

Histoire de l'Intelligence Artificielle

3- naissance de l'Intelligence Artificielle

Frédéric Fürst - www.u-picardie.fr/~furst

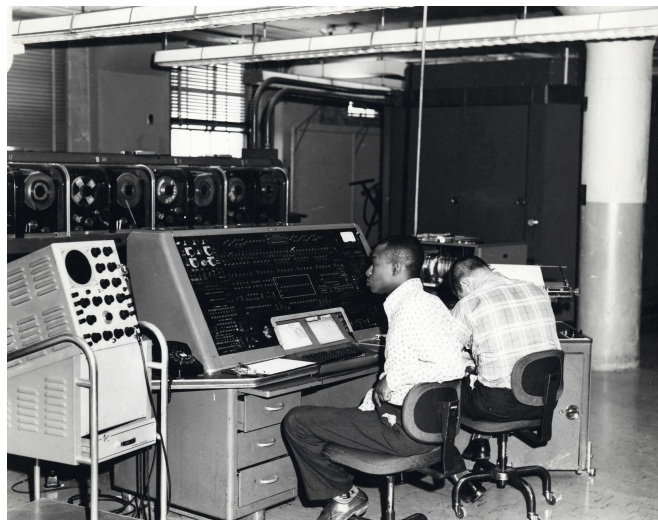
Les années 1940 et 1950 voient l'apparition des premiers véritables **ordinateurs**.

Ils sont Turing complets et **électroniques**, donc (relativement) rapides. Les entrées/sorties se font par cartes perforées et impression papier.

Il n'existe que très peu d'ordinateurs, uniquement dans quelques universités ou grandes entreprises.

Ces machines coûteuses servent surtout à faire des **calculs massifs** (statistiques pour l'état, calculs scientifique pour la recherche nucléaire, calculs balistiques pour l'armée, ...).

Par exemple, l'**UNIVAC I** (Universel Automatic Computer) est installé en 1951 au bureau du recensement américain.



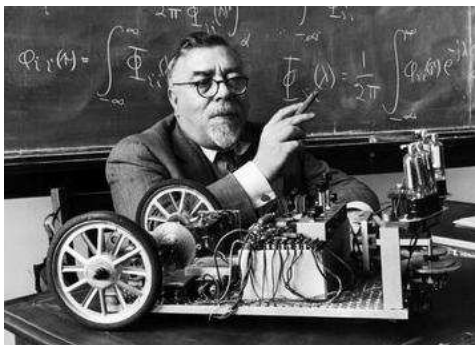
La machine **ORACLE** (Oak Ridge Automatic Computer and Logical Engine) est développée en 1954 au laboratoire de recherche nucléaire d'Oak Ridge.

L'informatique n'existe pas encore : les spécialistes des ordinateurs sont essentiellement des **mathématiciens**, ou des **électroniciens**.



Les langages de programmation au sens moderne du terme n'existent pas encore : le premier langage évolué, FORTRAN (FORmula TRANslator) ne verra le jour qu'en 1954. La programmation de telles machines est compliquée, très longue.

Mais l'apparition des ordinateurs semble rendre possible le rêve de l'IA. Les deux approches de l'IA vont émerger dans les années 1940 : le **connexionnisme** et le **cognitivisme**.



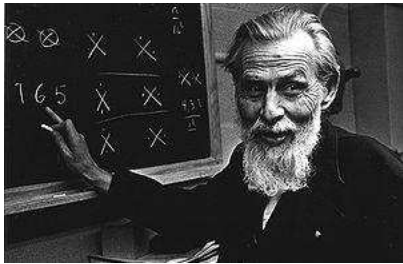
Norbert Wiener (1894-1964), mathématicien, lance dans les années 1940 la **cybernétique**, comme science du fonctionnement de l'esprit humain.

Wiener veut modéliser l'esprit comme une « boîte noire » dotée d'un comportement dépendant de mécanismes de **rétroaction** (feedback).

De 1942 à 1953, une dizaine de **conférences Macy** (du nom d'une fondation consacrée à la santé) vont se tenir à New York et réunir des mathématiciens, des neurologues, des psychologues ou des économistes avec pour but de dégager les principes de la cybernétique. Ces conférences vont surtout montrer à quel point l'esprit humain est une terra incognita.

Paradoxalement, alors que la modélisation globale de l'esprit comme boîte noire n'aboutit pas, certains participants à la conférence vont obtenir des résultats concernant la modélisation des constituants de base du cerveau, les neurones.

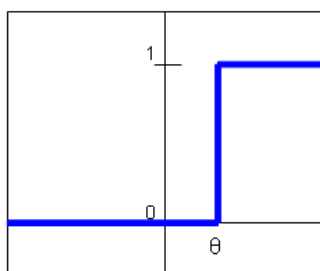
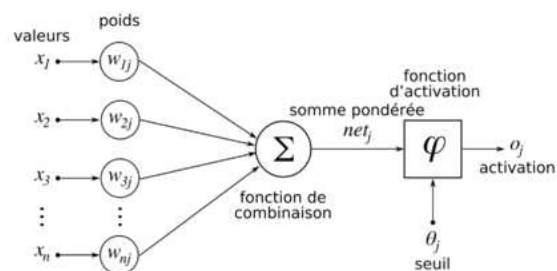
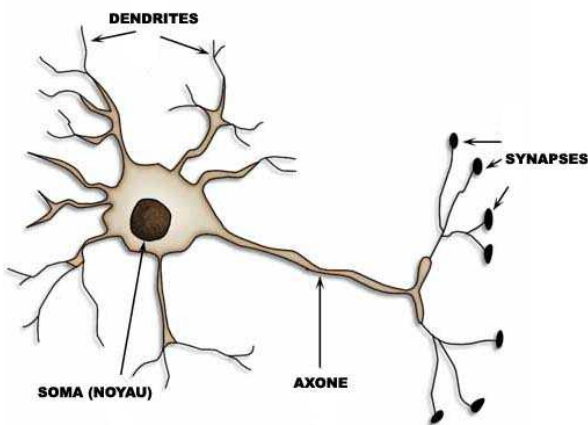
Warren McCulloch (1898-1969), et **Walter Pitts** (1923-1969), neurologues, inventent en 1943 le premier modèle mathématique du neurone biologique, le **neurone formel**.



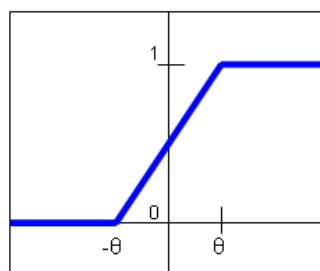
Une nouvelle approche émerge : le **connexionisme**. Puisqu'il semble hors de portée de modéliser le comportement global du cerveau, McCulloch et Pitts proposent de reproduire dans une machine le fonctionnement interne du cerveau humain.

Mais cette approche postule que le cerveau et l'architecture de ses centaines de milliards de cellules est mathématiquement modélisable, et qu'il puisse exister un ordinateur assez puissant pour faire tourner un tel modèle.

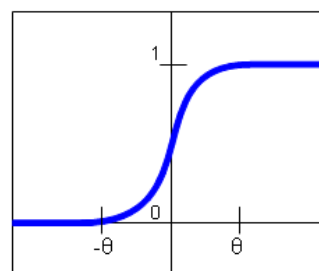
Un **neurone formel** modélise mathématiquement un neurone biologique



Fonction d'activation binaire



Fonction d'activation linéaire



Fonction d'activation sigmoïde

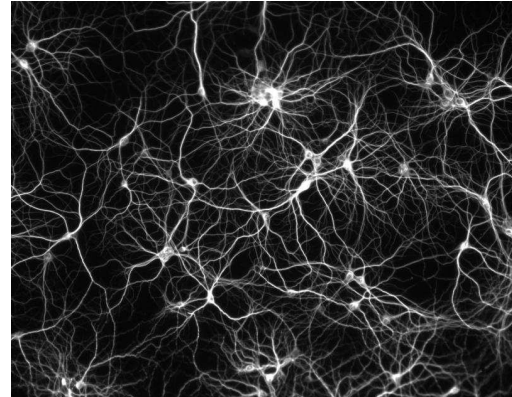


Donald Hebb (1904-1985), neuropsychologue, invente en 1949 la règle qui porte son nom et permet de doter les neurones formels de **capacités d'apprentissage** :

When an axon of cell A is near enough to excite B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased (The Organization of Behavior, 1949)

Ce principe qui explique la **mémoire** va être formalisé comme un processus de rétroaction dans les réseaux de neurones formels.

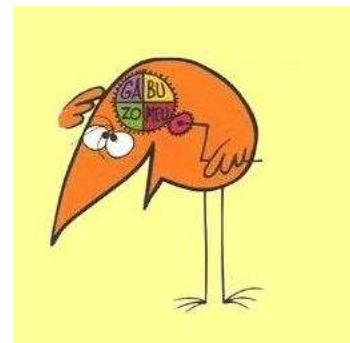
Mais cette approche « par le bas » (simuler le cerveau pour reproduire la pensée) semble moins prometteuse que l'approche « par le haut » (simuler la pensée en tant que manipulation symbolique).



Le **cognitivism** considère que la pensée peut être décrite à un niveau abstrait comme **manipulation de symboles**, indépendamment du support matériel de cette manipulation (cerveau, machine électronique, etc).

Cette approche établit un lien entre la pensée et le langage (système de symboles).

Une approche « par le haut » du cognitivism consiste à permettre aux machines de manipuler le langage, en supposant que cela les rendra intelligentes.



En 1949, Warren Weaver (1894-1978) écrit un mémorandum où il spécul, à la suite de Leibniz, sur l'existence de **règles linguistiques communes** à toutes les langues et veut appliquer ce principe à la **traduction automatique** sur ordinateur.

Dans le contexte de la guerre froide, la traduction automatique du russe en anglais ou de l'anglais au russe paraît cruciale.

En 1954, un premier programme, écrit à l'université de Georgetown permet de traduire plusieurs dizaines de phrases simples.

Le programme utilise 250 mots et seulement 6 règles de grammaire et tourne sur un IBM 701.



Exemples :

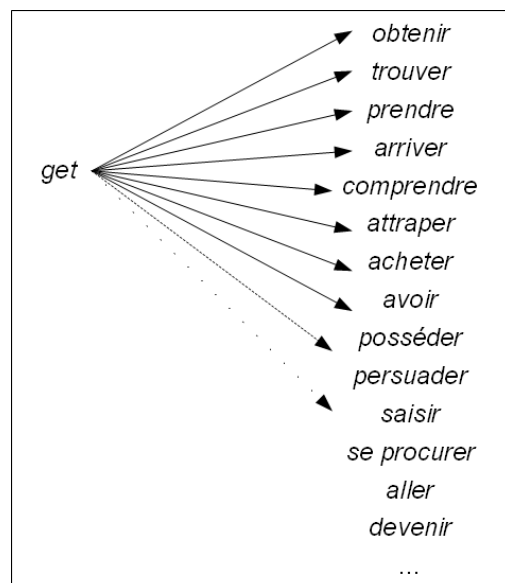
Mi pyeryedayem mislyi posryedstvom ryechyi → *We transmit thoughts by means of speech*

Vyelyichyina ugla opryedyelyayetsya otnoshenyiyem dlyini dugi k radiusu → *Magnitude of angle is determined by the relation of length of arc to radius*

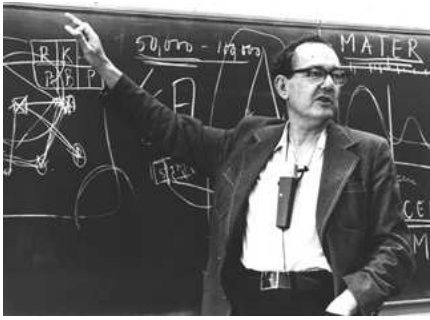
Des crédits sont rapidement alloués aux recherches sur la traduction automatique (aussi bien aux USA qu'en URSS). Les premiers travaux visent la **traduction directe**, presque mot à mot, à l'aide de dictionnaires bilingues et de règles simples.

Les problèmes de **polysémie** (amateur, blanc) ou d'**homonymie** (mousse, avocat, ...) apparaissent très rapidement.

La traduction directe de *The spirit is willing but the flesh is weak* (l'esprit est fort mais la chair est faible), après un aller-retour anglais-russe-anglais, devient *The vodka is strong but the meat is rotten* (la vodka est forte mais la viande est pourrie).



L'approche cognitive « par le bas » consiste à représenter dans la machine les mécanismes de la pensée, indépendamment de tout langage à l'aide de la **logique**. Le premier système logique de reproduction du raisonnement humain est le **Logic Theorist** de Newell et Simon.



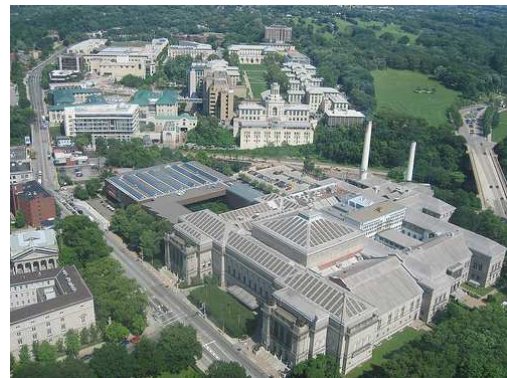
Herbert Simon (1916-2001) étudie l'économie à Chicago et y obtient son doctorant en science politique en 1943.

Simon cherche à comprendre comment les individus prennent des **décisions**, en particulier dans le cadre des organisations.

Simon introduit la notion de **rationalité limitée** : la rationalité d'un individu est limitée par son environnement, son histoire et par le manque d'information nécessaire à la prise de décision (*Administrative behavior*, 1947).

A partir de 1949 il travaille à l'université **Carnegie-Mellon** de Pittsburg sur le comportement des acteurs économiques et l'utilisation des ordinateurs pour à la fois comprendre et aider la prise de décision.

Simon recevra le prix Nobel d'économie en 1978, mais également le prix Turing en 1975, conjointement avec Allen Newell.



Allen Newell (1927-1992), après un Bachelor of Science en physique à Stanford, rejoint Princeton en 1949 pour mener une thèse en mathématiques.

Durant ses études, il a été fortement influencé par le mathématicien hongrois **Georges Polya** (1887-1985), qui avait introduit la notion d'heuristique pour la résolution de problème (*How to solve it*, 1945).

Une **heuristique** (du grec eurisko, *je trouve*) est une méthode empirique de résolution de problème, dont la validité ou l'efficacité n'est pas prouvée. Exemples : protéger la reine aux échecs, choisir la caisse où la file est la plus courte, ...

Trouvant finalement les mathématiques trop abstraites, Newell accepte en 1950 un poste à la **RAND Corporation** de Santa Monica, pour mener des travaux plus concrets, sur l'aéronautique de défense notamment.



Simon est également consultant à la **RAND (Research AND Development)** association de recherche à but non lucratif mais liée au complexe militaro-industriel américain.

La RAND, créée pour étudier la mise au point d'un satellite artificiel, va peu à peu étendre ses travaux à l'informatique, l'économie, la géopolitique, etc.



Les idées de **Simon** et de **Newell** convergent :

La **rationalité limitée** de Simon implique que la prise de décision repose sur des procédures permettant de palier aux manques d'information en tenant compte du contexte.

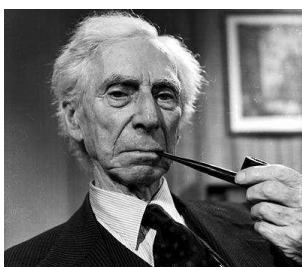
Pour Newell, ces procédures sont des **heuristiques**.

Simon et Newell considèrent que la condition *nécessaire et suffisante* pour qu'une machine puisse faire preuve d'intelligence est qu'elle soit un **système physique de symboles**.

Mais ils mettent au coeur de leurs travaux la notion d'**heuristique** : être intelligent, c'est aussi être capable de construire des heuristiques, de les tester, de les faire évoluer.

Aidés par un programmeur de la RAND, Cliff Shaw, ils développent **Logic Theorist** en 1956, un programme de démonstration automatique de théorème (voir l'article *The logic theory machine: A Complex Information Processing System*, 1956).

Newell et Simon s'inspirent des Principia Mathematica de **Russel** et **Whitehead**, des mathématiciens de Cambridge qui ont tentés de fonder toutes les mathématiques sur des bases logiques.



Logic Theorist utilise les **connecteurs** logiques NON (noté $\neg$) et OU (noté $\vee$). L'implication (notée $\Rightarrow$) se construit à partir d'eux car $P \Rightarrow Q$ équivaut à $\neg P \vee Q$.

Logic Theorist contient 5 **axiomes** :

$(P \vee P) \Rightarrow P$ (idempotence du OU)

$P \Rightarrow (Q \vee P)$

$(P \vee Q) \Rightarrow (Q \vee P)$ (commutativité du OU)

$(P \vee (Q \vee R)) \Rightarrow (Q \vee (P \vee R))$ (associativité du OU)

$(P \Rightarrow Q) \Rightarrow ((P \vee R) \Rightarrow (Q \vee R))$

Deux **règles d'inférences** sont utilisées :

Règle de substitution : si $A(P)$ est une formule vraie contenant la variable P , alors quelque soit la formule B , $A(B)$ est une formule vraie.

Règle de détachement : si A est une formule vraie et $A \Rightarrow B$ est une implication vraie, alors B est une formule vraie.

Comment Logic Theorist démontre t-il que $(P \Rightarrow \neg P) \Rightarrow \neg P$?

Il part du premier axiome : $(P \vee P) \Rightarrow P$

Il substitue $\neg P$ à P : $(\neg P \vee \neg P) \Rightarrow \neg P$

Il utilise la définition de l'implication : $(P \Rightarrow \neg P) \Rightarrow \neg P$

Tout le problème est de savoir quelles règles appliquer, et à quels endroits dans les formules, pour éviter des calculs inutiles et trop longs.

Newell et Simon introduisent donc des règles heuristiques supplémentaires pour guider le raisonnement, par analogie au comportement du mathématicien qui sélectionne certaines pistes pour démontrer un théorème.

Logic Theorist est considéré comme le premier programme informatique relevant du domaine de l'IA.

Logic Theorist **démontre 38 des 52 premiers théorèmes** du 2e chapitre des Principia Mathematica et découvre même pour le théorème 2.85 une preuve plus élégante que celle donnée par Russel et Whitehead.



Théorème 2.85 :

$$((P \vee Q) \Rightarrow (P \vee R)) \Rightarrow (P \vee (Q \Rightarrow R))$$

Newell et Simon envoie un article avec cette nouvelle démonstration au Journal of Symbolic Logic, en ajoutant Logic Theorist comme coauteur. Mais l'article est refusé au motif que ce théorème est déjà démontré depuis longtemps.

Simon écrit dans son autobiographie : « *nous avons inventé un programme informatique capable de penser de façon non numérique et, de ce fait, avons résolu le vénérable **problème de l'âme et du corps**, en expliquant comment un système composé de matière pouvait exhiber les propriétés de l'esprit* » (Models of My Life, H. Simon, 1991).

Les travaux de Newell et Simon illustrent également les apports croisés entre **informatique et IA** :

- le développement de l'informatique rend possible des expériences en IA
- les problèmes posées par les expériences en IA conduisent à produire des outils qui servent le développement de l'informatique

Pour faciliter la programmation du Logic Theorist, Newel, Simon et Shaw développent le langage **IPL** (Information Processing Language) en 1956.

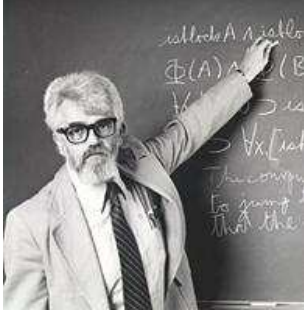
Ils inventent pour cela les **structures de listes**, permettant la manipulation d'éléments enchainés les uns aux autres, pour reproduire le caractère associatif de la mémoire humaine. Ces structures sont aujourd'hui utilisées dans tous les langages de programmation.

Un exemple de programme écrit en IPL :

L1	9-1	100
100	S4	101
101	S5	0
9-1	0	200
200	A1	201
201	V1	202
202	A2	203
203	V2	0

1956

McCarthy et Minsky se font aider par leurs aînés Shannon et Rochester pour obtenir de la fondation Rockefeller \$7500 afin d'organiser à l'université de Dartmouth, durant les 2 mois de l'été 1956, un **atelier de travail sur les machines pensantes**.



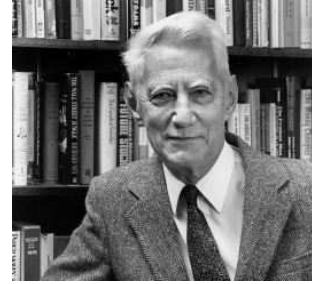
J. McCarthy
(Dartmouth College)



M. L. Minsky
(Harvard University)



N. Rochester
(I.B.M. Corporation)



C.E. Shannon
(Bell Telephone Laboratories)

Pour le texte complet de l'appel, voir : www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html



Marvin Minsky (né en 1927) se trouve en 1956 à l'université d'Harvard où il est chargé de cours.

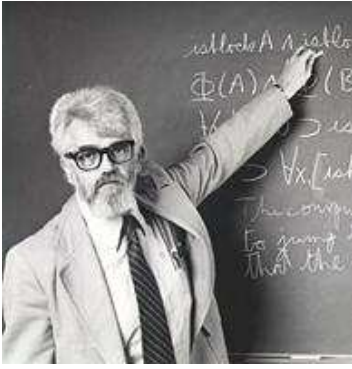
Après avoir suivi des études de mathématiques à Harvard (Bachelor of Arts en 1950), où il cotoie les premiers cybernéticiens, il a passé une thèse à Princeton intitulée *Neural Nets and the Brain Model Problem* (1954).

Son travail de thèse a porté sur la réalisation d'un **réseau de neurones** artificiels, en collaboration avec un doctorant en électronique, Dean Edmonds.

Leur machine SNARC (Stochastic Neural Analog Reinforcement Calculator) utilise 300 tubes à vide contrôlés par un pilote automatique de bombardier et simule le comportement d'un rat apprenant à se déplacer dans un labyrinthe.



Mais Minsky se tourne peu à peu vers l'IA symbolique, qui lui semble bien plus prometteuse. Il obtiendra le prix Turing en 1969.



John McCarthy (1927-2011) obtient un Bachelor of Science en mathématiques au California Institute of Technology en 1948, puis un doctorat à Princeton en 1951.

Sa thèse porte sur un certain type d'équations aux dérivées partielles, mais son passage à Princeton lui fait rencontrer Minsky avec qui il se découvre une passion commune pour l'idée de **machine pensante**.

En 1952, il travaille quelques mois aux Bell laboratories où il rencontre Shannon avec qui il lance un appel à papier sur le thème des machines pensantes. Mais le titre ambigu qu'ils utilisent, *Automata Studies*, n'attire pas les chercheurs intéressés par l'IA.

En 1956, il est chargé de cours au **Dartmouth College**, et propose à Shannon, Minsky et Rochester d'y organiser durant l'été une conférence qui portera explicitement sur les machines pensantes.

McCarthy recevra le prix Turing en 1971.



Claude Shannon (1916-2001) obtient en 1936 à l'université du Michigan un Bachelor of Science en ingénierie électrique et un autre en mathématique, puis il poursuit ses études au Massachusetts Institute of Technology.

Dans son mémoire de master, il établit le lien entre algèbre booléenne et circuits électriques (*A Symbolic Analysis of Relay and Switching Circuits*, 1937), et fonde ainsi l'**électronique numérique**.

Il travaille durant sa thèse sur la formalisation mathématique de la génétique (*An Algebra for Theoretical Genetics*, 1940).

Durant la guerre, il travaille aux Bell Labs, le département R&D de la compagnie American Telephone & Telegraph. Il y travaille sur la cryptographie et y élabore la **théorie de l'information** (*Mathematical Theory of Communication*, 1948).

Tout en continuant à collaborer avec les Bell Labs, il obtient un poste au MIT en 1956, et poursuit ses travaux sur la communication et le langage.



Nathaniel Rochester (1914-2001) obtient un Bachelor of Science en ingénierie électrique au MIT en 1941 et rejoint International Business Machine (IBM) en 1948.

Il y développe le premier **langage assembleur** symbolique.

Rochester sera ingénieur en chef pour toute la série des IBM 7xx, en particulier l'**IBM701** le premier **ordinateur généraliste** produit en série à partir de 1953 (19 exemplaires seront fabriqués).

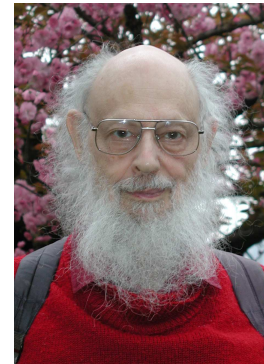


Il contribuera ponctuellement à plusieurs projets en IA (développement du langage LISP, du Geometry Theorem Prover)

6 personnes rejoignent les organisateurs de la conférence de Dartmouth :

Allen Newell et Herbert Simon

Ray Solomonoff (1926-2009) travaille au MIT et connaît Minsky et McCarthy depuis 1952. Il deviendra un pionnier de l'apprentissage automatique et inventera la notion de probabilité algorithmique.



Oliver Selfridge (1926-2008), un ancien étudiant de Wiener, travaille au MIT. Il sera le père de la reconnaissance de formes et un précurseur des systèmes experts.



Arthur Samuel (1901-1990) travaille à IBM sur la réalisation d'un programme de jeu de dames.

Trenchard More est un doctorant du MIT qui travaille sur la démonstration automatique de théorème

Les organisateurs de la conférence de Dartmouth prévoient que soient abordés diverses questions autour de l'idée de **machine pensante** :

Comment simuler la pensée et le langage au travers de règles formelles?

Comment faire penser un réseau de neurones?

Comment doter une machine de capacité d'apprentissage automatique?

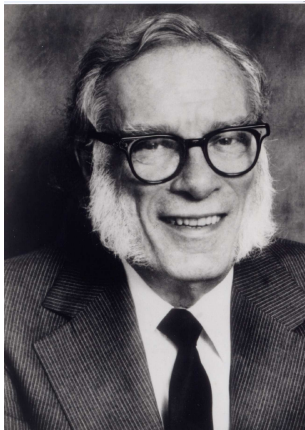
Comment doter une machine de créativité?



Mais les discussions seront assez limitées, et la conférence de Dartmouth n'aura servi qu'à forger un embryon de **communauté de recherche** autour des problématiques évoquées.

McCarthy propose le terme d'**intelligence artificielle** pour désigner la nouvelle discipline qui va connaître son âge d'or durant les 15 années qui suivent.

Parallèlement, le thème de l'IA émerge peu à peu dans la littérature et au cinéma.



Issac Asimov (1920-1992), surnommé *l'Empereur de la Science-Fiction*, traite, entre autres, du thème des robots humanoïdes intelligents au travers de dizaines de nouvelles et de romans.

En 1941, dans une de ces nouvelles, il invente le terme *Robotics*. Mais c'est la **psychologie humaine** qu'il met en scène au travers du comportement des robots.

Minsky a dit avoir été amené à l'IA par les histoires de robots d'Asimov, avec qui il est devenu ami.

Autre écrivain phare de la SF, **Philip K. Dick** (1928-1982) va aussi écrire dans les années 50 et 60 sur les machines pensantes, qu'il appelle *androïdes*.

Les androïdes servent à Dick à mettre en scène le problème des **simulacres** et des réalités fictives. Il a par exemple écrit la nouvelle *Les androïdes rêvent-ils de moutons électriques?* qui sert de scénario au film *Blade Runner*.

