

1 Traitement des données informatiques

● Définition d'un ordinateur

On appelle **ordinateur** un système capable de traiter de l'information selon un programme écrit dans un langage informatique, et comportant au minimum un **processeur** et une **mémoire vive** (Fig. 1).

● Les données informatiques

Les données informatiques que peut traiter un ordinateur sont de natures diverses : textes, sons, images mais également programmes informatiques. Elles sont toujours numériques, ce qui nécessite parfois un processus de **numérisation** (Fig. 2).

Exemple : La norme internationale ASCII définit la façon de coder les caractères d'un texte. Chaque caractère y est codé par un octet (valeur binaire codée sur 8 bits). La valeur 1000001 sera attribuée à A, 1000010 à B, etc. La taille du fichier (en nombre d'octets) est alors égale au nombre de caractères du texte.

Les programmes peuvent être transformés en fichiers exécutables qui ont l'extension .exe sous le système d'exploitation Windows ou .app sous Mac OS.

Les données informatiques sont stockées sur divers supports (disques durs, clés USB, etc.) et peuvent être transportées sur un réseau informatique (dont internet).

2 Les programmes informatiques

Pour élaborer des programmes, les programmeurs utilisent des langages informatiques traduits ensuite par la machine en suites de 0 et de 1.

Un programme informatique peut comporter jusqu'à plusieurs centaines de millions de lignes de code. La probabilité d'erreurs de codage (appelées **bogues** ou **bugs**) est importante. Cela peut parfois avoir de graves conséquences.

Il existe de nombreux types de processeurs. Certains langages de programmation sont spécifiques à un type de processeur et sont très difficiles à utiliser.

Les **langages de haut niveau** sont indépendants du matériel utilisé et beaucoup plus simples à utiliser et à déchiffrer, donc à déboguer*.

Exemples : Les langages Java, Python ou Scratch sont des langages de haut niveau. Autour de nous, on trouve de nombreux ordinateurs programmés par des programmeurs dans les sociétés qui les commercialisent et/ou des ordinateurs programmables directement par l'utilisateur grâce à des interfaces simples à utiliser (Fig. 3).

3 L'intelligence artificielle

● Définition

Le terme d'**intelligence artificielle** (IA) désigne un ensemble de théories et de techniques de programmation dont le but est de simuler l'intelligence humaine, que ce soit dans la capacité d'apprentissage, les perceptions ou la prise d'initiatives.

Exemples : Les résultats de recherches en intelligence artificielle permettent aujourd'hui à des ordinateurs de reconnaître des formes et des visages, de conduire une voiture, de traduire un texte, etc. (Fig. 4).

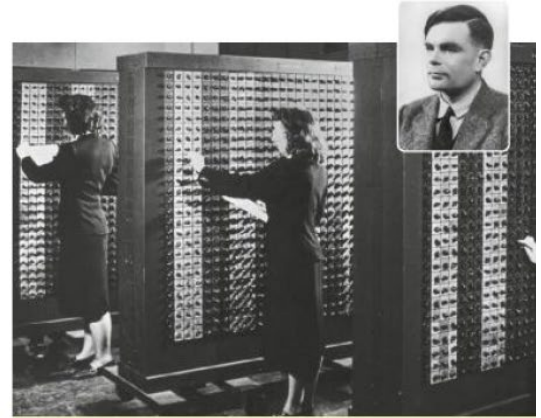


Fig. 1 : C'est le mathématicien Alan Turing (1912-1954) qui imagine le concept d'ordinateur moderne. L'ENIAC, qui pèse plus de 30 tonnes, est le premier ordinateur construit selon ses principes (1945).



Fig. 2 : Grâce à un scanner, des textes ou images peuvent être convertis en données informatiques, c'est-à-dire en données numériques.



◀ a Puce de GPS.



Fig. 3 : L'utilisateur peut (b) ou non (a) interagir avec l'ordinateur.

b Thermostat réglable.

Les recherches en IA n'en sont cependant actuellement qu'à leur début, aucun ordinateur actuel ne pouvant être qualifié encore de véritablement « intelligent ».

● L'apprentissage machine

L'apprentissage machine permet aux ordinateurs de réagir dans des situations inédites à partir de données d'entraînement sélectionnées par le programmeur.

Des méthodes mathématiques de repérage de tendance (corrélations, similarités) et de probabilités sont utilisées sur de grandes quantités de données (*big data*). Elles permettent ensuite d'effectuer des prédictions ou de prendre des décisions dans des situations nouvelles.

La qualité et la représentativité des données d'entraînement sont essentielles pour la fiabilité des réponses.

Si les données d'entraînement comportent un **biais**, les résultats donnés par l'ordinateur seront de mauvaise qualité, ce qui peut avoir de graves conséquences.

Certains usages de l'IA peuvent ainsi poser des problèmes éthiques. Des biais dans l'apprentissage pourraient par exemple entraîner des discriminations.

Exemples : Des algorithmes utilisant l'apprentissage machine peuvent être utilisés pour sélectionner automatiquement des CV pour un recruteur, calculer une probabilité de radicalisation pour la police ou de maladie pour une assurance.

● L'inférence bayésienne

L'inférence bayésienne est une méthode de calcul de probabilités de causes à partir des probabilités de leurs effets. Elle sert notamment à modéliser des relations au sein de systèmes complexes.

Exemples : L'inférence bayésienne peut servir à se prononcer sur la crédibilité de diagnostics dans les domaines médicaux et industriels ou l'analyse de spams.

Dans un test médical, si la fréquence de la maladie dans la population générale est f_m , alors le nombre de personnes réellement malades s'écrit (Fig. 5) :

$$\frac{f_m \cdot \text{fréquence(VP)}}{f_m \cdot \text{fréquence(VP)} + (1 - f_m) \cdot \text{fréquence(FP)}}$$



Fig. 4 : Spot, le chien robot de la société Boston Dynamics. Doté d'un bras dorsal, il est déjà utilisé pour explorer de façon autonome des zones dangereuses pour l'Homme.

Test	Malade	Sain
Positif	vrai positif (VP)	faux positif (FP)
Négatif	faux négatif (FN)	vrai négatif (VN)

Fig. 5 : Exemple de tableau de contingence utilisé dans l'approche bayésienne.



Le vocabulaire à retenir

- **Apprentissage machine :** technique qui permet à un ordinateur d'apprendre à réagir dans des situations nouvelles à partir de données d'entraînement entrées par un programmeur.
- **Biais :** dans l'apprentissage machine, erreur systématique dans les prédictions due à des données d'entraînement mal choisies.
- **Bogue (ou bug) :** erreur de codage entraînant un comportement de l'ordinateur aberrant.
- **Inférence bayésienne :** à partir de faits observés, cette méthode permet de définir la probabilité de causes possibles.
- **Intelligence artificielle :** techniques de programmation informatique visant à imiter l'intelligence humaine.
- **Langage de haut niveau :** langage informatique conçu pour être facilement compréhensible et utilisé par l'être humain.
- **Mémoire vive :** mémoire informatique présente dans un ordinateur et permettant de stocker temporairement des données.
- **Numérisation :** conversion d'une information (texte, image, etc.) en données numériques.
- **Ordinateur :** machine de traitement d'information comprenant au moins un processeur et une mémoire vive.
- **Processeur :** partie d'un ordinateur qui traite les données et exécute les instructions des programmes.
- **Tableau de contingence :** représentation de données dans un tableau selon deux critères, l'un en ligne, l'autre en colonne.