Interdisciplinary Cognitive Science

Discipline-based and interdisciplinary research

What does a discipline contribute to the development of science?

What are the purposes of working beyond disciplinary boundaries?

What risks/costs are born by interdisciplinary pursuits?
Chomsky’s Review of Skinner’s *Verbal Behavior*

Emphasis on the novelty of linguistic constructions

Inadequacy of probabilistic models (Markov processes) and need for a **generative system governed by rules**

**Poverty of the stimulus**: from data too impoverished to support behaviorist learning, young children learn their language
Setting the Task for a Grammar

There are an infinite number of grammatically correct sentences in a natural language (English, French)

Want a finite account (using recursion) that can generate all and only the grammatically well-formed sentences of the language (e.g., something that could be run on a computer)

- Hypothesize a grammar
- Determine what would be legitimate sentences given that grammar
- Test whether those are in fact grammatically correct sentences of the language
- If not, revise the grammar

Early Chomsky: Tree Structures and Transformations: Wh Fronting

I can eat what

NP

V

VP

Aux

VP

N

S

Comp

S

Comp

S

I

can
cat

what

NP
Competence vs. Performance

Chomsky uses ordinary speakers (e.g., himself) to test his grammars.

But ordinary speakers make grammatical mistakes all the time.

Proposes that these are due to performance limitations.

Claim: we all possess perfect linguistic competence, and hence can evaluate sentences even if our performance is flawed.

Universal Grammar and Nativism

The underlying grammatical processes are same for all languages.
   The implementation differs (different parameters).
   Thus, grammar is universal.

Grammar is too difficult to learn in restricted time given the linguistic evidence available to children (Poverty of the Stimulus).
   Universal Grammar must be innate.
   Children only have to figure out which implementation is found in their language.
The Impact and Continuing Legacy of Chomsky

The idea that language could be characterized in terms of rules specifying operations on symbols inspired psychological research on cognition.

The development of grammars, especially ones more oriented toward processing (e.g., Augmented Transition Networks or ATNs) contributed to ongoing psychological research on language processing and language learning.

Miller and the Psychological Reality of Grammar

Chomsky’s arguments about the inadequacy of finite state automata for constructing grammars adequate to natural languages and the need for transformational grammars led Miller to redirect his program.

Does transformational grammar characterize language processing?
- Early evidence that processing difficulty corresponded to number of transformations in sentence’s derivation

Subsequent evidence undermined this direct inference from grammar to processing

But the metaphor of cognition operating on structures remained compelling
The Turing Machine

Turing took as his model for a machine capable of computing all decidable functions humans who were called computers.

Individuals who calculated sums on paper for a living

Turing Machine: finite state device operating on a potentially infinite tape

- Argued that there was a Turing machine for any decidable function
- And a (Universal) Turing machine that could realize any actual Turing machine

Introducing the Digital Computer

- The first general function digital computer, Electronic Numerical Integrator and Computer (ENIAC) was developed by John Mauchly and J. Presper Eckert at the Moore School of Electrical Engineering, U. Penn
  - Purpose: calculate artillery firing tables
  - Went into operation too late for WWII (15 February 1946)
  - First used for problems related to design of the hydrogen bomb
  - Decommissioned on 2 October 1955
- John von Neumann advanced the idea of a stored program in his proposal for EDVAC
  - Von Neumann architecture
Inspiration for Linking Computers with Thought: Logic

George Boole: 1854: *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probability*: Natural deduction as a model of thought

1. A v –B : Premise
2. (-B & C) ⊃ D : Premise
3. C & -D : Premise
| 4. -A : Assump
| 5. -B : 1, 4 v-elim
| 6. C : 3 &-elim
| 7. -B & C : 5, 6 &-intro
| 8. D : 2, 7 ⊃-elim
| 9. -A ⊃ D : 4, 8 ⊃-intro

Logic Theorist: Computer program to prove theorems of logic

Human Problem Solving

Take verbal protocols as humans solve problems such as those of cryptoarithmetic

\[
\begin{array}{c}
\text{DONALD} \\
+ \text{GERALD} \\
\text{ROBERT}
\end{array}
\]

D=5

or Tower of Hanoi

Figure out general strategies that would enable computer to perform these tasks

Importance of means-ends reasoning and reasoning backwards
Production Systems

Rules
- If (A & B) → -A & +D
- If C → -C & +D & +E
- If (B & D) → -D & +J
- If (G & J) → -J & +A

Working Memory

G
B
D
H

Working Memory

G
B
J
H

The Physical Symbol System Hypothesis

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).

A physical symbol system is a machine that produces through time an evolving collection of symbol structures.

A physical symbol system has the necessary and sufficient means for general intelligent action.

Allen Newell and Herbert Simon, 1975
Advances in Symbolic AI

Success of limited purpose AI programs

Recognition of need for larger-scale representational structures to model real-world cognitive activities such as understanding stories

Oliver and Cleo went to Tony’s. Cleo slipped the maitre d’ a $20, and they were directed to a very nice table. They considered the entrées on the menu, but decided to order the salmon special. They asked for the salmon to be well cooked. They waited a long time for their dinner to come, and consumed nearly all their wine while waiting. When the waiter brought the salmon, it was nearly raw. They complained to the waiter but he insulted them for their unsophisticated taste. They finished their entrées, but decided to skip desert. They left a very small tip.

Answering Questions

On this evening:
Were Oliver and Cleo seated at their table?
Did the waiter bring them menus?
Did they read them?
Did they order a bottle of wine?
Did they eat the salmon?
Did they pay the check?
Were they unhappy when they left?

Even though this information was not stated in the story, all of us are able to answer these questions.
Roger Schank’s Restaurant Script

Schank proposed that we reason about such problems using larger-scale knowledge structures, into which we fit the information we are given. They specify what typically happens in events such as going to a restaurant. In addition to typical props, roles, entry conditions, etc. they are comprised of a sequence of primitive actions such as:

S MTRANS signal to W
W PTRANS W to table
S MTRANS 'need menu' to W
W PTRANS W to menu (from the coffee shop track)

Scripts contain tracks for common variations, such as going to a fast food restaurant, coffee shop, fancy restaurant.

Do Machines Really Think?

1. In one room ... a machine
2. In another room ... a human
3. In a third room ... the interrogator...
Implications of AI for Understanding the Human Mind

The mind is a symbol processing system
It manipulates symbolic structures in accordance with rules
The mind's native symbols constitute a language—the language of thought
This language must be innate—all learning depends upon constructing and testing hypotheses
Evidence for the language of thought: thought is
- productive
- systematic
Only a system with a composition syntax and semantics will exhibit these properties

Opposition to AI: Searle and the Chinese Room

Imagine yourself as a monolingual English speaker locked in a room.
You are given three sets of paper on which strange inscriptions are written.
You are also given some directions in English.
Following the directions, you match the first set of inscriptions with the second, and the third with the first two, and produce a sequence of inscriptions and slide these through a slot in the door.

You follow the directions much as a computer follows its program
Carrying on a Chinese “Conversation”

Unbeknownst to you, the symbols you were given and which you produced were Chinese.

The first set of symbols in fact constituted a script
The second constituted a story
The third constituted questions
By operating on these symbols following the English rules (match the top symbol of the second set with one in the first set), you were able to give cogent answers to the questions about the story

Native Chinese speakers outside believe they are conversing with a fellow Chinese speaker. *The Turing Test is passed!*
Implications of the Chinese Room

The Chinese speakers were wrong that they were having a conversation with anyone in Chinese—you don’t know Chinese.

But you were doing just what the computer running Schank’s program would do!

So it doesn’t understand either. It is not intelligent, and does not constitute a mind.

Challenge: what would it take for a machine to use symbols meaningfully?

A non-Symbolic Alternative: Artificial Neural Networks (Connectionism)

McCulloch and Pitts (1943) saw how to build sentential logic networks out of artificial neurons: negation, and-gates, or-gates

Pitts and McCulloch (1947) saw the potential to model perception, etc. with less structured networks
Rosenblatt’s Perceptrons

Artificial Retina  Associator Units  Response Units

Minsky and Papert and the Demise of Perceptrons

Exclusive Or

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>A xor B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Failure of linear separability
Connectionism Returns

New learning rules: Delta Rule and Backpropagation

Solving XOR with Backpropagation
Nets learn to talk: NETtalk

Corpus presented to network

Started with random weights

Error backpropogated through network to adjust weights

Figure 1: Schematic drawing of the NETtalk network architecture. A window of letters in an English text is fed to an array of 203 input units. Information from these units is transformed by an intermediate layer of 80 “hidden” units to produce patterns of activity in 20 output units. The connections in the network are specified by a total of 18020 weight parameters (including a variable threshold for each unit).