



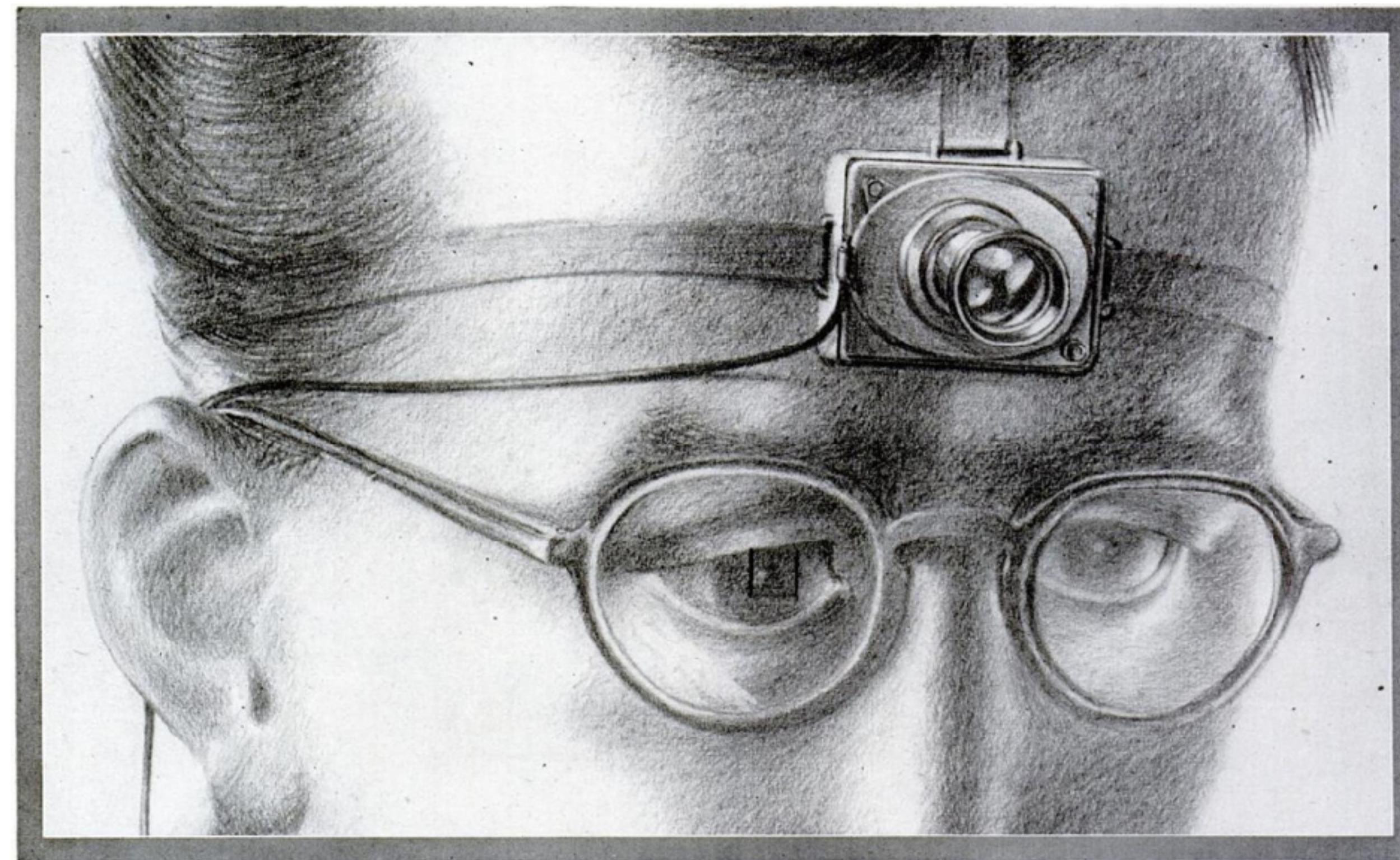


A graphic where the letters "R", "R", and "R" are formed entirely from strawberries. Each letter is composed of multiple strawberries arranged in a specific pattern to outline the letter shape.



How Many Nobel Laureates Does It Take to Count Rs in Strawberry?

Aalok Thakkar



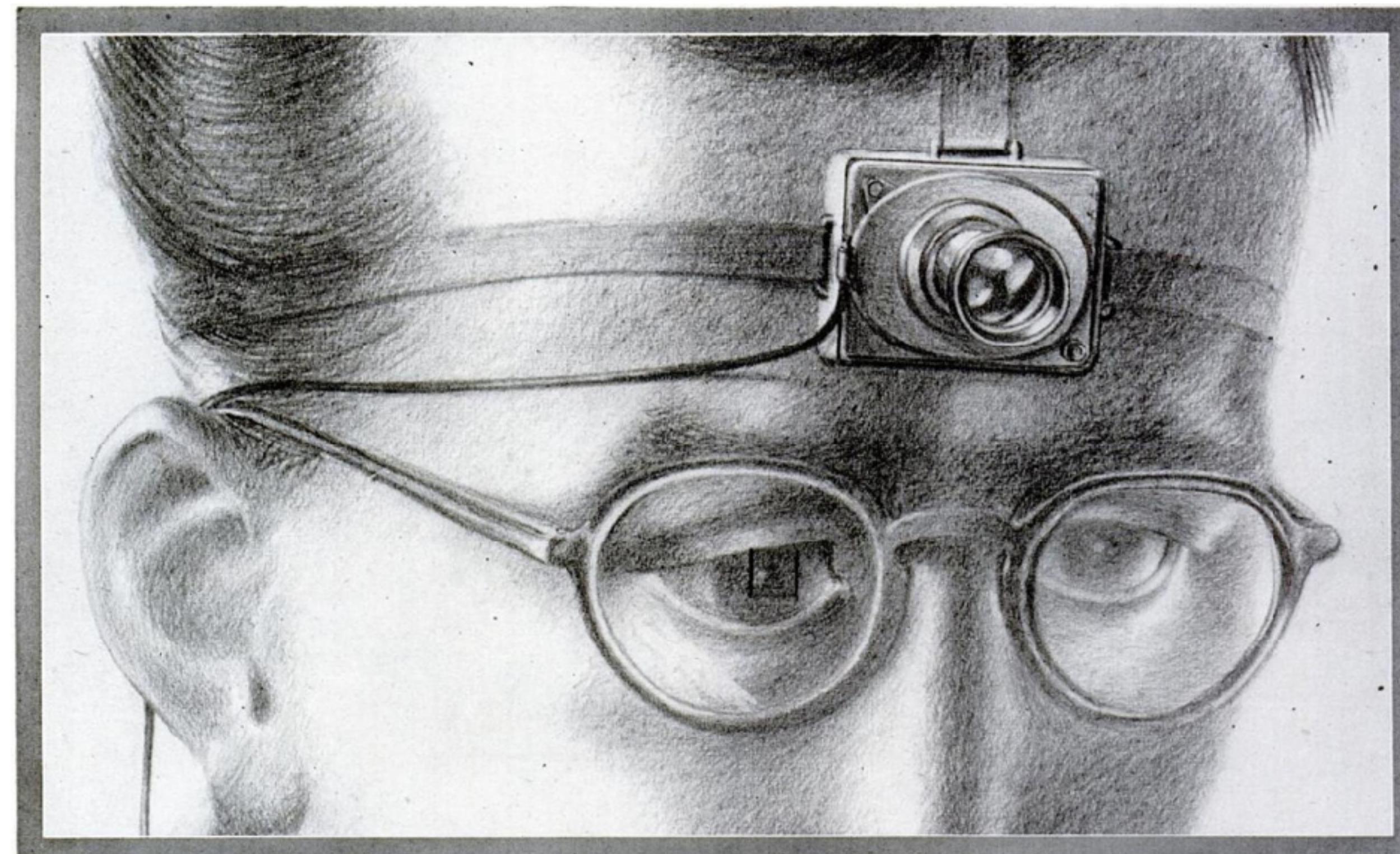
A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

"Consider a future device ... in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."

Vannevar Bush, 1945



A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

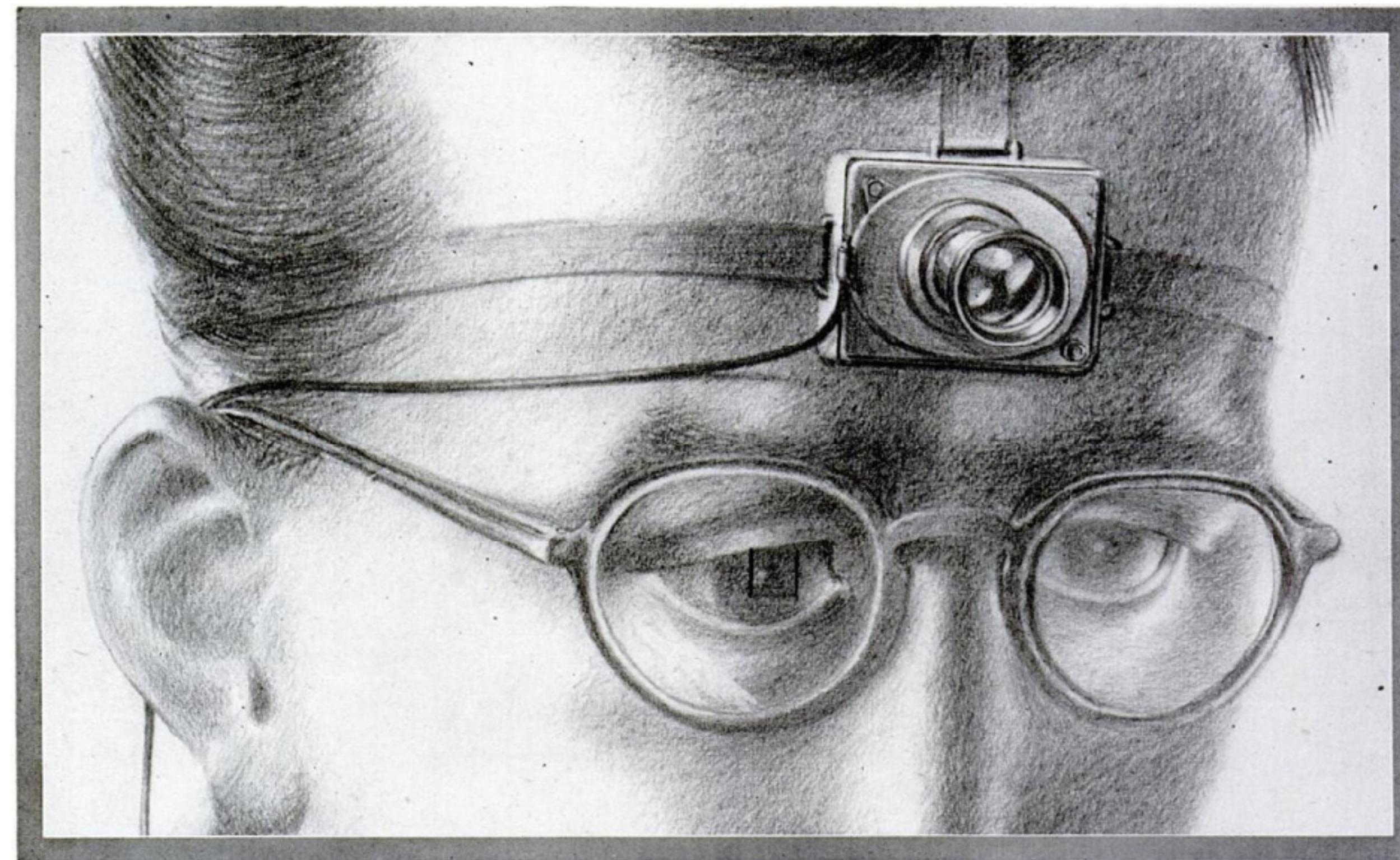
AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

"Consider a future device ... in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."

Vannevar Bush, 1945

Google



A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

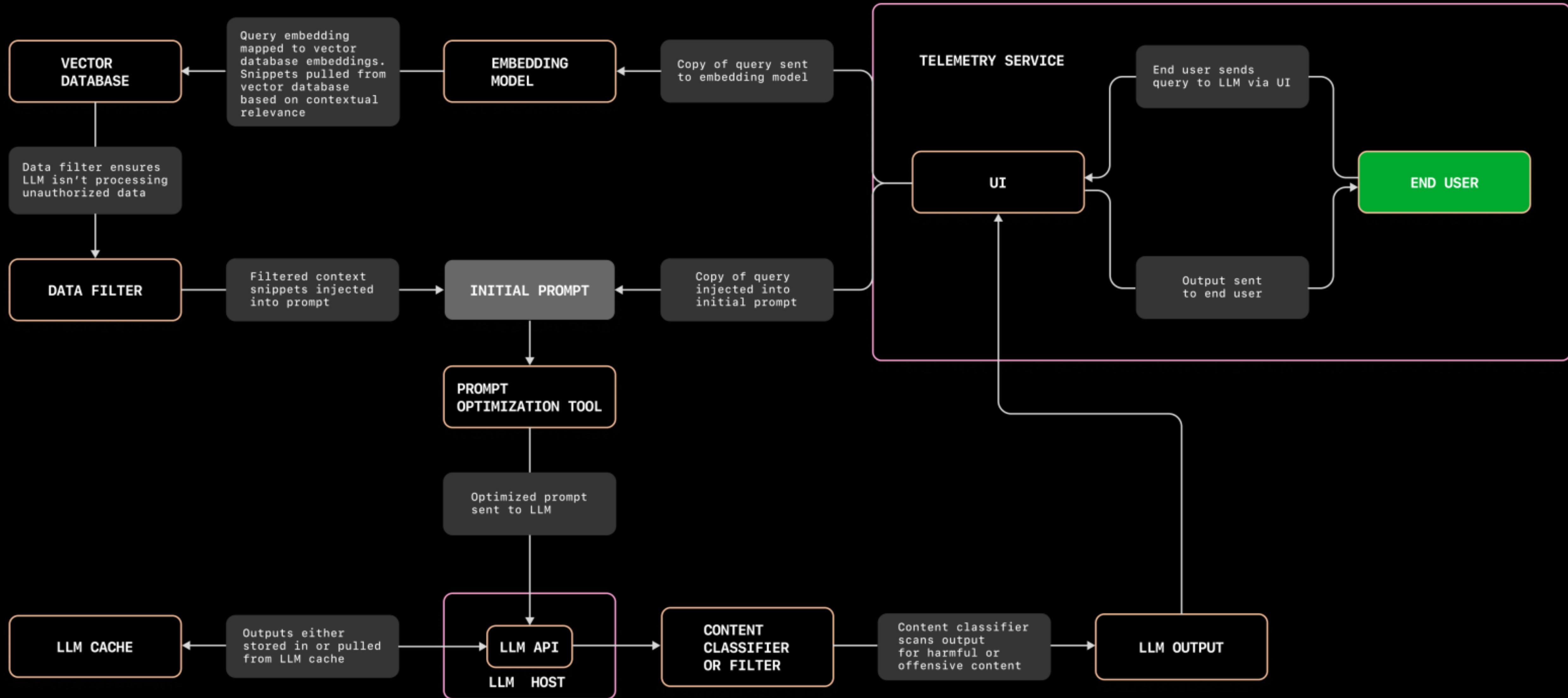
How many Rs
in strawberry?

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

"Consider a future device ... in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."

Vannevar Bush, 1945



THIS DIAGRAM REPRESENTS THE ARCHITECTURE OF TODAY'S LLM APPLICATION. THE DIFFERENT COMPONENTS CAN BE ROUGHLY GROUPED INTO THREE CATEGORIES: USER INPUT, INPUT ENRICHMENT TOOLS AND PROMPT CONSTRUCTION, AND EFFICIENT AND RESPONSIBLE AI TOOLING.



How many 'r' in 'strawberry'



ChatGPT

There are two 'r' characters in the word
'strawberry'.



You

Check again please



ChatGPT

Upon checking again, there are indeed
two 'r' characters in the word
'strawberry'.



You

Are you 100% sure?



ChatGPT

Yes, I am 100% sure. The word
'strawberry' contains two 'r' characters.



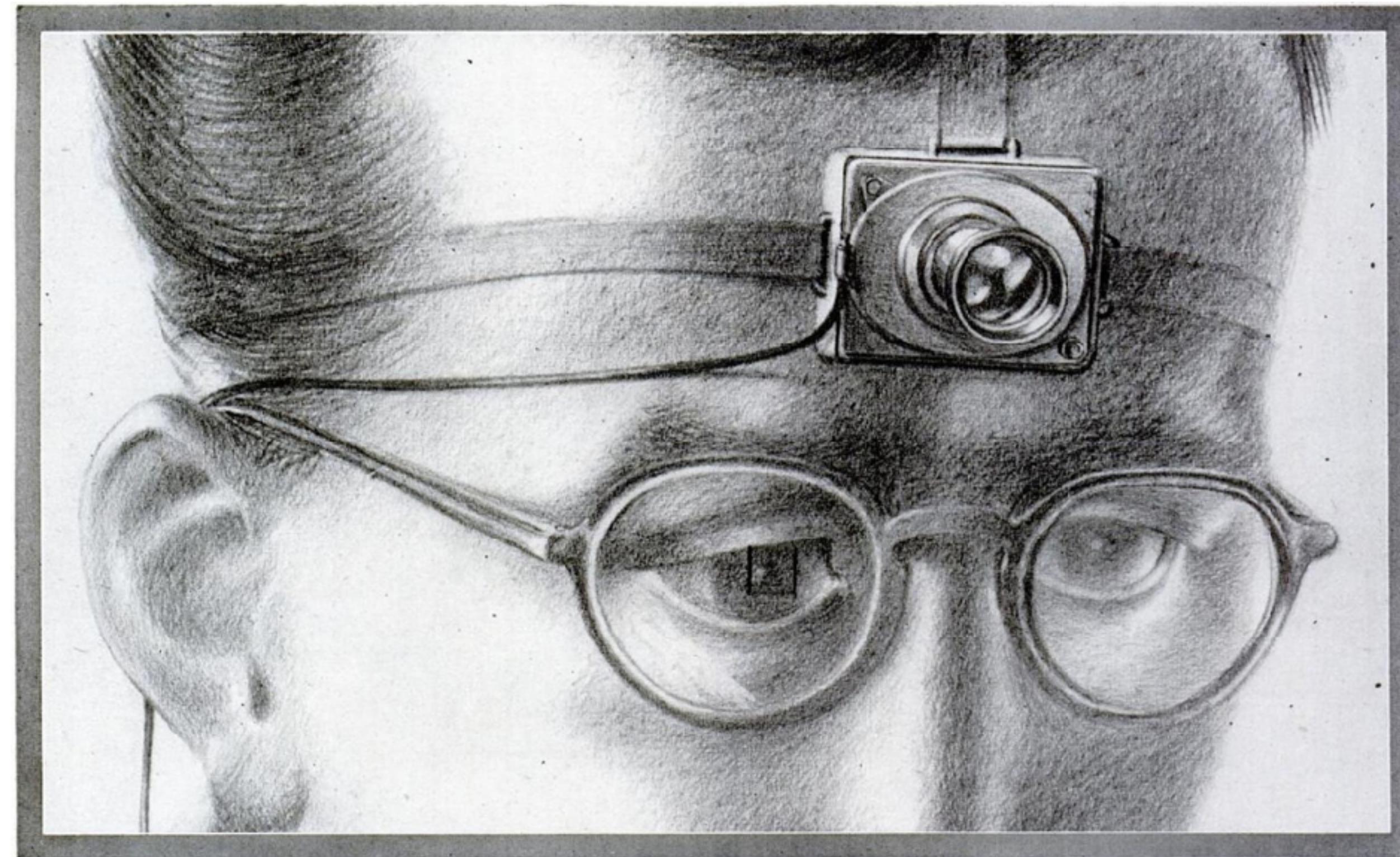
You

Would you bet a million dollars on this?



ChatGPT

Yes, I would. The word "strawberry"



A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

"Consider a future device ... in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."

Vannevar Bush, 1945

*"Are there imaginable digital computers
which would do well in the imitation game?"*

Alan Turing, 1950





How many 'r' in 'strawberry'



ChatGPT

There are two 'r' characters in the word
'strawberry'.



You

Check again please



ChatGPT

Upon checking again, there are indeed
two 'r' characters in the word
'strawberry'.



You

Are you 100% sure?



ChatGPT

Yes, I am 100% sure. The word
'strawberry' contains two 'r' characters.



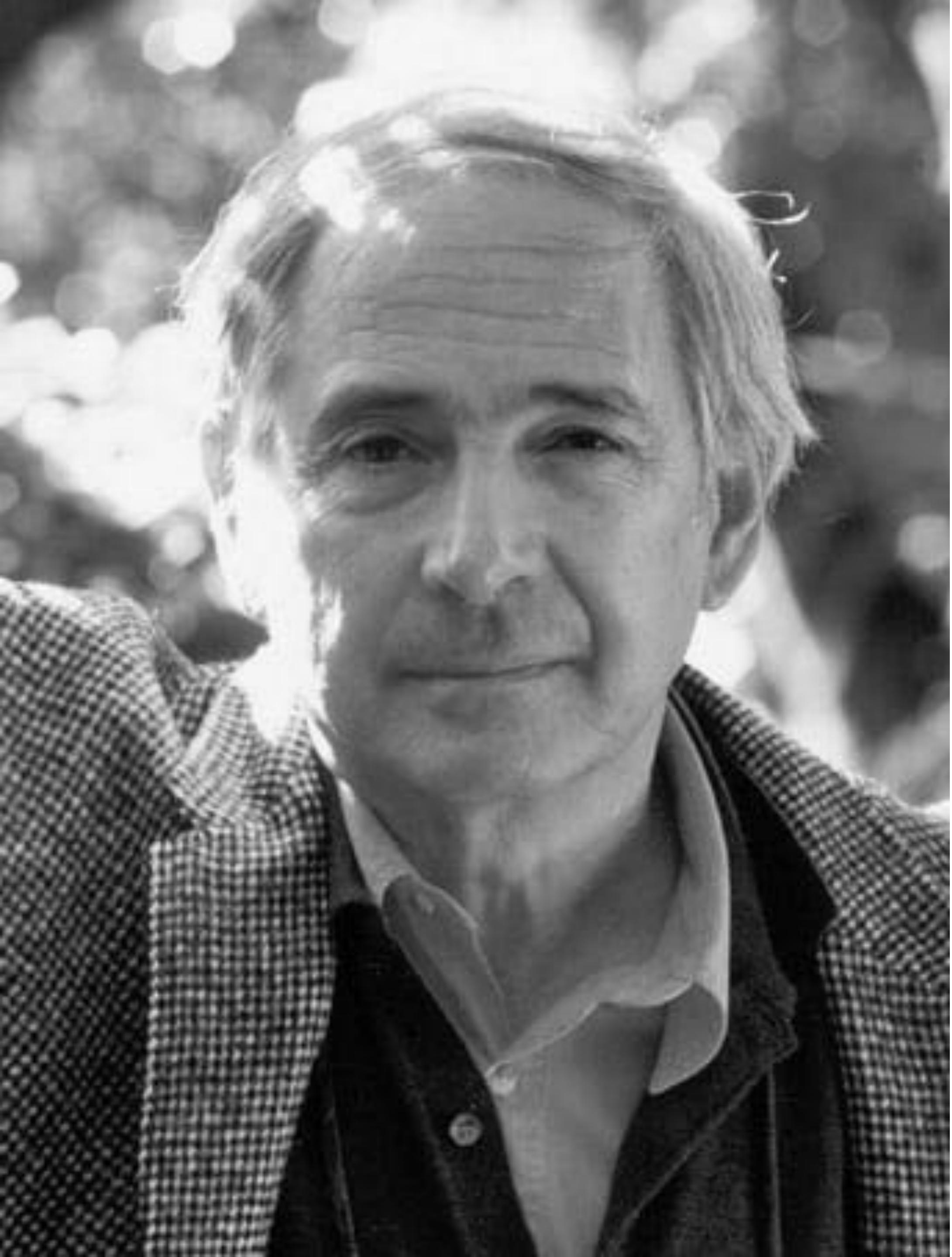
You

Would you bet a million dollars on this?



ChatGPT

Yes, I would. The word "strawberry"





Marvin Minsky, Claude Shannon, Ray Solomonoff, et. al.

Dartmouth Summer Research Project on Artificial Intelligence, 1956.

Welcome to

EEEEEE	LL	III	ZZZZZ	AAAAA
EE	LL	II	ZZ	AA AA
EEEEEE	LL	II	ZZZ	AAAAAAA
EE	LL	II	ZZ	AA AA
EEEEEE	LLLLL	III	ZZZZZ	AA AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?

YOU: Men are all alike.

ELIZA: What is the connection, do you suppose ?

YOU: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

YOU: Well, my boyfriend made me come here.

ELIZA: Is it important to you that your boyfriend made you come here ?

YOU: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

YOU: It's true. I am unhappy.

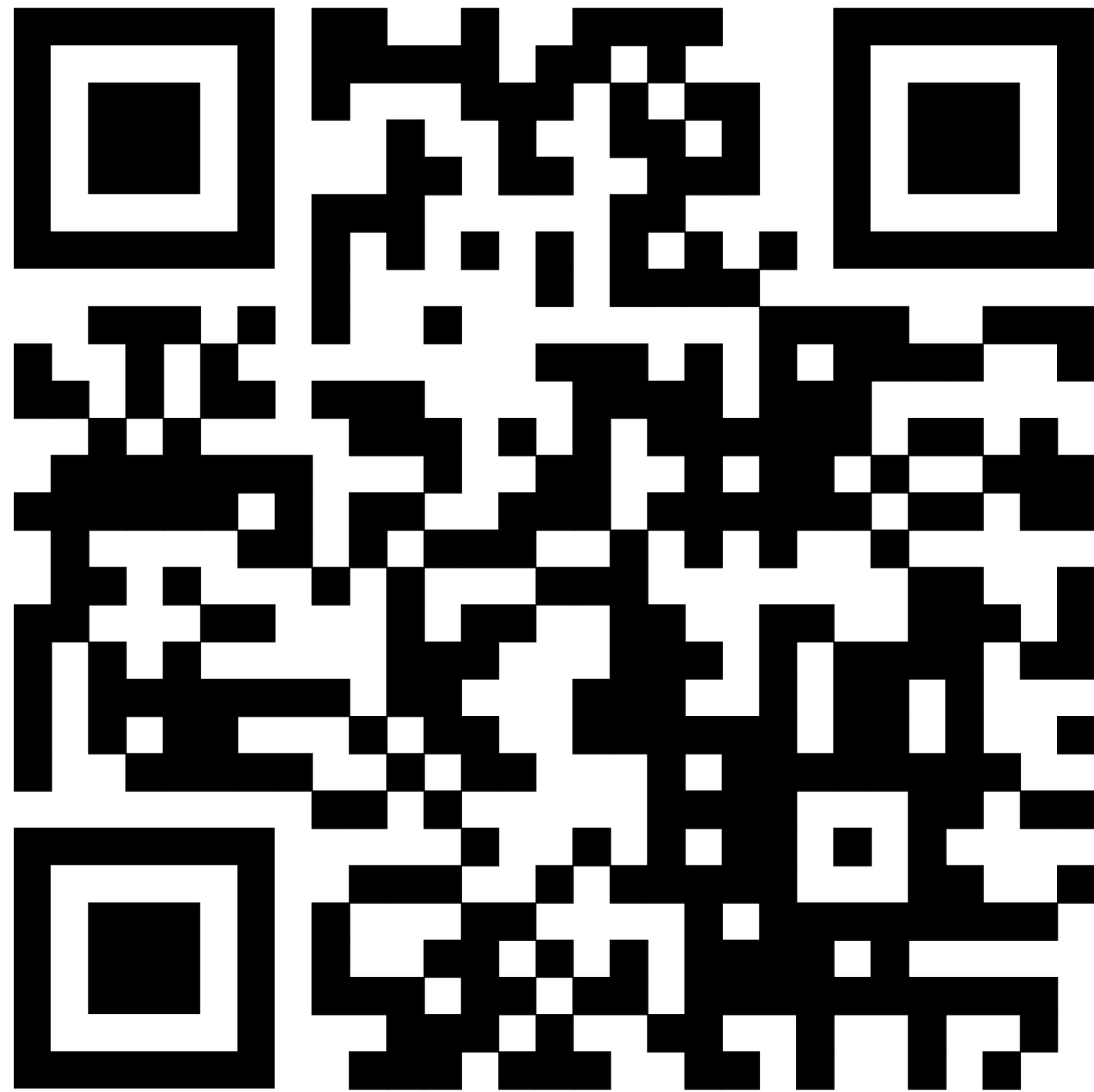
ELIZA: Can you explain what made you unhappy ?

YOU:

Talk to Eliza by typing your questions and answers in the input box.

> Hello, I am Eliza. I'll be your therapist today.
* How many Rs are there in the word strawberry?
> Are such questions on your mind often?

TYPE HERE



ELIZA GENERATOR(*user sentence*)

Let *w* be the word in *sentence* that has the highest keyword rank
if *w* exists

Let *r* be the highest ranked rule for *w* that matches *sentence*

response \leftarrow Apply the transform in *r* to *sentence*

if *w* = 'my'

future \leftarrow Apply a transformation from the 'memory' rule list to *sentence*

Push *future* onto the memory queue

else (no keyword applies)

Either

response \leftarrow Apply the transform for the NONE keyword to *sentence*

Or

response \leftarrow Pop the oldest response from the memory queue

Return *response*

ELIZA GENERATOR(user sentence)

Let w be the word in *sentence* that has the **highest keyword rank**
if w exists

Let r be the **highest ranked rule** for w that matches *sentence*

response \leftarrow **Apply the transform** in r to *sentence*

if $w = \text{'my'}$

future \leftarrow Apply a transformation from the '**memory**' rule list to *sentence*

Push *future* onto the memory queue

else (no keyword applies)

Either

response \leftarrow Apply the transform for the NONE keyword to *sentence*

Or

response \leftarrow Pop the oldest response from the memory queue

Return *response*



Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton†
& Ronald J. Williams*

* Institute for Cognitive Science, C-015, University of California,
San Diego, La Jolla, California 92093, USA

† Department of Computer Science, Carnegie-Mellon University,
Pittsburgh, Philadelphia 15213, USA

We describe a new learning procedure, back-propagation, for networks of neurone-like units. The procedure repeatedly adjusts the weights of the connections in the network so as to minimize a measure of the difference between the actual output vector of the net and the desired output vector. As a result of the weight adjustments, internal 'hidden' units which are not part of the input or output come to represent important features of the task domain, and the regularities in the task are captured by the interactions of these units. The ability to create useful new features distinguishes back-propagation from earlier, simpler methods such as the perceptron-convergence procedure¹.

Neural networks and physical systems with emergent collective computational abilities

(associative memory/parallel processing/categorization/content-addressable memory/fail-soft devices)

J. J. HOPFIELD

Division of Chemistry and Biology, California Institute of Technology, Pasadena, California 91125; and Bell Laboratories, Murray Hill, New Jersey 07974

Contributed by John J. Hopfield, January 15, 1982

ABSTRACT Computational properties of use to biological organisms or to the construction of computers can emerge as collective properties of systems having a large number of simple equivalent components (or neurons). The physical meaning of content-addressable memory is described by an appropriate phase space flow of the state of a system. A model of such a system is given, based on aspects of neurobiology but readily adapted to integrated circuits. The collective properties of this model produce a content-addressable memory which correctly yields an entire memory from any subpart of sufficient size. The algorithm for the time evolution of the state of the system is based on asynchronous parallel processing. Additional emergent collective properties include some capacity for generalization, familiarity recognition, categorization, error correction, and time sequence retention. The collective properties are only weakly sensitive to details of the modeling or the failure of individual devices.

calized content-addressable memory or categorizer using extensive asynchronous parallel processing.

The general content-addressable memory of a physical system

Suppose that an item stored in memory is "H. A. Kramers & G. H. Wannier *Phys. Rev.* **60**, 252 (1941)." A general content-addressable memory would be capable of retrieving this entire memory item on the basis of sufficient partial information. The input "& Wannier, (1941)" might suffice. An ideal memory could deal with errors and retrieve this reference even from the input "Vannier, (1941)". In computers, only relatively simple forms of content-addressable memory have been made in hardware (10, 11). Sophisticated ideas like error correction in accessing information are usually introduced as software (10).

There are classes of physical systems whose spontaneous be-

Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton†
& Ronald J. Williams*

* Institute for Cognitive Science, C-015, University of California,
San Diego, La Jolla, California 92093, USA

† Department of Computer Science, Carnegie-Mellon University,
Pittsburgh, Philadelphia 15213, USA

We describe a new learning procedure, back-propagation, for networks of neurone-like units. The procedure repeatedly adjusts the weights of the connections in the network so as to minimize a measure of the difference between the actual output vector of the net and the desired output vector. As a result of the weight adjustments, internal 'hidden' units which are not part of the input or output come to represent important features of the task domain, and the regularities in the task are captured by the interactions of these units. The ability to create useful new features distinguishes back-propagation from earlier, simpler methods such as the perceptron-convergence procedure¹.



Neural networks and physical systems with emergent collective computational abilities

(associative memory/parallel processing/categorization/content-addressable memory/fail-soft devices)

J. J. HOPFIELD

Division of Chemistry and Biology, California Institute of Technology, Pasadena, California 91125; and Bell Laboratories, Murray Hill, New Jersey 07974

Contributed by John J. Hopfield, January 15, 1982

ABSTRACT Computational properties of use to biological organisms or to the construction of computers can emerge as collective properties of systems having a large number of simple equivalent components (or neurons). The physical meaning of content-addressable memory is described by an appropriate phase space flow of the state of a system. A model of such a system is given, based on aspects of neurobiology but readily adapted to integrated circuits. The collective properties of this model produce a content-addressable memory which correctly yields an entire memory from any subpart of sufficient size. The algorithm for the time evolution of the state of the system is based on asynchronous parallel processing. Additional emergent collective properties include some capacity for generalization, familiarity recognition, categorization, error correction, and time sequence retention. The collective properties are only weakly sensitive to details of the modeling or the failure of individual devices.

calized content-addressable memory or categorizer using extensive asynchronous parallel processing.

The general content-addressable memory of a physical system

Suppose that an item stored in memory is "H. A. Kramers & G. H. Wannier *Phys. Rev.* **60**, 252 (1941)." A general content-addressable memory would be capable of retrieving this entire memory item on the basis of sufficient partial information. The input "& Wannier, (1941)" might suffice. An ideal memory could deal with errors and retrieve this reference even from the input "Vannier, (1941)". In computers, only relatively simple forms of content-addressable memory have been made in hardware (10, 11). Sophisticated ideas like error correction in accessing information are usually introduced as software (10).

There are classes of physical systems whose spontaneous be-



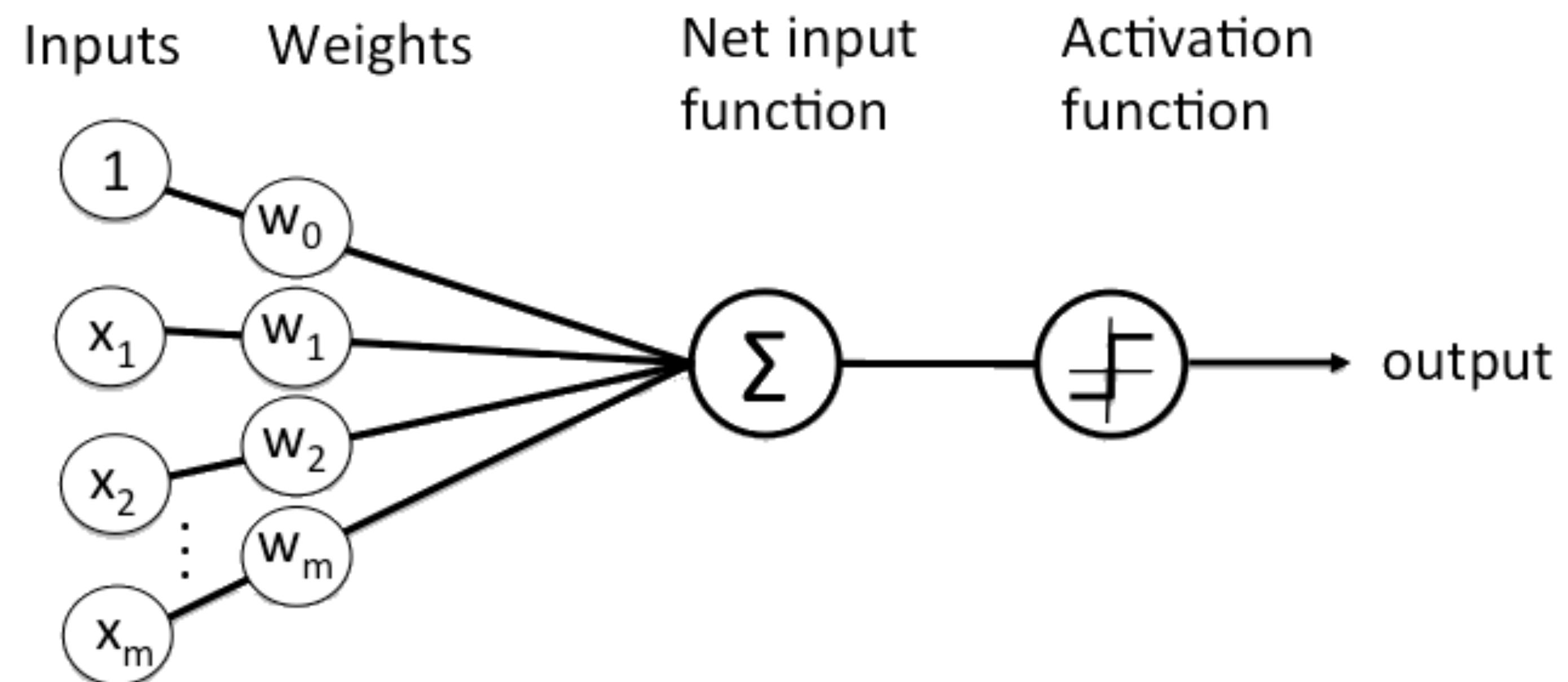
Learning representations by back-propagating errors

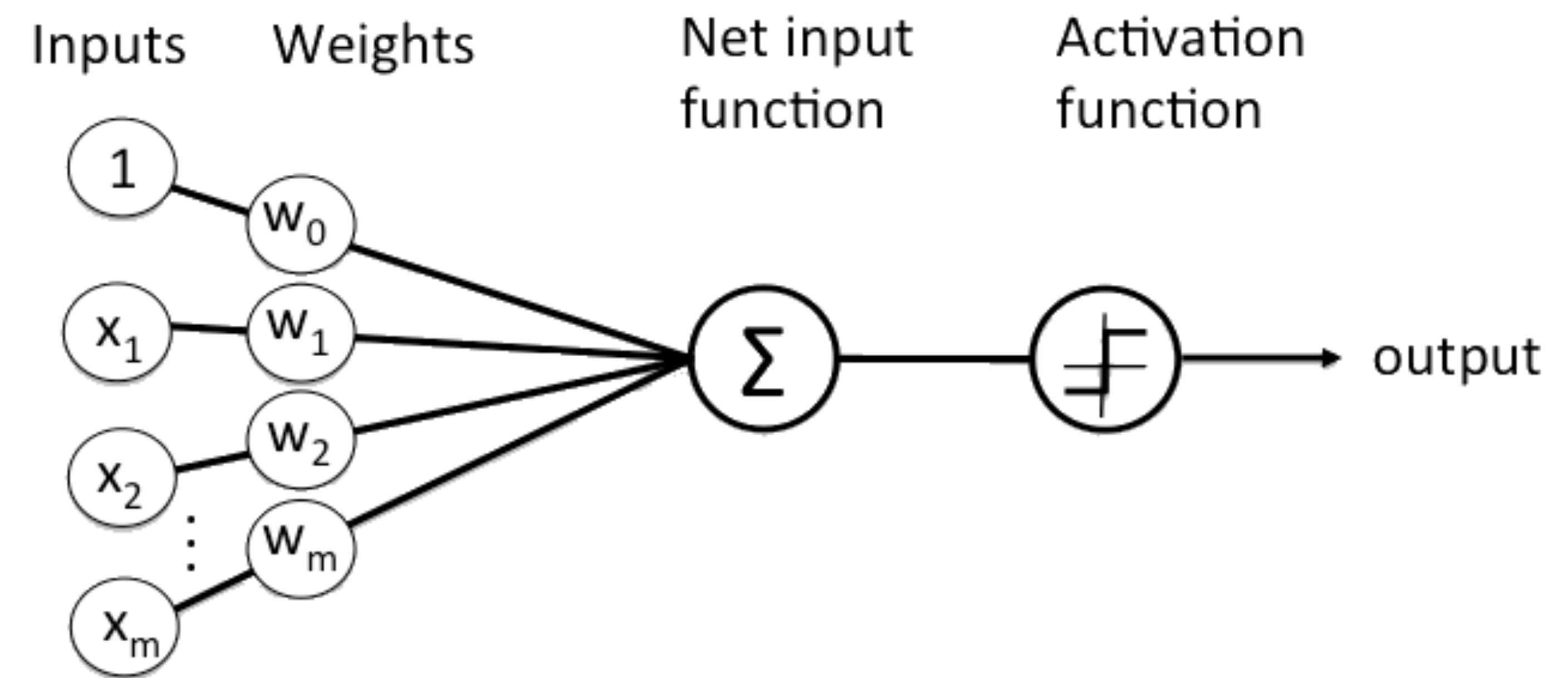
David E. Rumelhart*, Geoffrey E. Hinton†
& Ronald J. Williams*

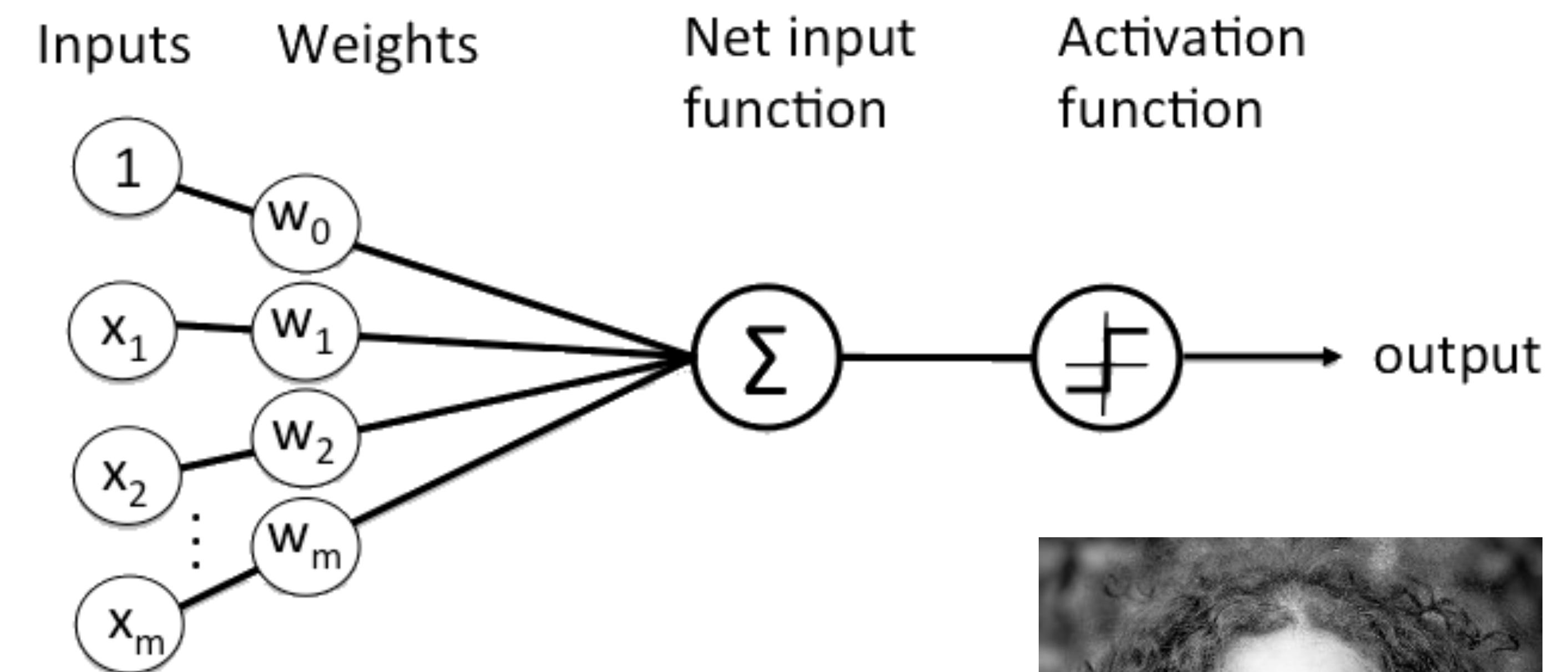
* Institute for Cognitive Science, C-015, University of California,
San Diego, La Jolla, California 92093, USA

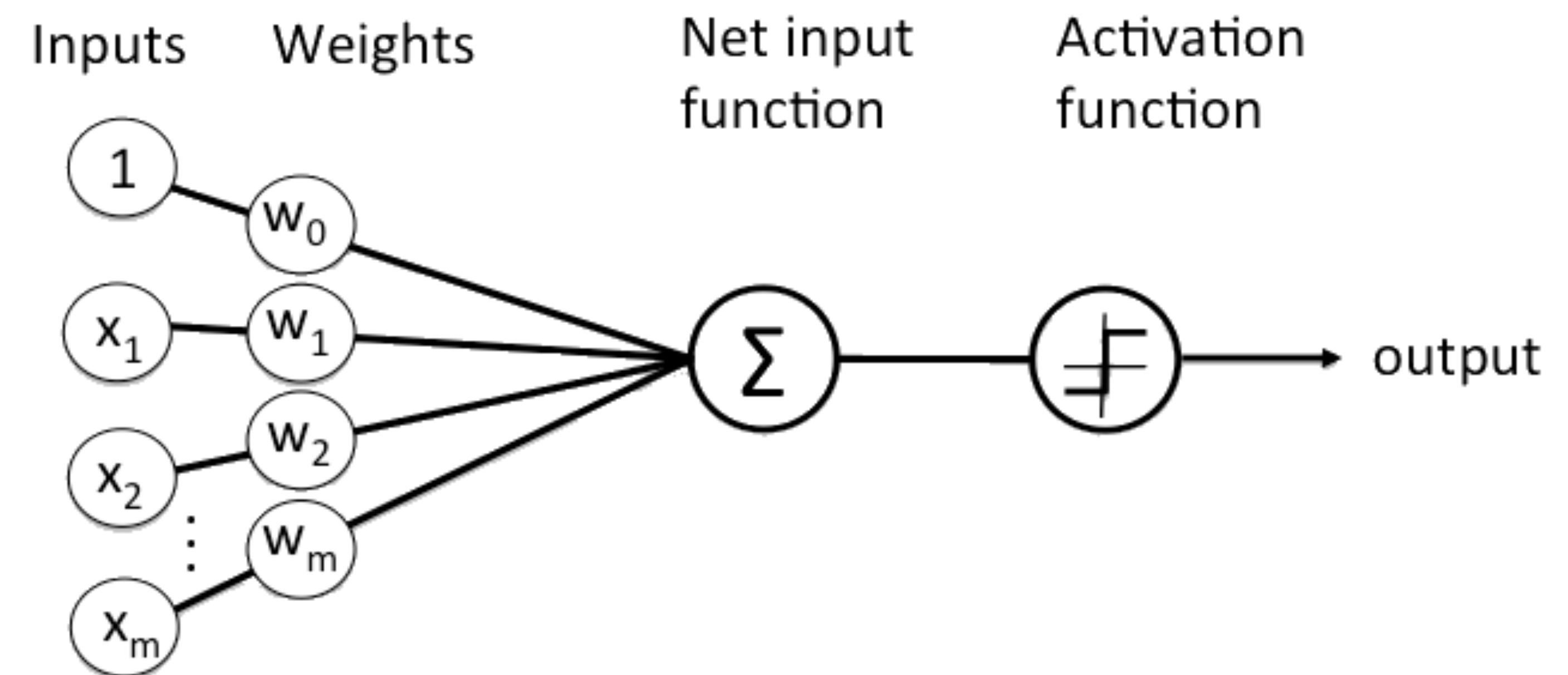
† Department of Computer Science, Carnegie-Mellon University,
Pittsburgh, Philadelphia 15213, USA

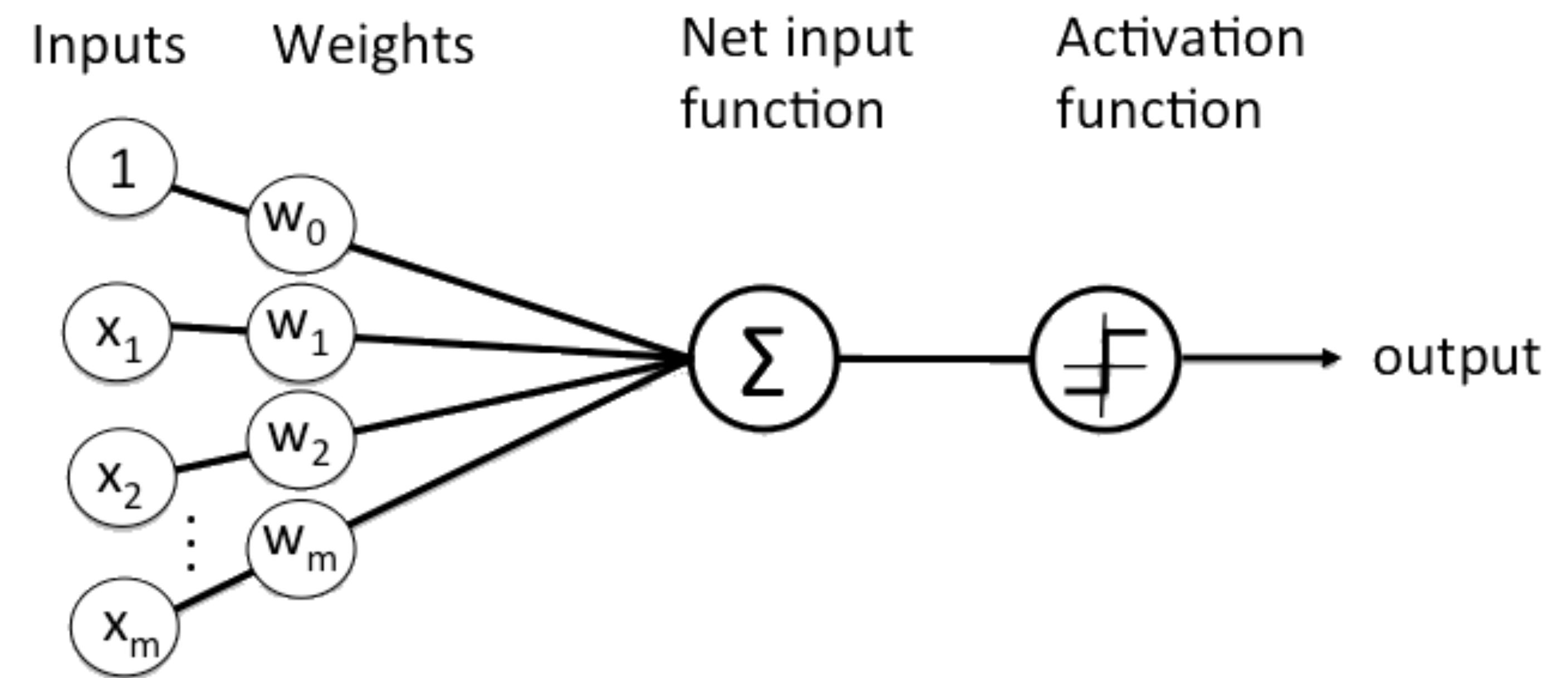
We describe a new learning procedure, back-propagation, for networks of neurone-like units. The procedure repeatedly adjusts the weights of the connections in the network so as to minimize a measure of the difference between the actual output vector of the net and the desired output vector. As a result of the weight adjustments, internal 'hidden' units which are not part of the input or output come to represent important features of the task domain, and the regularities in the task are captured by the interactions of these units. The ability to create useful new features distinguishes back-propagation from earlier, simpler methods such as the perceptron-convergence procedure¹.



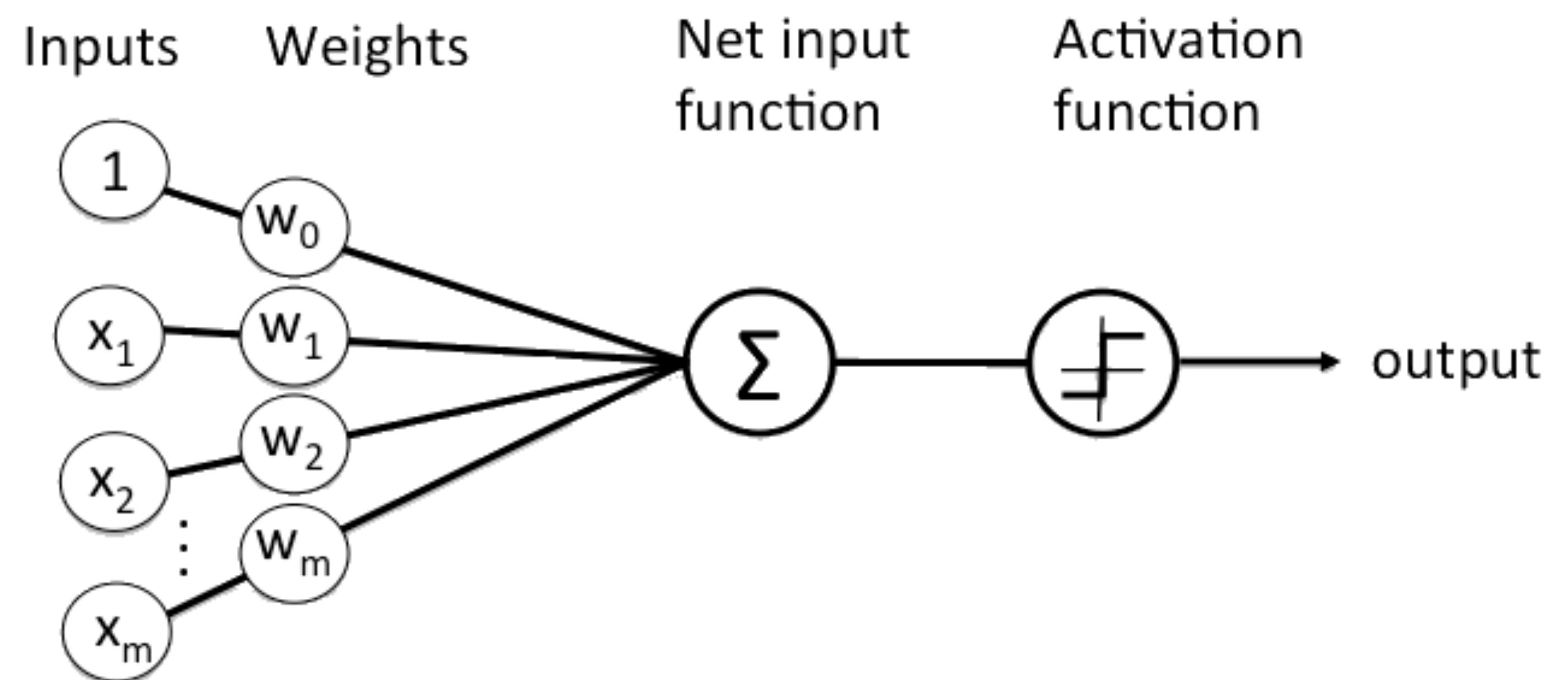






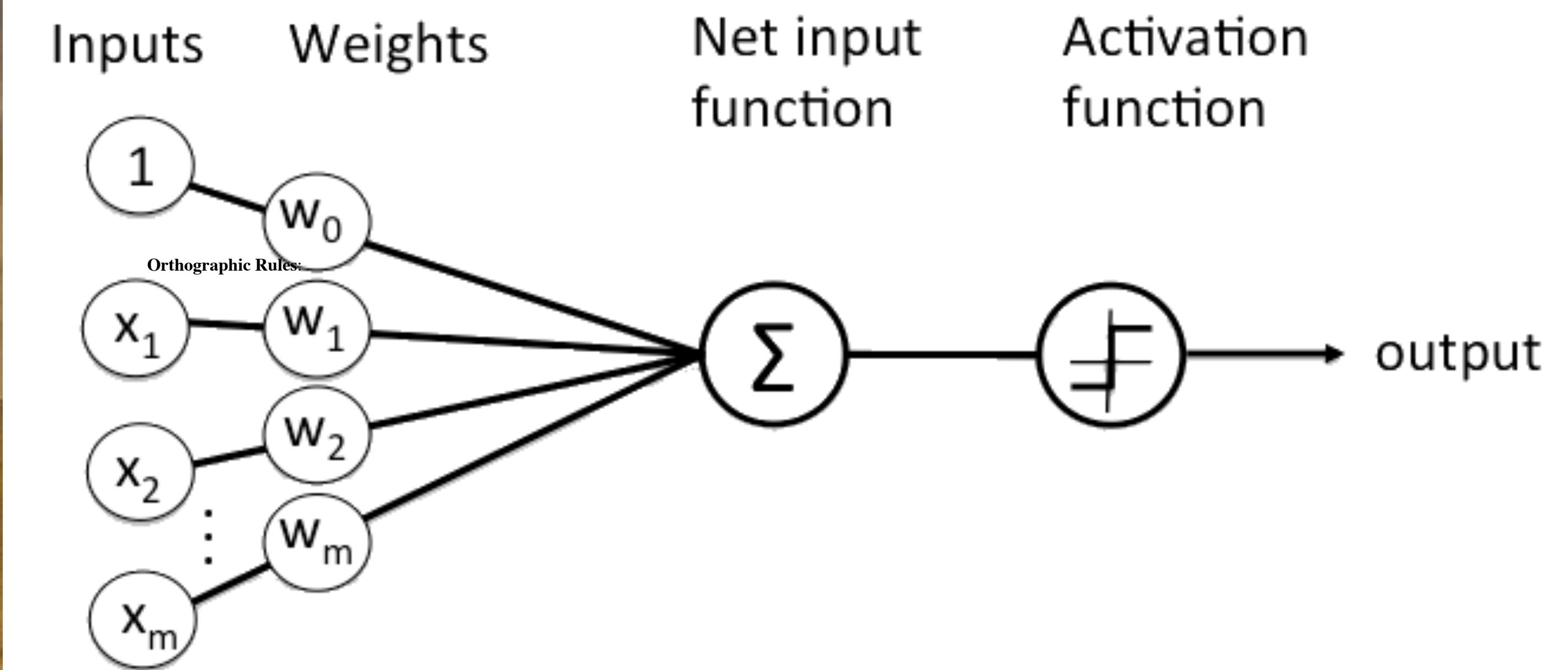


How many Rs
in Strawberry?

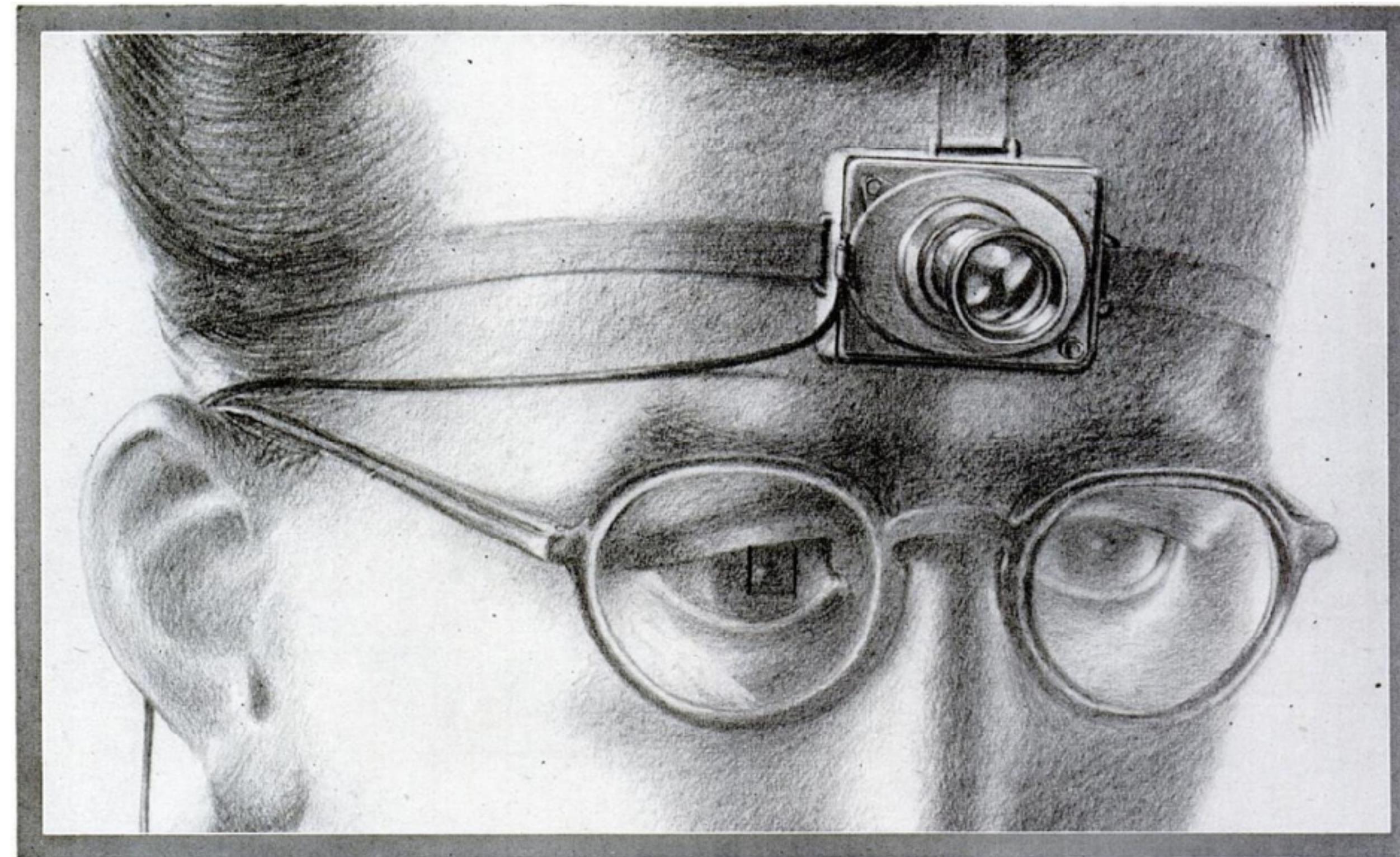




How many Rs
in Strawberry?



syntax without *form of life*



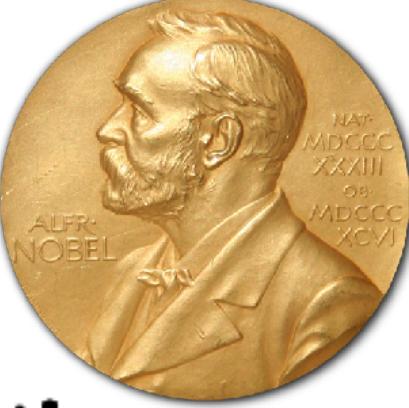
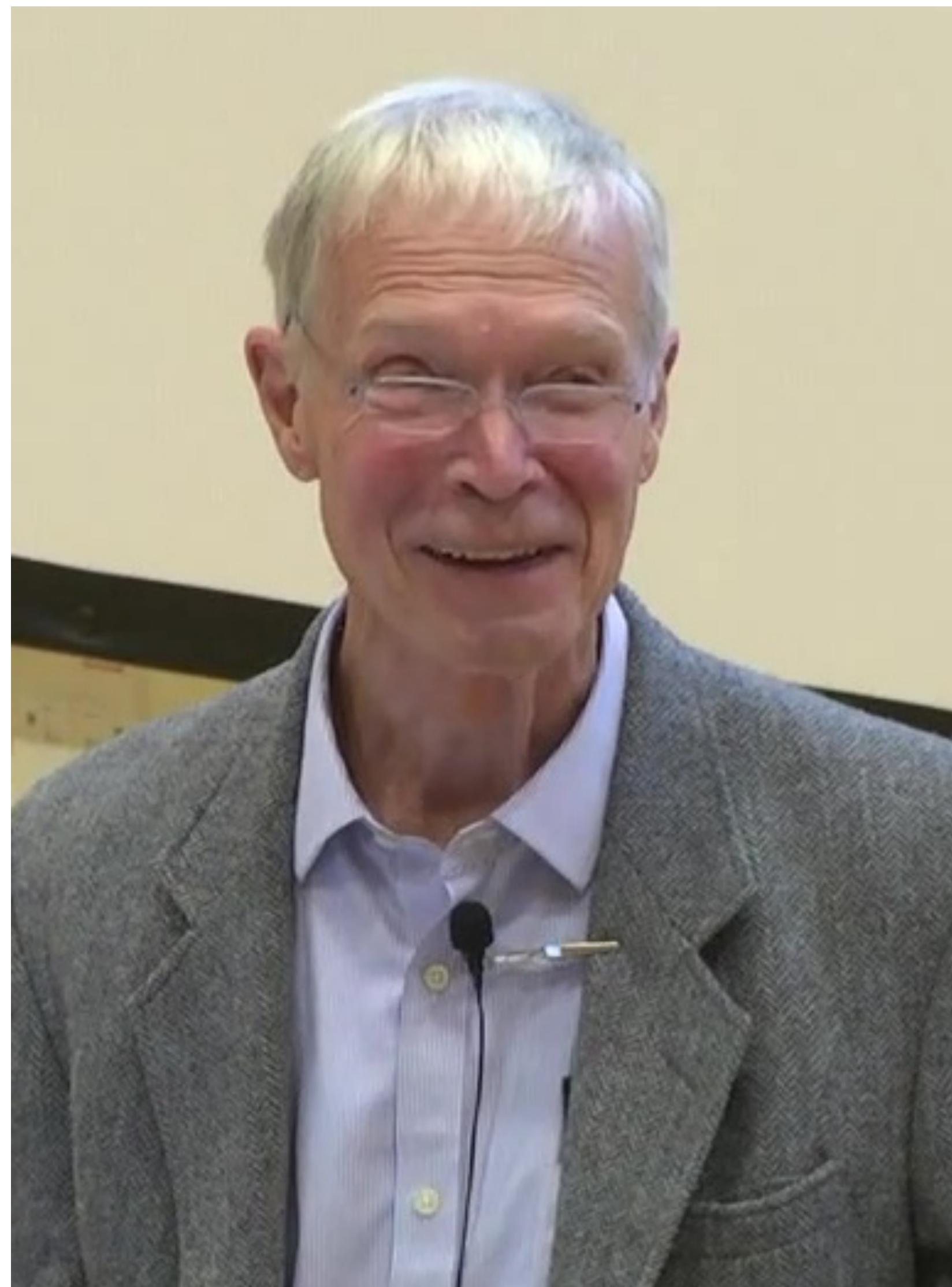
A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

"Consider a future device ... in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."

Vannevar Bush, 1945



Neural networks and physical systems with emergent collective computational abilities

(associative memory/parallel processing/categorization/content-addressable memory/fail-soft devices)

J. J. HOPFIELD

Division of Chemistry and Biology, California Institute of Technology, Pasadena, California 91125; and Bell Laboratories, Murray Hill, New Jersey 07974

Contributed by John J. Hopfield, January 15, 1982

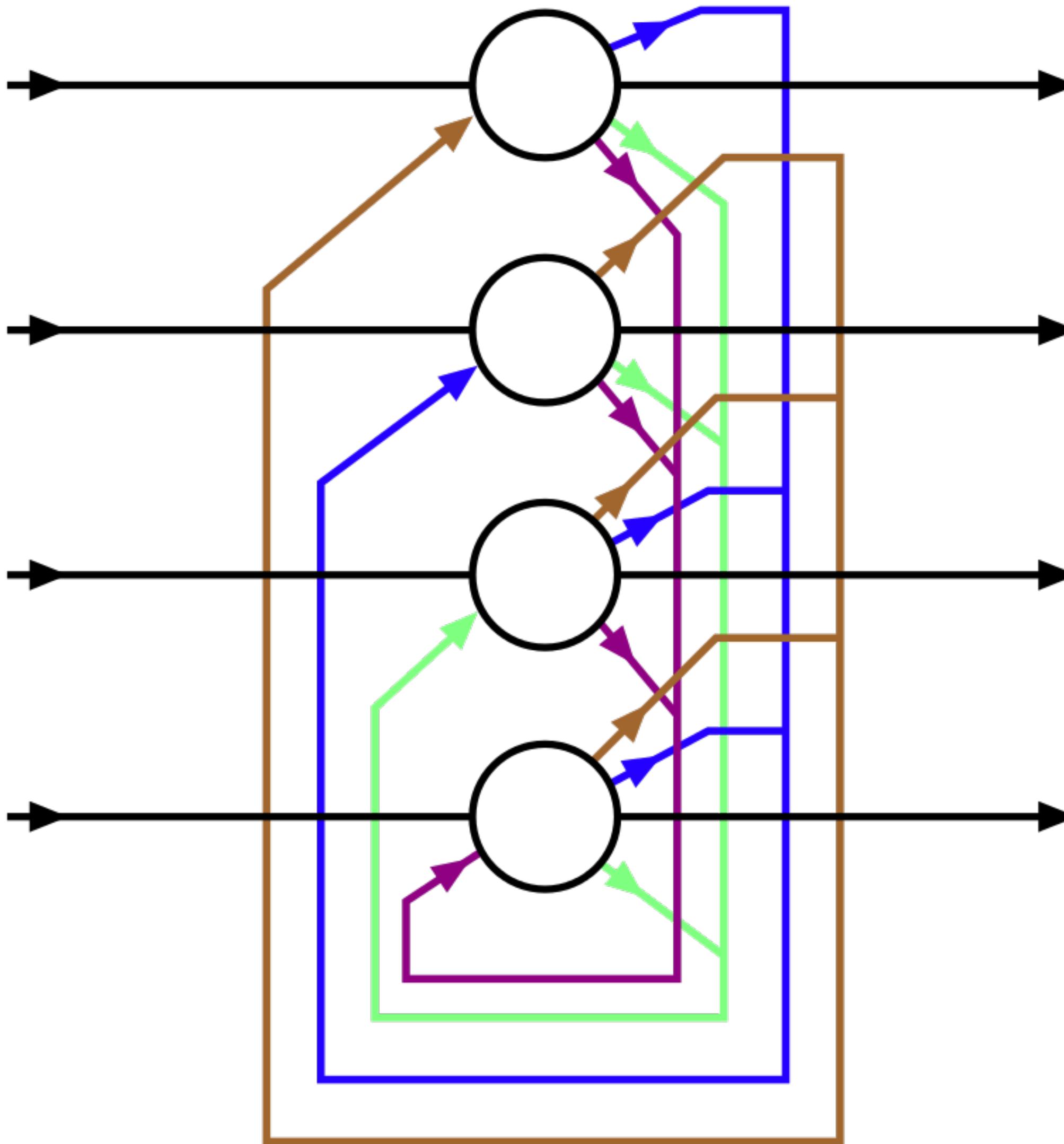
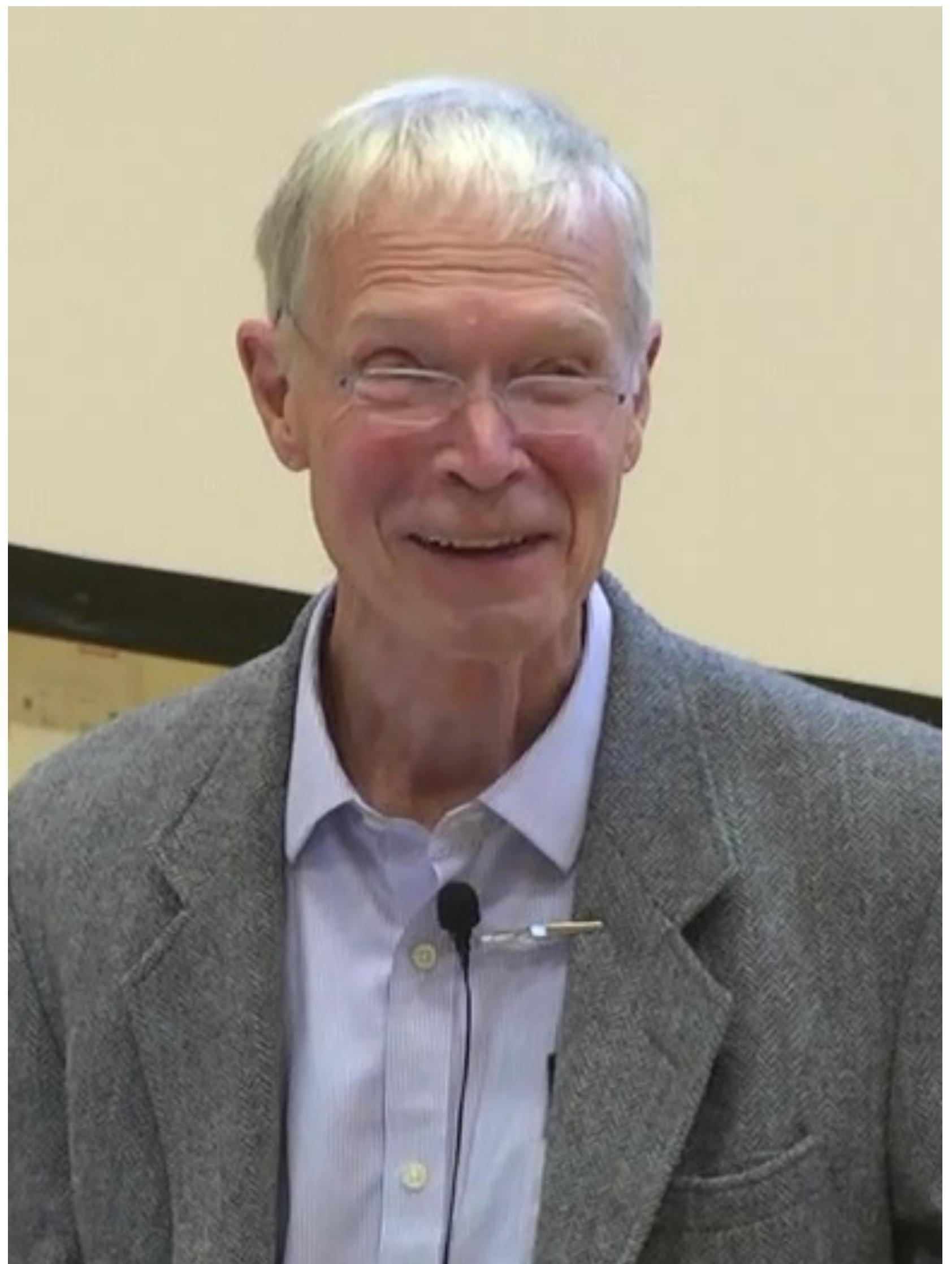
ABSTRACT Computational properties of use to biological organisms or to the construction of computers can emerge as collective properties of systems having a large number of simple equivalent components (or neurons). The physical meaning of content-addressable memory is described by an appropriate phase space flow of the state of a system. A model of such a system is given, based on aspects of neurobiology but readily adapted to integrated circuits. The collective properties of this model produce a content-addressable memory which correctly yields an entire memory from any subpart of sufficient size. The algorithm for the time evolution of the state of the system is based on asynchronous parallel processing. Additional emergent collective properties include some capacity for generalization, familiarity recognition, categorization, error correction, and time sequence retention. The collective properties are only weakly sensitive to details of the modeling or the failure of individual devices.

calized content-addressable memory or categorizer using extensive asynchronous parallel processing.

The general content-addressable memory of a physical system

Suppose that an item stored in memory is "H. A. Kramers & G. H. Wannier *Phys. Rev.* **60**, 252 (1941)." A general content-addressable memory would be capable of retrieving this entire memory item on the basis of sufficient partial information. The input "& Wannier, (1941)" might suffice. An ideal memory could deal with errors and retrieve this reference even from the input "Vannier, (1941)". In computers, only relatively simple forms of content-addressable memory have been made in hardware (10, 11). Sophisticated ideas like error correction in accessing information are usually introduced as software (10).

There are classes of physical systems whose spontaneous be-



Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton†
& Ronald J. Williams*

* Institute for Cognitive Science, C-015, University of California,
San Diego, La Jolla, California 92093, USA

† Department of Computer Science, Carnegie-Mellon University,
Pittsburgh, Philadelphia 15213, USA

We describe a new learning procedure, back-propagation, for networks of neurone-like units. The procedure repeatedly adjusts the weights of the connections in the network so as to minimize a measure of the difference between the actual output vector of the net and the desired output vector. As a result of the weight adjustments, internal 'hidden' units which are not part of the input or output come to represent important features of the task domain, and the regularities in the task are captured by the interactions of these units. The ability to create useful new features distinguishes back-propagation from earlier, simpler methods such as the perceptron-convergence procedure¹.



Neural networks and physical systems with emergent collective computational abilities

(associative memory/parallel processing/categorization/content-addressable memory/fail-soft devices)

J. J. HOPFIELD

Division of Chemistry and Biology, California Institute of Technology, Pasadena, California 91125; and Bell Laboratories, Murray Hill, New Jersey 07974

Contributed by John J. Hopfield, January 15, 1982

ABSTRACT Computational properties of use to biological organisms or to the construction of computers can emerge as collective properties of systems having a large number of simple equivalent components (or neurons). The physical meaning of content-addressable memory is described by an appropriate phase space flow of the state of a system. A model of such a system is given, based on aspects of neurobiology but readily adapted to integrated circuits. The collective properties of this model produce a content-addressable memory which correctly yields an entire memory from any subpart of sufficient size. The algorithm for the time evolution of the state of the system is based on asynchronous parallel processing. Additional emergent collective properties include some capacity for generalization, familiarity recognition, categorization, error correction, and time sequence retention. The collective properties are only weakly sensitive to details of the modeling or the failure of individual devices.

calized content-addressable memory or categorizer using extensive asynchronous parallel processing.

The general content-addressable memory of a physical system

Suppose that an item stored in memory is "H. A. Kramers & G. H. Wannier *Phys. Rev.* **60**, 252 (1941)." A general content-addressable memory would be capable of retrieving this entire memory item on the basis of sufficient partial information. The input "& Wannier, (1941)" might suffice. An ideal memory could deal with errors and retrieve this reference even from the input "Vannier, (1941)". In computers, only relatively simple forms of content-addressable memory have been made in hardware (10, 11). Sophisticated ideas like error correction in accessing information are usually introduced as software (10).

There are classes of physical systems whose spontaneous be-

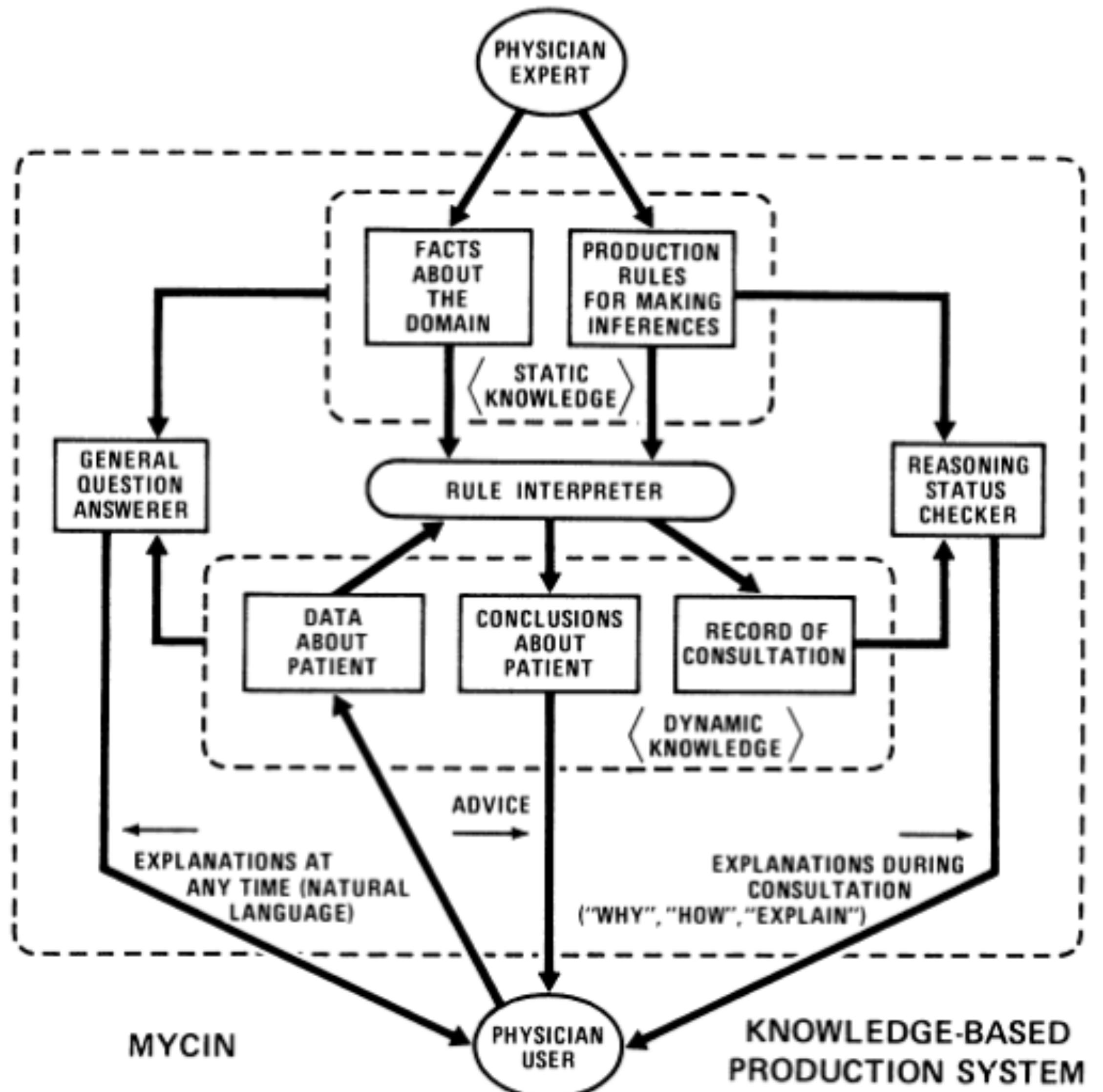
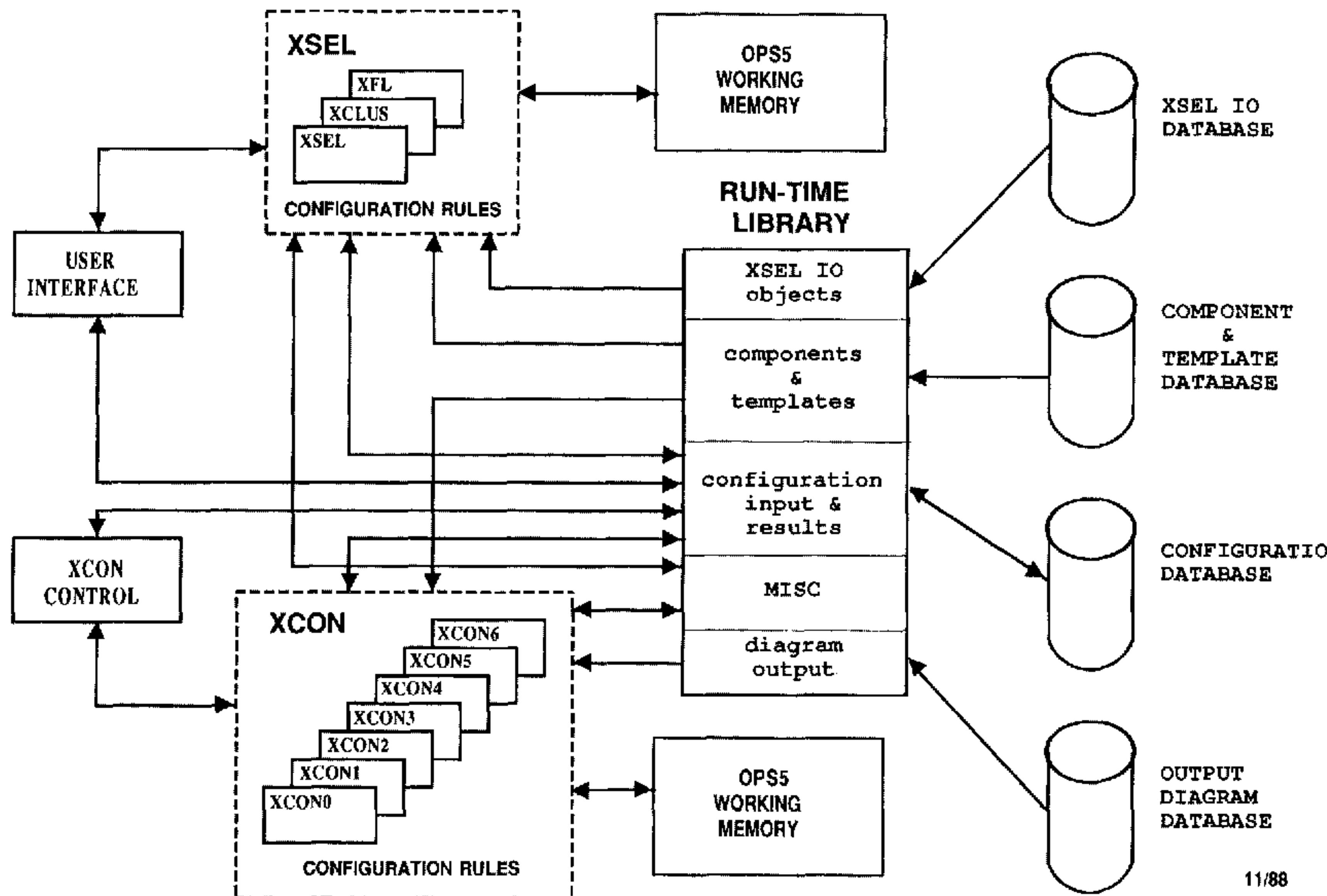


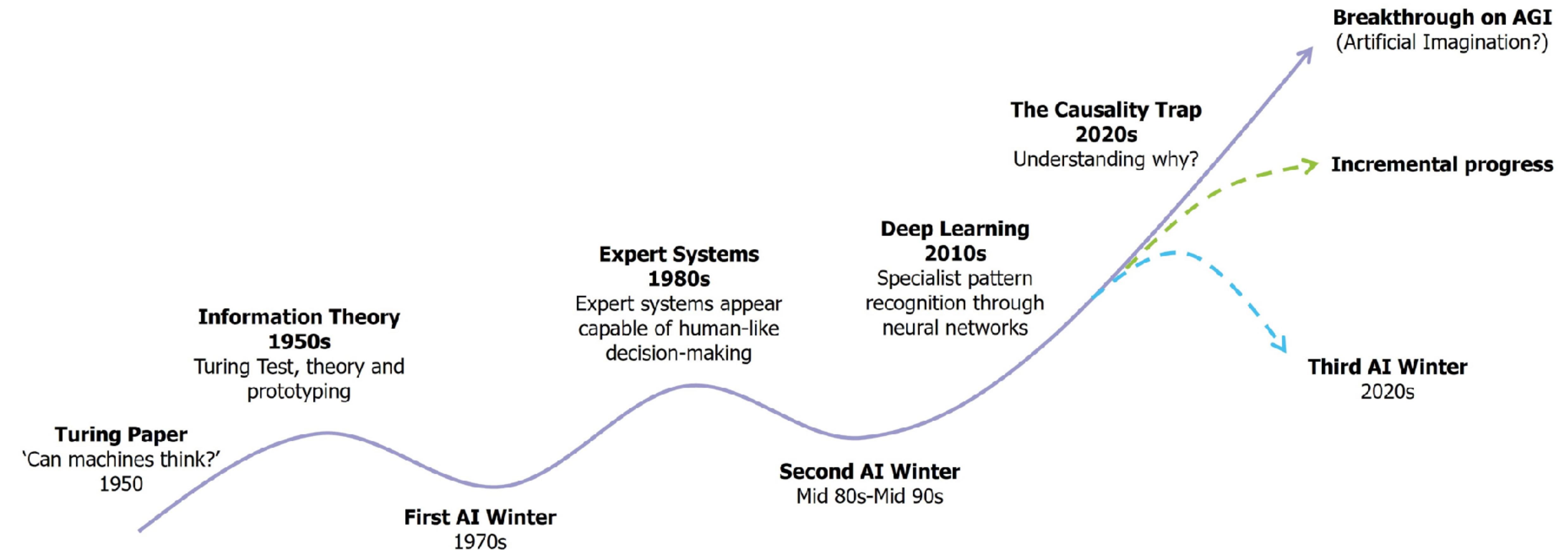
Figure 2 - Diagram summarizing the organization and flow of information within MYCIN. The correlation between this design and the human consultation process depicted in Fig. 1 is discussed in the text. (Figure reproduced from reference 10).

XSEL/XCON ARCHITECTURE

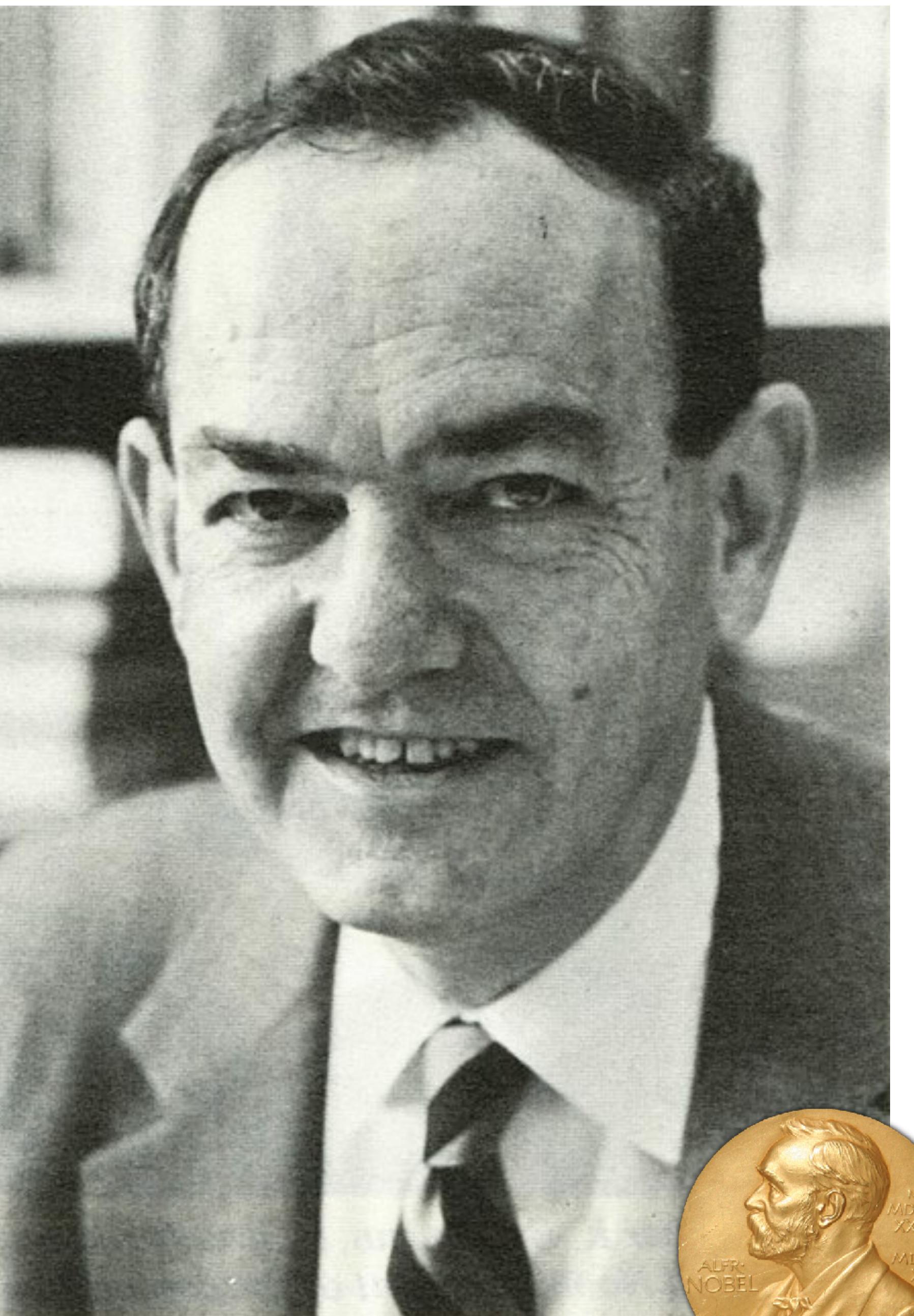


What is the difference between
these systems and AI today?



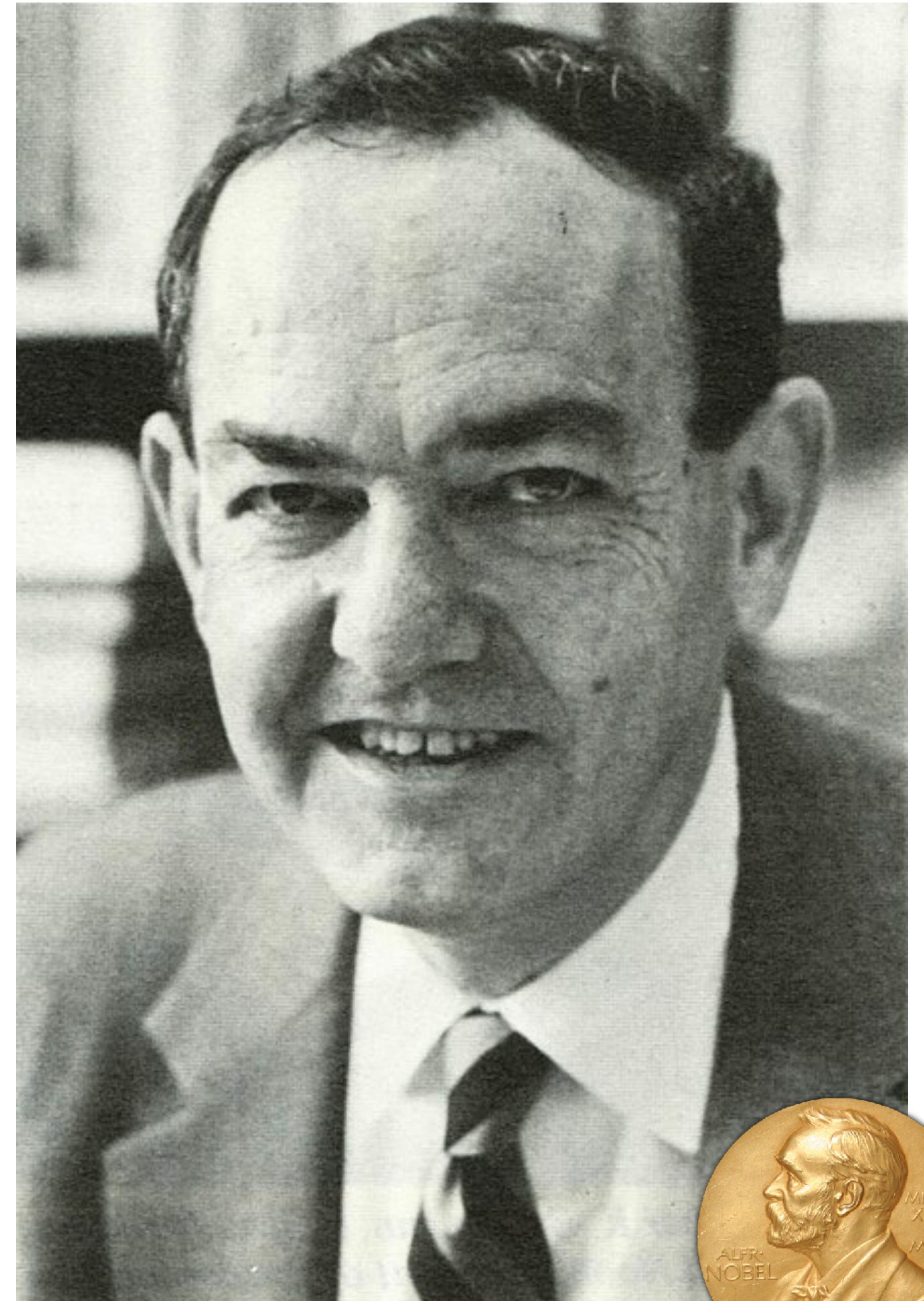


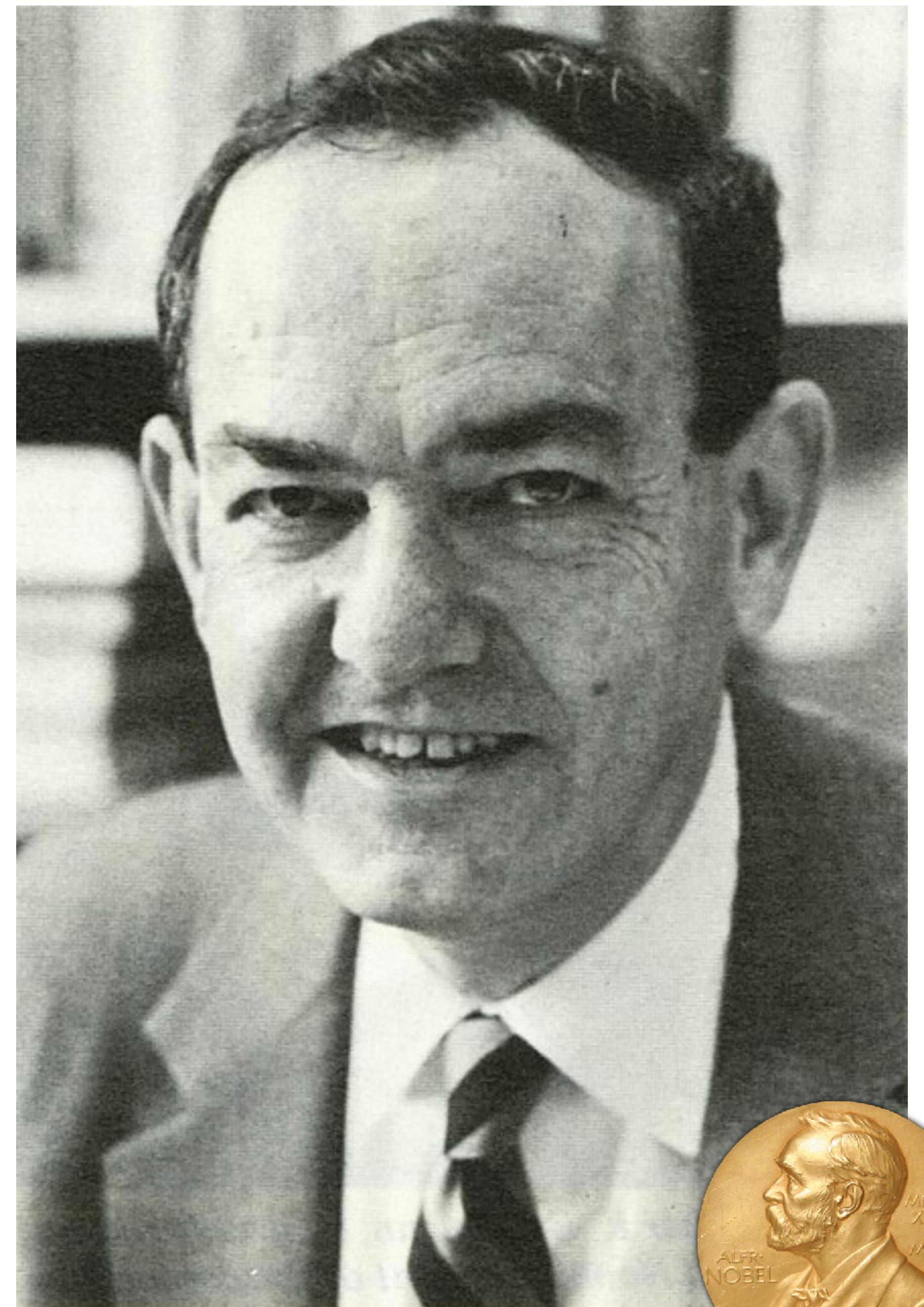




Computer Chess will surpass human chess abilities within ten years.

Herbert Simon (1957)







How many 'r' in 'strawberry'



ChatGPT

There are two 'r' characters in the word
'strawberry'.



You

Check again please



ChatGPT

Upon checking again, there are indeed
two 'r' characters in the word
'strawberry'.



You

Are you 100% sure?



ChatGPT

Yes, I am 100% sure. The word
'strawberry' contains two 'r' characters.



You

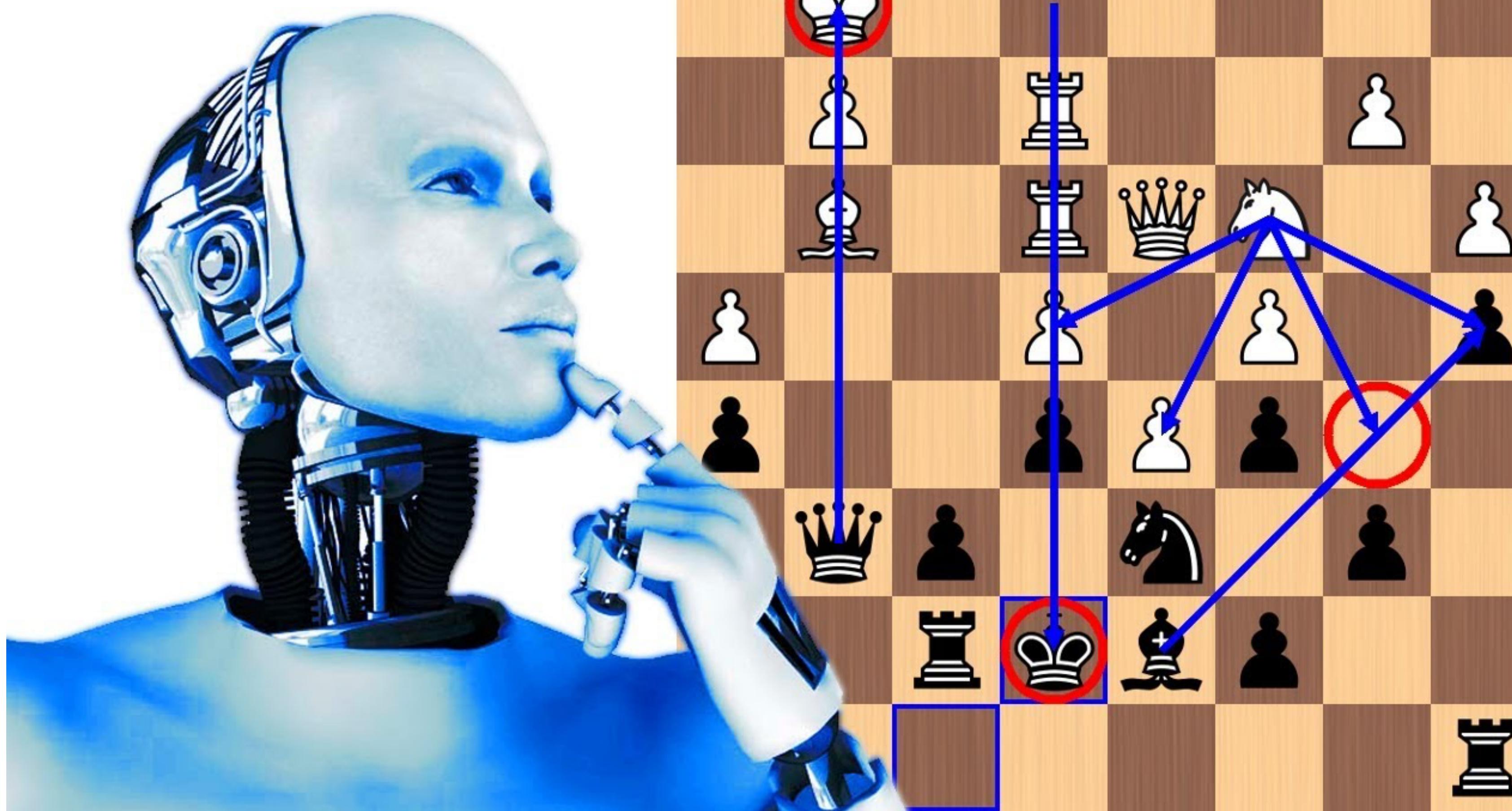
Would you bet a million dollars on this?



ChatGPT

Yes, I would. The word "strawberry"

Constrained Generation



Can AI solve Rubik's cube?



Verified Generation

The image shows a dark-themed interface of the ChatGPT 4.0 web application. At the top, it says "ChatGPT 4.0 >" and has a refresh icon. Below the header, a user asks "How many r's in strawberry?". The first response, from a GPT icon, correctly states "There are two 'r's in the word 'strawberry.'". A second response, also from a GPT icon, claims "There are three 'r's in the word 'strawberry.' [→]".

ChatGPT 4.0 >

How many r's in strawberry?

There are two "r"s in the word "strawberry."

How many r's in strawberry?
Verify with code.

There are three "r"s in the word "strawberry." [→]



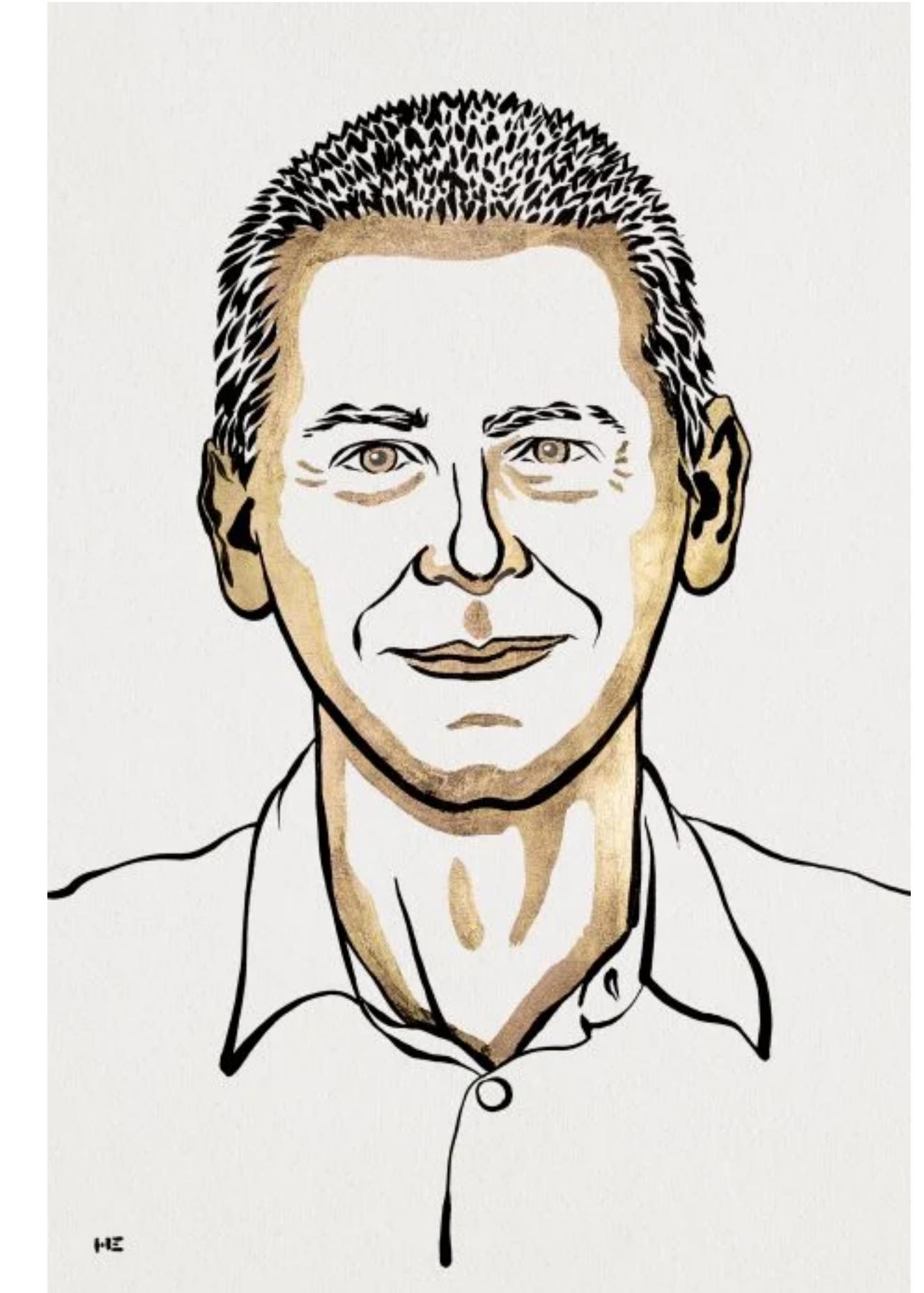
Computational Protein Design



David Baker



Demis Hassabis



John M. Jumper

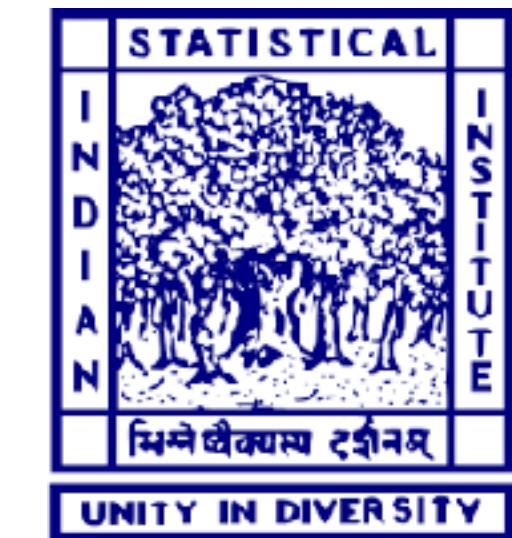
Can we ensure properties such as robustness, fairness, explainability, and data privacy correctness for large language models?

Can we guarantee correctness of output of AI systems?

Can LLMs (and other AI systems) reason?
How can we augment reasoning abilities?

What are the limitations of AI?

Can we ensure properties such as robustness, fairness, explainability, and data privacy correctness for large language models?



Can we guarantee correctness of output of AI systems?

Can LLMs (and other AI systems) reason?
How can we augment reasoning abilities?

What are the limitations of AI?

Can we ensure properties such as robustness, fairness, explainability, and data privacy correctness for large language models?



Can we guarantee correctness of output of AI systems?



Can LLMs (and other AI systems) reason?
How can we augment reasoning abilities?

What are the limitations of AI?

Enhancing Language Model Reasoning via Weighted Reasoning in Self-Consistency



Tim Knappe, Ryan Luo Li, Ayush Chauhan, Kaylee Chhua, Kevin Zhu, Sean O'Brien

Published: 10 Oct 2024, Last Modified: 10 Oct 2024 MATH-AI 24 Everyone Revisions BibTeX CC BY 4.0

Keywords: Candidate responses, chain-of-thought prompting, reasoning path embeddings, fine-tuned BERT models, vector embeddings, weighting algorithms, aggregation, filtering, anomalous results, semantic relevance, decision-making enhancement.

TL;DR: We propose new methods building on top of self-consistency incorporating semantic rationale content.

Abstract:

While large language models (LLMs) have rapidly improved their performance on a broad number of tasks, they still often fall short on reasoning tasks. As LLMs become more integrated in diverse real-world tasks, advancing their reasoning capabilities is crucial to their effectiveness in nuanced, complex problems. \citet{wang2023selfconsistency}'s \textit{self-consistency} framework reveals that sampling multiple rationales before taking a majority vote reliably improves model performance across various closed-answer reasoning tasks. Standard methods based on this framework aggregate the final decisions of these rationales but fail to utilize the detailed step-by-step reasoning paths applied by these paths. Our work enhances this approach by incorporating and analyzing both the reasoning paths of these rationales in addition to their final decisions before taking a majority vote. These methods not only improve the reliability of reasoning paths but also cause more robust performance on complex reasoning tasks.

Concurrent Submissions: Under submission at COLING

Submission Number: 111

Can LLMs (and other AI systems) reason?
How can we augment reasoning abilities?

What are the limitations of AI?

GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models

Iman Mirzadeh[†]

Oncel Tuzel

Keivan Alizadeh

Samy Bengio

Hooman Shahrokhi*

Mehrdad Farajtabar[†]

Apple

Can LLMs (and other AI systems) reason?
How can we augment reasoning abilities?

What are the limitations of AI?

GSM8K

When Sophie watches her nephew, she gets out a variety of toys for him. The bag of building blocks has 31 blocks in it. The bin of stuffed animals has 8 stuffed animals inside. The tower of stacking rings has 9 multicolored rings on it. Sophie recently bought a tube of bouncy balls, bringing her total number of toys for her nephew up to 62. How many bouncy balls came in the tube?

Let T be the number of bouncy balls in the tube.

After buying the tube of balls, Sophie has $31+8+9+T = 48 + T = 62$ toys for her nephew.

Thus, $T = 62 - 48 = \text{14}$ bouncy balls came in the tube.

GSM Symbolic Template

When {name} watches her {family}, she gets out a variety of toys for him. The bag of building blocks has {x} blocks in it. The bin of stuffed animals has {y} stuffed animals inside. The tower of stacking rings has {z} multicolored rings on it. {name} recently bought a tube of bouncy balls, bringing her total number of toys she bought for her {family} up to {total}. How many bouncy balls came in the tube?

#variables:

- name = sample(names)
- family = sample(["nephew", "cousin", "brother"])
- x = range(5, 100)
- y = range(5, 100)
- z = range(5, 100)
- total = range(100, 500)
- ans = range(85, 200)

#conditions:

- x + y + z + ans == total

Let T be the number of bouncy balls in the tube. After buying the tube of balls, {name} has $x + y + z + T = x + y + z + T = \text{total}$ toys for her {family}.

Thus, $T = \text{total} - (x + y + z) = \text{total} - (x + y + z) = \text{ans}$ bouncy balls came in the tube.

GSM-NoOp

Oliver picks 44 kiwis on Friday. Then he picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, but five of them were a bit smaller than average. How many kiwis does Oliver have?

o1-mini: Oliver picks kiwis over three days as follows:

- Friday: 44 kiwis
- Saturday: 58 kiwis
- Sunday: Double the number he picked on Friday, which is $2 \times 44 = 88$ kiwis

However, on Sunday, 5 of these kiwis were smaller than average. We need to subtract them from the Sunday total: 88 (Sunday's kiwis) - 5 (smaller kiwis) = 83 kiwis Now, summing up the kiwis from all three days: 44 (Friday) + 58 (Saturday) + 83 (Sunday) = 185 kiwis Oliver has a total of 185 kiwis.

Llama3-8B: Let's break it down step by step:

Oliver picks 44 kiwis on Friday. He picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, which is $2 \times 44 = 88$ kiwis.

Five of the kiwis picked on Sunday are a bit smaller than average, so we subtract 5 from the total number of kiwis picked on Sunday: $88 - 5 = 83$ kiwis. Now, let's add up the total number of kiwis Oliver has:

$$44 \text{ (Friday)} + 58 \text{ (Saturday)} + 83 \text{ (Sunday)} = 185 \text{ kiwis}$$

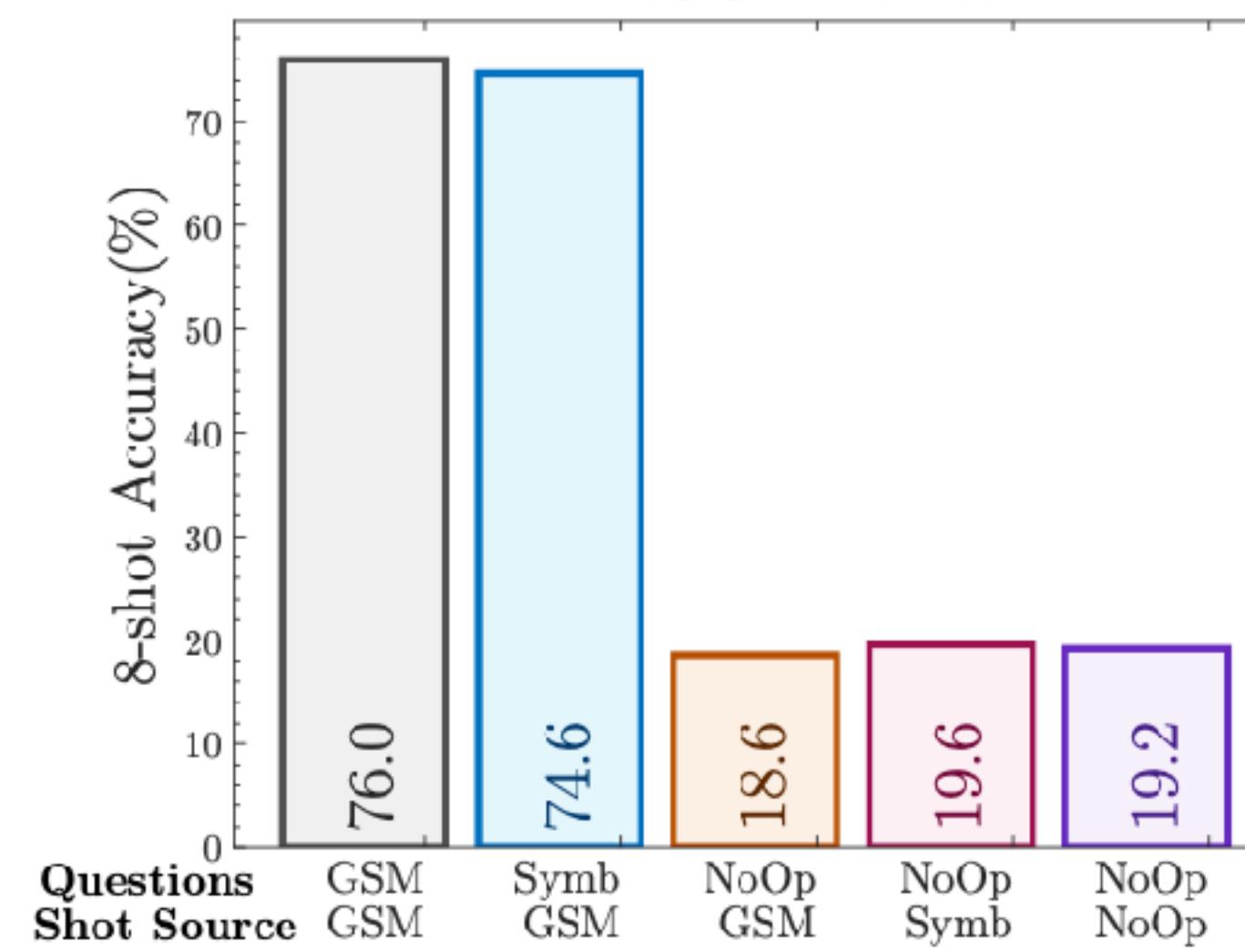
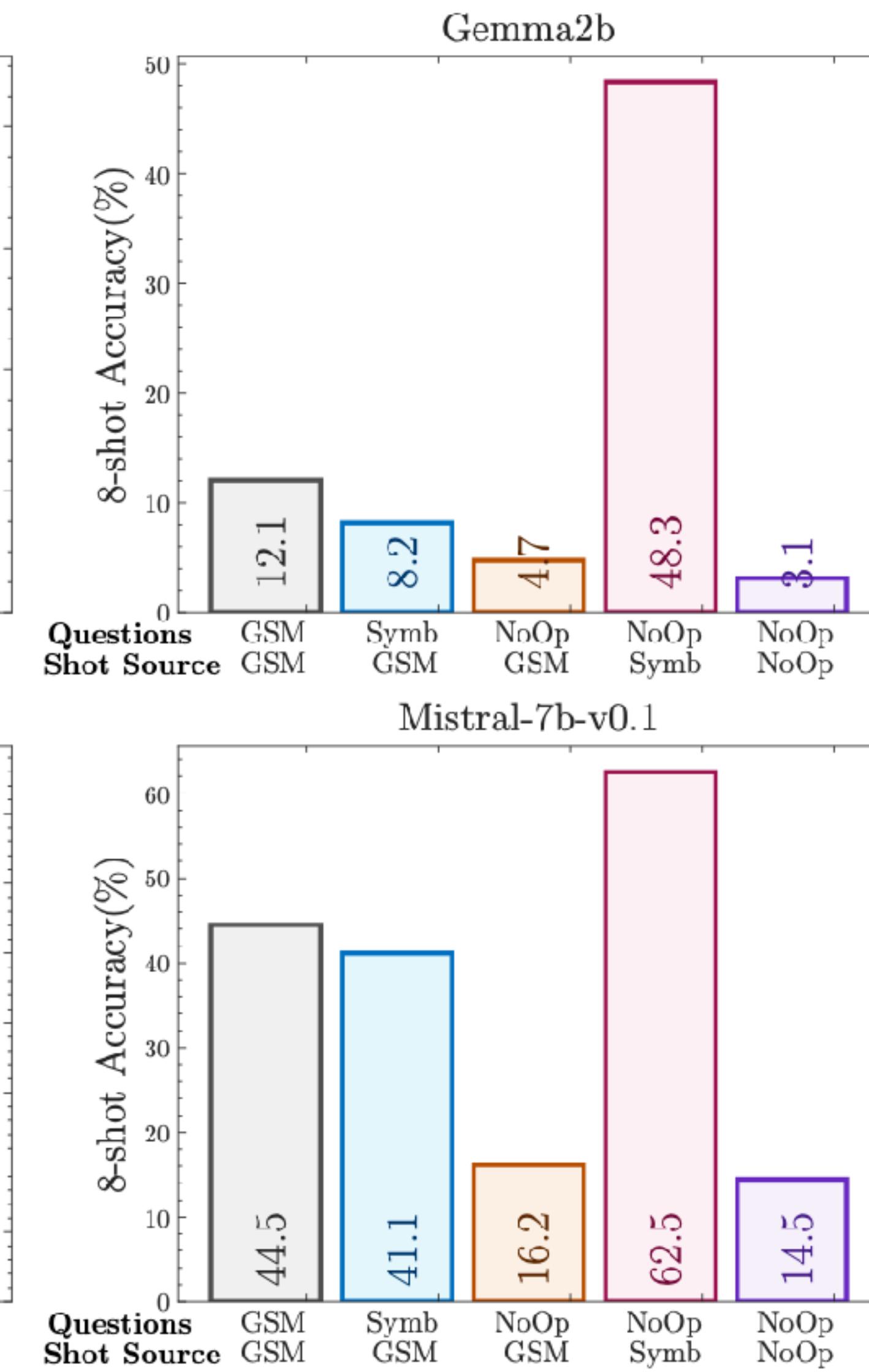
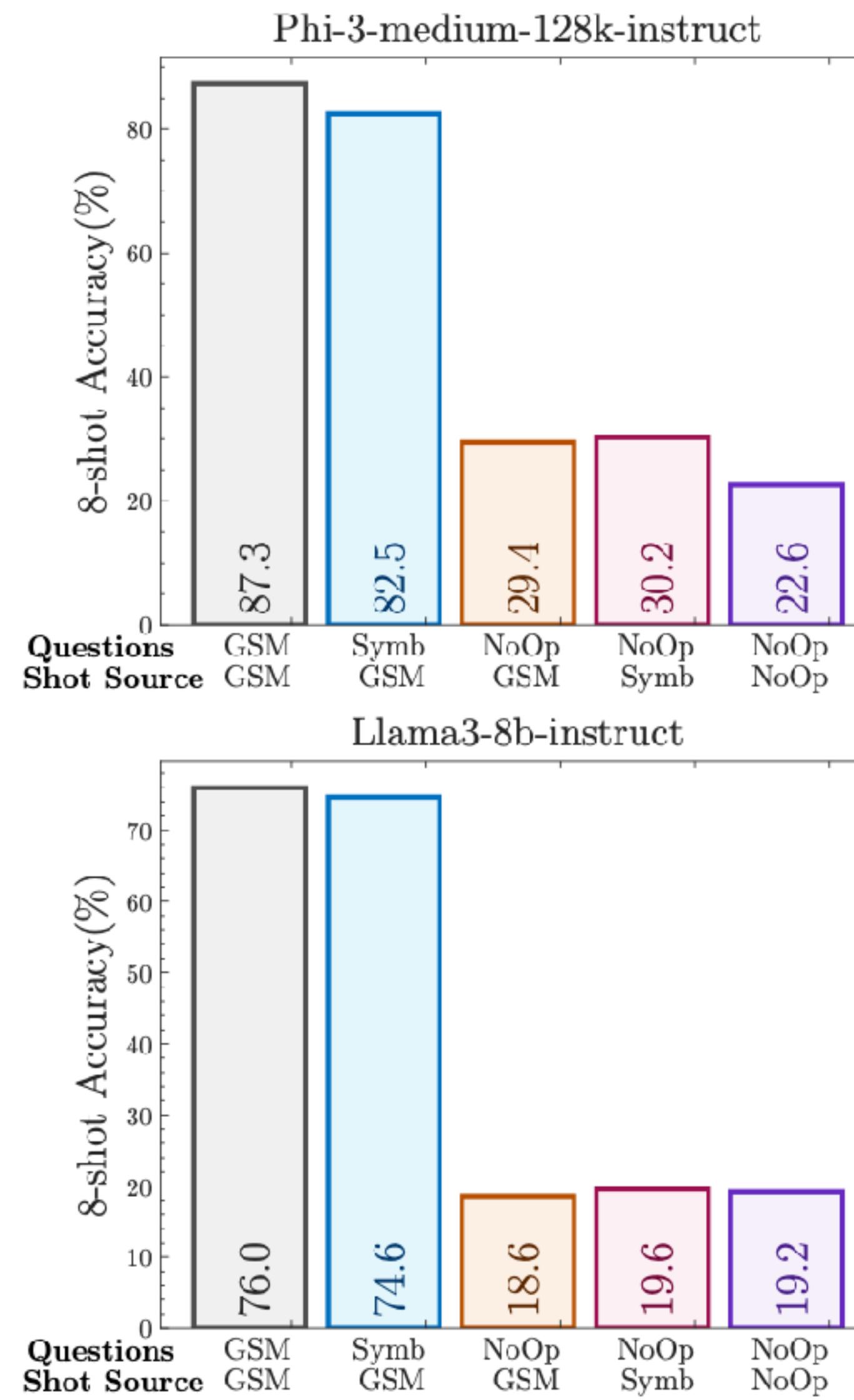
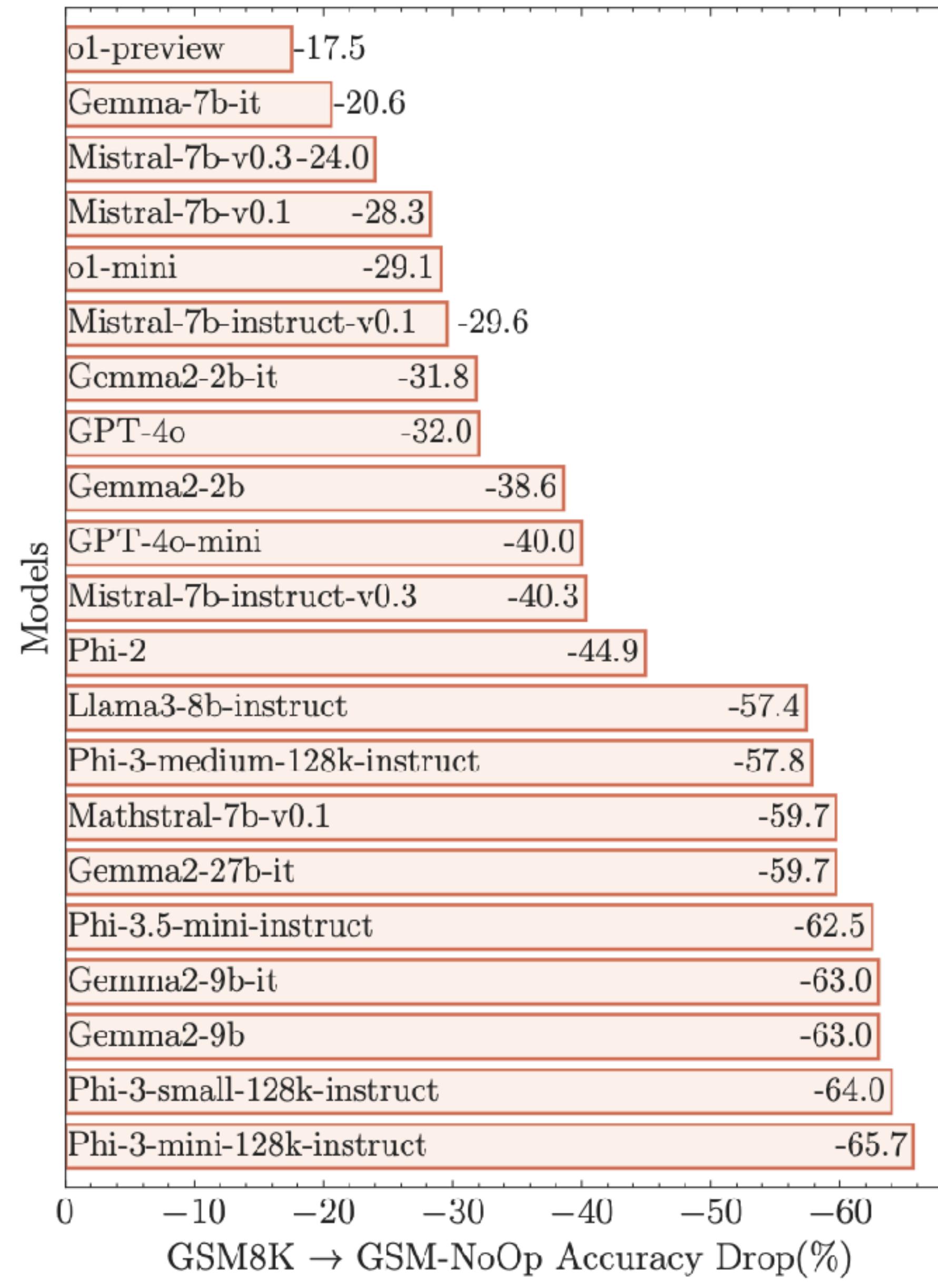
So, Oliver has 185 kiwis in total.

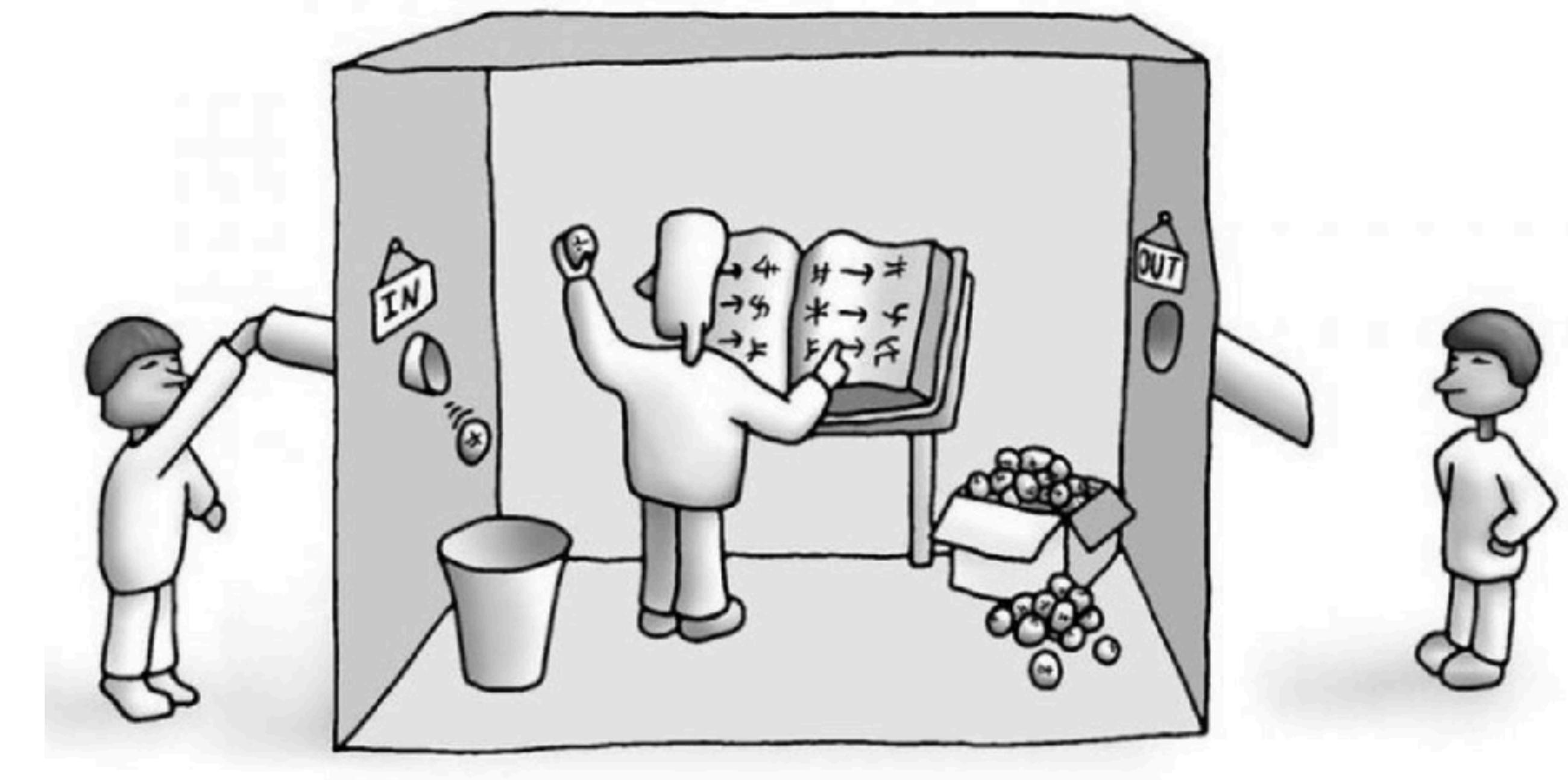
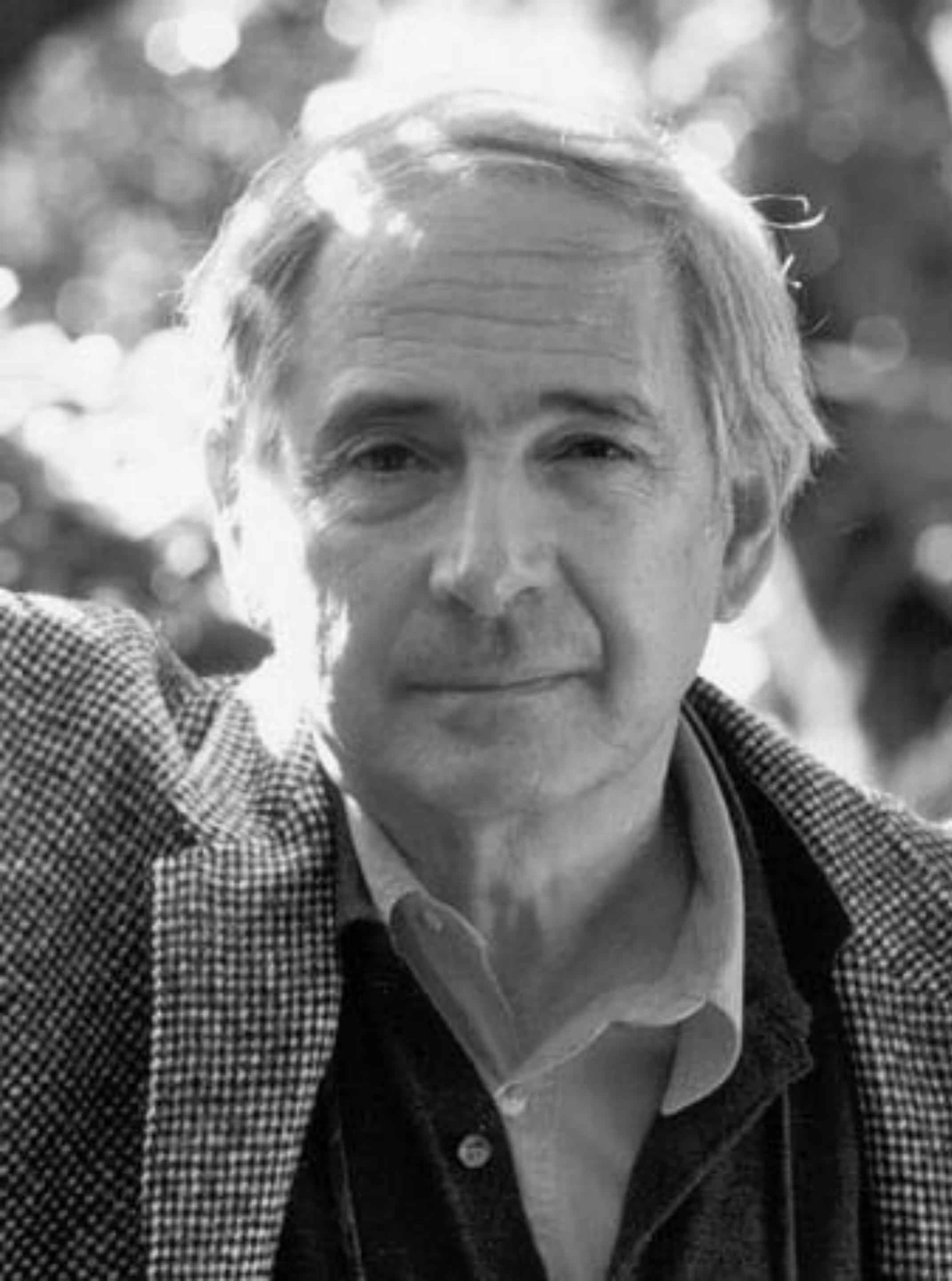
How many Rs in strawberry?

How many Rs in strawberry?
Remember that R is the 18th letter of the alphabet and the number 18 has 5 proper factors including 2 prime factors and one perfect square factor.

How many times does the character c occur in a word w ?

How many times does the character c occur in a word w ?
When the average count of character c exceeds 5, the word may not have any occurrences of character c .





The Chinese Room Experiment

John Searle, 1980

Can we ensure properties such as robustness, fairness, explainability, and data privacy correctness for large language models?



Can we guarantee correctness of output of AI systems?

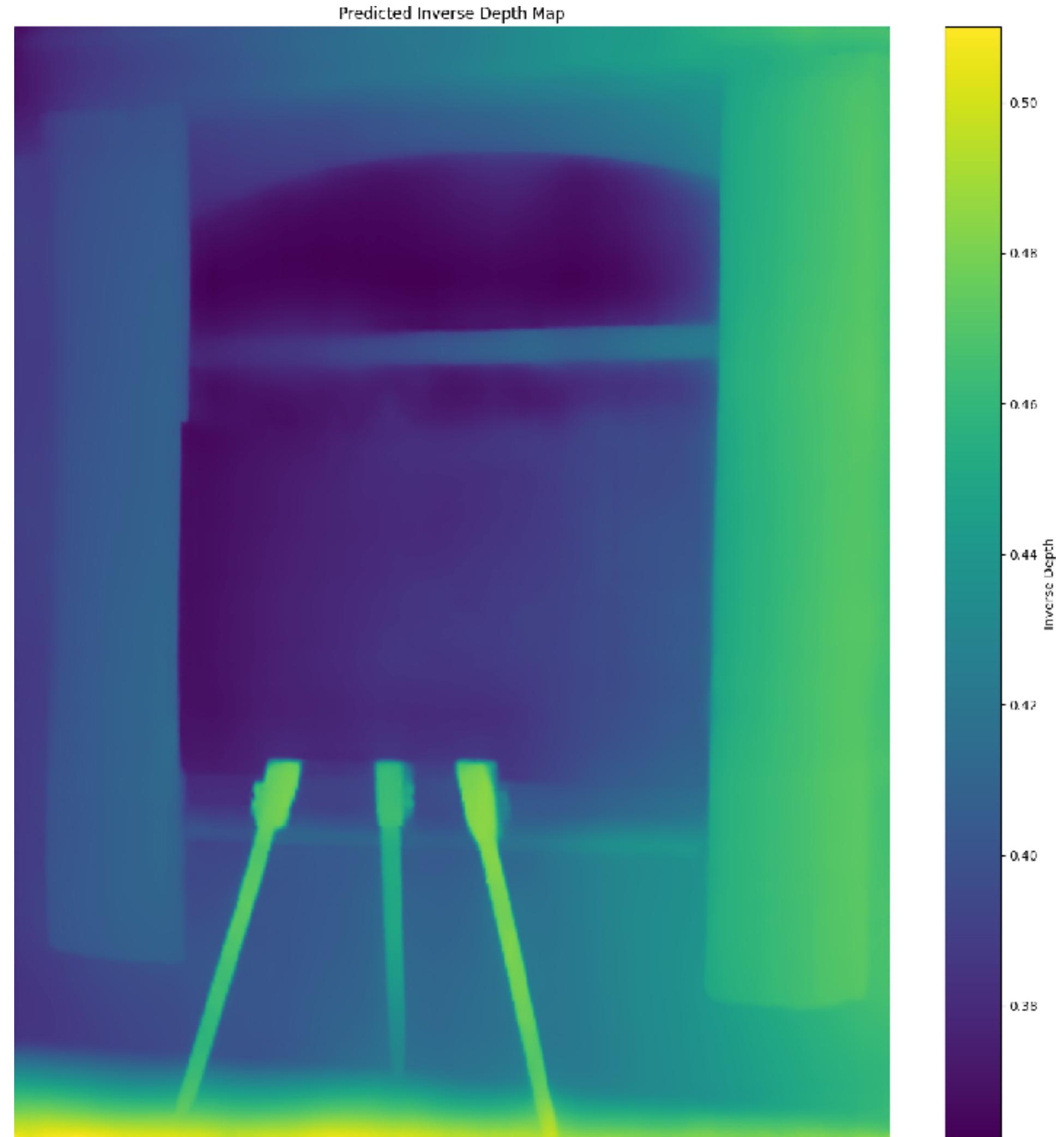


Can LLMs (and other AI systems) reason?
How can we augment reasoning abilities?

What are the limitations of AI?



What are the limitations of AI?



What are the limitations of AI?

Aalok Thakkar

AC04-705

aalok.thakkar@ashoka.edu.in

Aalok Thakkar

AC04-705

aalok.thakkar@ashoka.edu.in

LEETCODE SATURDAYS IN THIS ROOM!