

Histoire, épistémologie et concepts fondamentaux des sciences cognitives

Jérôme Sackur
DEC / LSCP

Rationalisme et Innéisme

- Le paradigme est:
 - empiriste mais contre « l'empirisme » au sens épistémologique
 - innéiste mais contre le dualisme

Les Sciences cognitives sont donc:

- rationalistes
- matérialistes
- expérimentales

La notion de représentation

- Le paradigme cognitif est **mentaliste**:
 - il y a des niveaux de descriptions des organismes ou systèmes qui impliquent des états ou représentations mentales
 - Ces représentations ne sont pas éphiphénoménales: elles ont une efficacité causale
 - Les représentations ont une réalité matérielle

L'épouvantail béhavioriste

- Le béhaviorisme (J. Watson, 1913, B. F. Skinner, 1938) est le repoussoir idéal:
 - empiriste
 - anti-représentationnaliste
 - (Mais pas réductionniste)
- Chomsky, 1959: « A review of Skinner's *Verbal Behavior* »

Parenthèse: qu'est-ce que le béhaviorisme?

- Psychologie et philosophie de la psychologie matérialiste et objectiviste...

... mais ...

- anti-mentaliste et empiriste (au sens épistémologique)
- lois de l'apprentissage par conditionnement

Principes du béhaviorisme (« à la Skinner »)

- Le conditionnement pavlovien
 - avant conditionnement A^0, US^+
 - conditionnement: $A+US$
 - après conditionnement A^+, US^+
- Le conditionnement opérant:
 - Avant conditionnement: répertoire d'opérants: ($\langle A, p_{t_0} \rangle$, $\langle B, q_{t_0} \rangle$, $\langle C, r_{t_0} \rangle$, $\langle D, s_{t_0} \rangle$, ...)
 - Conditionnement: $A \rightarrow +$
 - Après conditionnement ($\langle A, p_{t_1} \rangle$, $\langle B, q_{t_1} \rangle$, $\langle C, r_{t_1} \rangle$, $\langle D, s_{t_1} \rangle$, ...) avec $\mathbf{p_{t_1} > p_{t_0}}$
- L'histoire passée d'un organisme explique son comportement présent

La superstition chez le pigeon

A pigeon is brought to a stable state of hunger by reducing it to 75 percent of its weight when well fed. It is put into an experimental cage for a few minutes each day. A food hopper attached to the cage may be swung into place so that the pigeon can eat from it. A solenoid and a timing relay hold the hopper in place for five sec. at each reinforcement.

If a clock is now arranged to present the food hopper at regular intervals *with no reference whatsoever to the bird's behavior*, operant conditioning usually takes place. In six out of eight cases the resulting responses were so clearly defined that two observers could agree perfectly in counting instances. One bird was conditioned to turn counter-clockwise about the cage, making two or three turns between reinforcements. Another repeatedly thrust its head into one of the upper corners of the cage. A third developed a 'tossing' response, as if placing its head beneath an invisible bar and lifting it repeatedly.

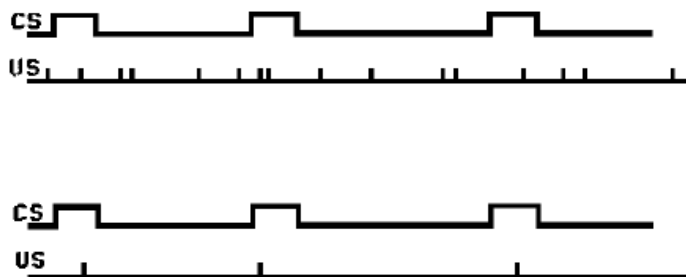
The experiment might be said to demonstrate a sort of superstition. The bird behaves as if there was a causal relation between its behavior and the presentation of food, although such a relation is lacking.

Skinner, 1948

Limites et mérites du béhaviorisme

- Semble démenti par les faits: observation learning, latent learning, etc.
- Non plausible mais non « falsifiable »; multiplication des hypothèses *ad hoc*.
- Théorie du conditionnement fausse: la contiguité spatio-temporelle ne fait pas le conditionnement. Exemple (le blocking):
 - $A^0, B^0, US^+ / A+US / A+B+US /$
 - Que croyez-vous qu'il arriva à B? B^0 ou B^+ ?

Figure 1
*Schematic of Two Conditioned Stimulus/
Unconditioned Stimulus (CS/US) Relations That
Share the Same Contiguity but Differ in the
Information the CS Gives About the US*



Rescorla, 1988

Le point d'appui formel: logique et informatique

Les grandes étapes:

- Du raisonnement au calcul (Frege, 1879: *Idéographie*)
- Du calcul au mécanisme:
 - La machine logique (Turing, 1936)
 - L'ordinateur (Von Neumann, 1945)
- Du mécanisme au cerveau (McCulloch et Pitts, 1943)

Le raisonnement comme calcul: langage et système formels

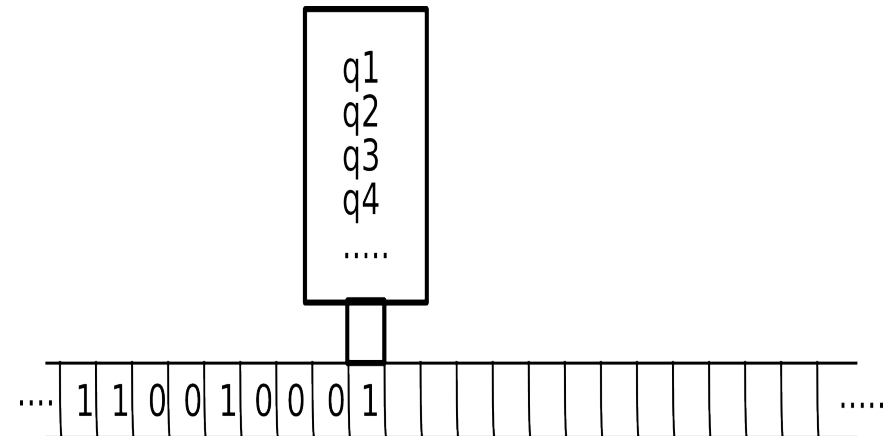
- Frege, 1879: un langage calculatoire pour des preuves « sans lacune ».
- Exemple, le *Modus Ponens*: « S'il fait beau, je sors le parasol; or il fait beau; donc je sors le parasol ».

$$\frac{\mathcal{X} \quad \mathcal{X} \rightarrow \mathcal{Y}}{\mathcal{Y}}$$

La machine logique: Turing, 1936

« Nous pouvons comparer un homme en train de calculer un nombre réel à une machine capable de se trouver seulement dans un nombre fini d'états q_1, q_2, \dots, q_R que l'on appellera « configurations- m ». La machine dispose d'un « ruban » (analogue à une feuille de papier) qui défile en elle et qui est divisé en sections (appelées « cases ») dont chacune est susceptible de recevoir un « symbole ». A un moment quelconque, il n'y a qu'une case, disons la r -ième, portant le symbole $S(r)$, qui soit « dans la machine ». Nous pouvons l'appeler la « case inspectée ». Le symbole inscrit dans la case inspectée est le « symbole inspecté ». Le « symbole inspecté » est le seul dont la machine soit, pour ainsi dire, « directement consciente ». (...) Le comportement possible de la machine à un moment quelconque est déterminé par la configuration- m q_n et le symbole inspecté $S(r)$. [Ce comportement se borne à l'écriture d'un symbole, son effacement, un changement de configuration- m , et/ou un déplacement de la bande.] (...) Je soutiens que ces opérations comprennent toutes celles que l'on utilise pour le calcul d'un nombre. »

- Tout ce qui est calculable est turing-calculable (thèse *spéculative* de Church-Turing).
- Il existe une machine de Turing universelle (théorème)



L'ordinateur de von Neumann

- Pas un simple calculateur...
- ... mais approximation **finie** d'une machine de Turing universelle
- Problèmes d'architecture
- Problème physique de la « discrétisation »

L'ordinateur pour le paradigme cognitif: un système physique symbolique

A physical symbol system consists of a set of entities, called symbols, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure). Thus a symbol structure is composed of a number of instances (or tokens) of symbols related in some physical way (such as one token being next to another). At any instant of time the system will contain a collection of these symbol structures. Besides these structures, the system also contains a collection of processes that operate on expressions to produce other expressions: processes of creation, modification, reproduction, and destruction. A physical symbol system is a machine that produces through time an evolving collection of symbol structures. Such a system exists in a world of objects wider than just these symbolic expressions themselves.

A. Newell, H. Simon « Computer Science as empirical inquiry », 1976

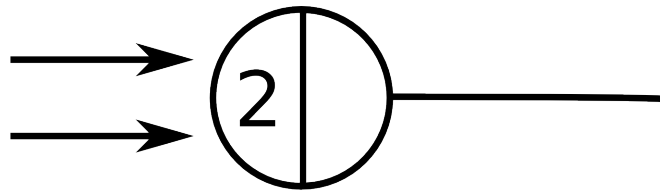
L'ordinateur: une métaphore? le fonctionnalisme

We wish to emphasize that we are not using the computer as a crude analogy to human behavior—we are not comparing computer structures with brains, nor electrical relays with synapses. Our position is that the appropriate way to describe a piece of problem-solving behavior is in terms of a program: a specification of what the organism will do under varying environmental circumstances in terms of certain elementary information processes it is capable of performing. This assertion has nothing to do—directly—with computers. Such programs could be written (now that we have discovered how to do it) if computers had never existed.²

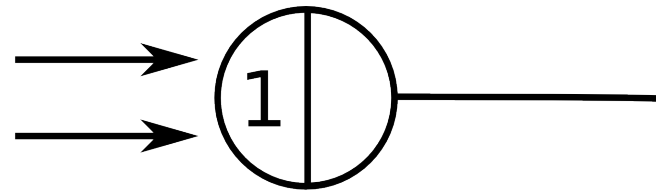
Newell, Shaw, Simon, 1958: « Elements of a theory of human problem solving »

Le cerveau comme machine logique: McCulloch et Pitts, 1943

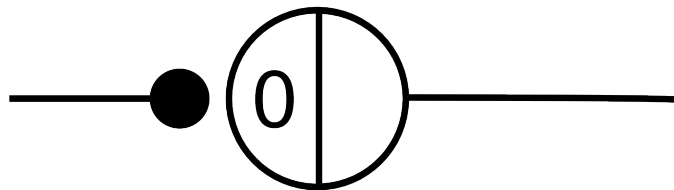
- Les neurone formels:
 - machines « tout ou rien »
 - en réseaux logiques



Connecteur "et"



Connecteur "ou"

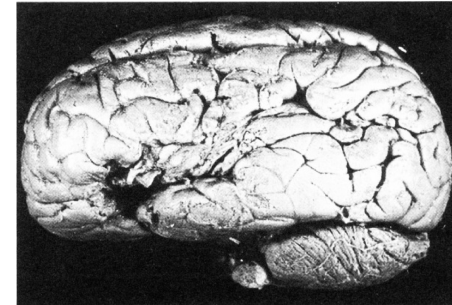


Connecteur "non"

Un résumé?



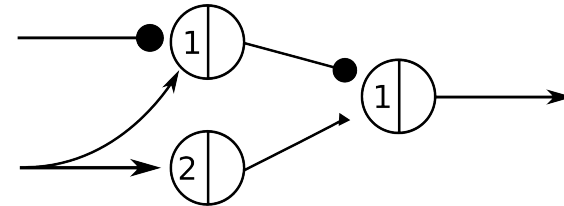
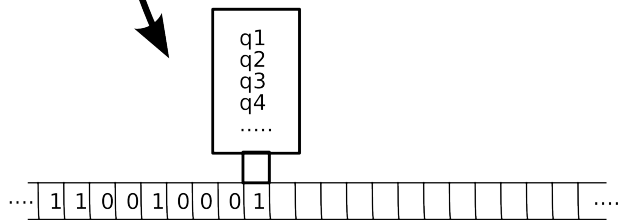
Systemes physiques



Machines logiques

S'il fait beau je vais à la plage
Or il fait beau
Donc je vais à la plage

Réseaux de neurones



Objets mathématiques



A quoi sert la logique?

- Au raisonnement

A quoi sert la logique?

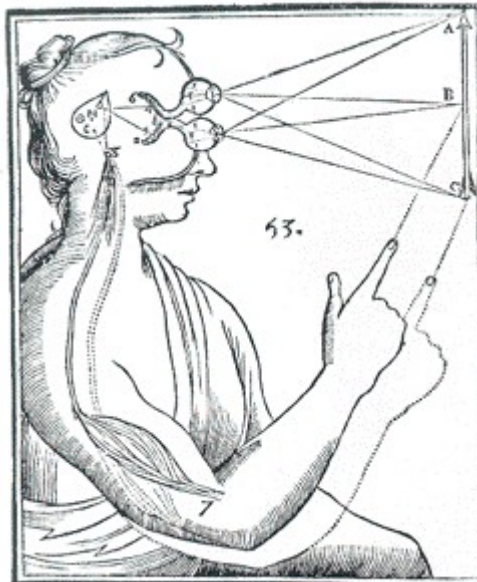
- Au raisonnement
- Au langage (Chomsky)

A quoi sert la logique?

- Au raisonnement
- Au langage (Chomsky)
- A la perception (Helmholtz, 1865): percevoir est interpréter les signes sensorielles au moyen d'inférences *inconscientes*

Trois « mécanismes »

- Le mécanisme de l'automate (Descartes, Pavlov, Watson)
- Le mécanisme flexible (Skinner)
- Le mécanisme des machines logiques: le cognitivisme

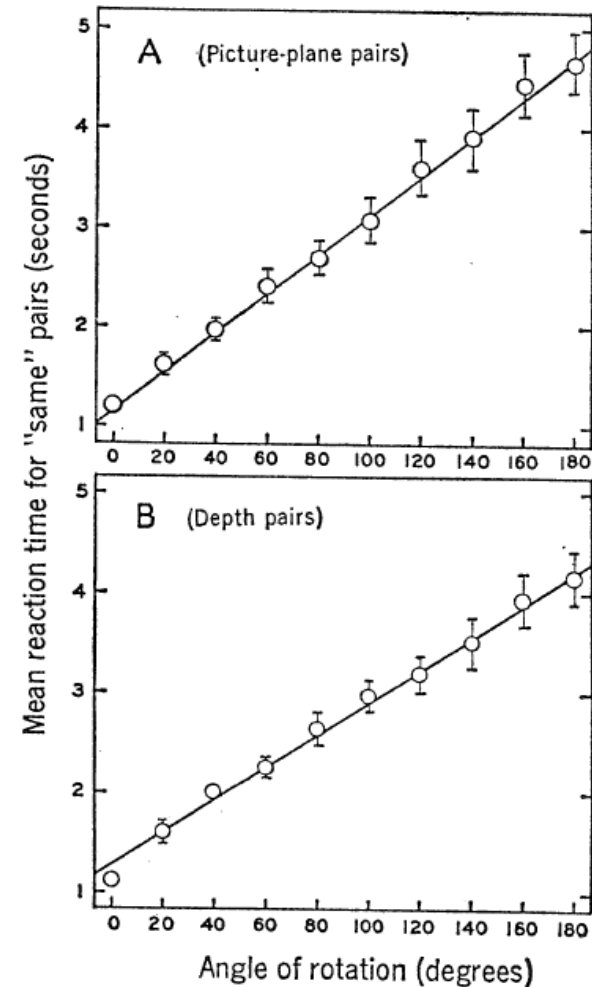
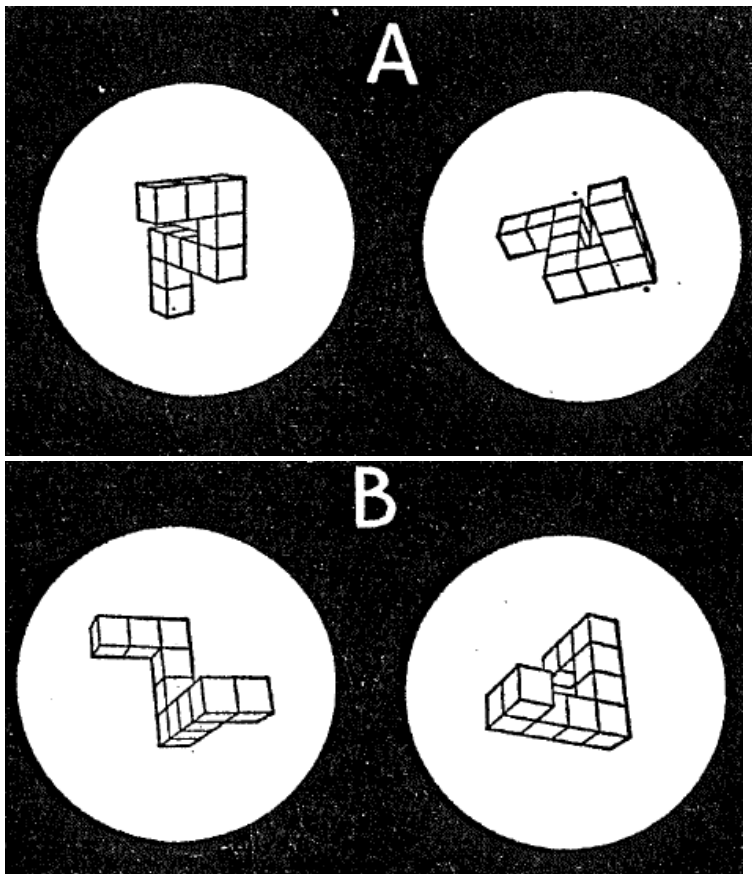


Limites du paradigme « classique »

- ambition généraliste: Newell et Simon « general problem solver »
- indifférence au stimulus: de la représentation à l'information
- Excès « logicistes » (computationalistes)
- au contexte biologique de l'évolution de l'esprit

Un exemple: l'imagerie mentale

- Manipule-t-on les images mentales de manière **analogue** à ce dont elles sont les images?



Shepard et Metzler, 1971

Mi dankas vim multe!