et des humains
Qu'elle comprime et métamorphose
Sous un format de représentation inédit, unifié et propice au calcul
En un immense code informatique qu'elle a affiné elle-même
A l'intérieur duquel on redécouvre des mécanismes élémentaires du logos humain
Capable de résoudre une quantité infinie et indéfinie de questions et de problèmes
Grâce à une érudition et une capacité de raisonnement très supérieures aux nôtres

Dont nous comprenons chaque jour un peu plus la nature
Mais probablement moins vite que ses progrès
Comme un horizon qui avance sur l'Ordinaterre
L'Ordinaterre toujours plus peuplée de machines, d'algorithmes et de données.