Simon, Newell.

$p \to q \overset{(\equiv \mathrm{Def})}{=_D} -p \vee q$ Page 23 a.

$\to$ implies
$-$ not

$\vee$ or

p
q
r
s
} propositions.

dot convention.

rules

1. Substitution

2.

$a \to b$ ∴

if a, ^is a prop^ then b is a prop.

(replacement) (of a symbol by its def.)

3.

2) $a \to b$ means if a is a valid prop., then b is. (detachment) [not used at first

3) replacement of a symbol by its definition. used much at first

$\vdash$ means that a prop. is being asserted (Russel)

! will be used here instead of $\vdash$.

1.2 $(p \vee p) \to p$ Axiom.

$(-p \vee -p) \to -p$

$(p \to -p) \to -p$

---

1) Find "similar" Theorems. (Page 29)
2) Match them. (?)

---

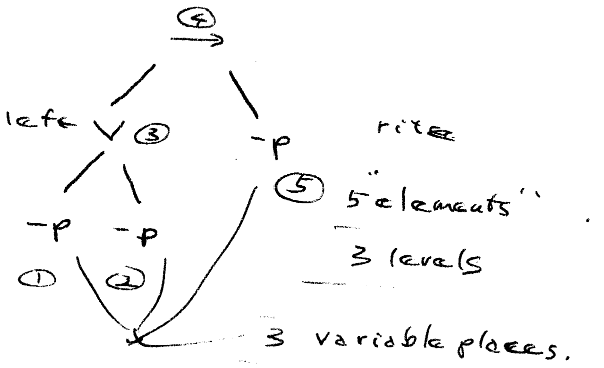
p 29

$(p \vee p) \to p$

K = no. levels. 3
J = no. names 1
H = no. variable places 3

tree notation →→

1 name.



"5 elements"

3 levels

3 variable places.

description is K J H
3 1 3.

① look for Theorems with v descriptions then try substitutions. (3):
Will prove many of Th. Theorems in Chapter 2 of Principia (about $\frac{1}{3}$).

try to prove

$(p \to -p) \to -p$

$\dot{A}$ (~~$\dot{p}$~~) means subs. has not been made.

$(A \vee A) \to A$

$(-B \vee -B) \to -B$

$(B \to -B) \to -B$

$\therefore p \to -p \to -p$

Page 38

2.06 ~~$(p \to q) \to ((q \to r) \to (p \to r))$~~

? $(p \to q) \to ((q \to r)\ (p \to r))$

[place paren. at bigest no. of dots first.

detachment

using 6 $\left(\begin{matrix} a \\ a \to b \end{matrix} \right) \to b$.

"similarity" means only Th./ ~~no.~~ triple descriptions are identical.

Machine would <u>not</u> consider

$a \to (b \to a)$

and $(p \to q) \to (b \to (p \to q))$

---

using detachment:
to prove b, try to find a $p \to q$ ∋ $q \sim b$ then xfm to $c \to b$ and try to prove c.

used around page 40

(1.3

1. substitution method

2. Detachment " p 350 , 39

3. Chaining method . page 41

to prove a → c

prove a → b and b → c

one can work from both ends.

Lattice theo.	▽	⌢	⌣				null	univ. classes.
Set algeb.	−	∩	∪		≤	≥	= 0	I
Intuitiv.	~	∧	∨	⇒ ⇐ ⇔	⊂	⊃	= Λ	V
Prop. Calc.	¬	&	♂	⇒ ⇐ ⇔	⊢	⊣	⊣⊢ 0	I
Sw. ckts	▽	·	+					

$* \quad p \overset{\vdash}{\cancel{\#}} q$ means "we haven't yet proved $p \vdash q$".

---

no. of stars befor a prop. indicates no. of assumptions it uses.

~~Auto branching~~. branching by simple process.

*	a			
*	*	b		
*	*	*	c	
*	*	*	×	d

1. substitution of any valid expression 2.
2. Deletion (ommiting intermediate ~~results~~ steps).

e.g. if $A \supset B \supset C \supset D$

Then $A \supset D$.

Some starting points:

$\cdots * \; p \wedge q$

$\cdots * \; p$

$* \; p \wedge q$

$* \; q$

$\cdots *$

$\cdots * * \; p$

$\cdots * * \; q$

$\cdots * \; p \vdash q$ ← indip of q.

$\cdots * \cdots * \; p$

$p \vdash q$

$* \; q$

Quine
Fitch
Cope
} proposed this method ~ 1952. (?)

To prove

$$(p \wedge q) \vdash (q \wedge p)$$

$\vdash$ suggest to look at $p \wedge q$ (because $\vdash$ is a main connective)

assume $p \wedge q$

then
* $p \wedge q$
* p
* $p \wedge q$
* q
* p
* $q \wedge p$

Then deletion :

*$(p \wedge q) \vdash (q \wedge p)$

Also work backwards

↑

↑

* q
* p
* $q \wedge p$

$(p \wedge q) \vdash (q \wedge p)$

This connects with

* $p \wedge q$
* p
* $p \wedge q$
* q

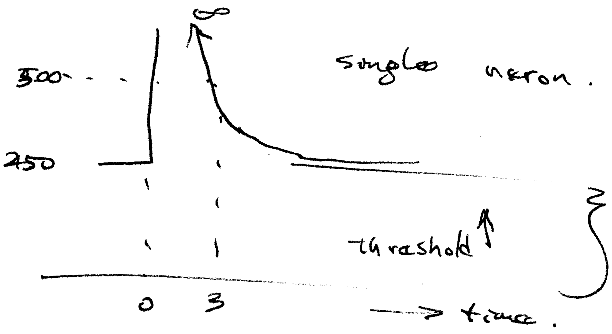
Simon, Newell.

subs. method

1) Str. first, most imp. metric.
2) If not matching connectives, try $A \to \sim A$
3) " " " try subs. definitions for ~~some~~ non-matching connectives)

N. Rochester:

Attempt to check Hebb's model of neural behavior with ~~701~~.



{ 250 is max output of neuron
" " " may be as low as 75.

each neuron ~~stims.~~ 10 others (completely random.

6 by ave. 63 are internal.

---

many { Axon
efferents

each neuron has ~~many~~ 10 outputs levels — 1 for each neuron.
it syns. upon.

{ Output ↑ 1 every time ~~A~~ from A to B synapses B. ~~Every~~ once in a while the Σ output are reduced to some fixed level ~~by subtraction of same small amt. from each~~

1) state of recovery
2) " " fatigue } 2 nos.

rel. size of 1, a, b is imp.

Fatigue ~~of~~ state is added to ~~B~~ threshold.

thresh ↑ a every firing
" ↓ b " at <u>not</u> firing. } but upper and lower limits.

Hebb wanted to account for

1) Lashley's results on brain exterpation,

~ 10K neurons in a "cell assembly". "Arrousal of "cell assembly" is used for short → time memory. Cell assembles are not <u>physically</u> localized.

inputs:
25%
12%
12%
50% random.

Model is for discrimination a/o abstraction

When A fires B many times A tends to fire B more easily

trys of ~~[illegible] perturbed~~ not perturbed slitely,
— nets diverged from original rapidly.

conclusion was that th. net would need more non-random structure if it was to exhibit cell-assembly.

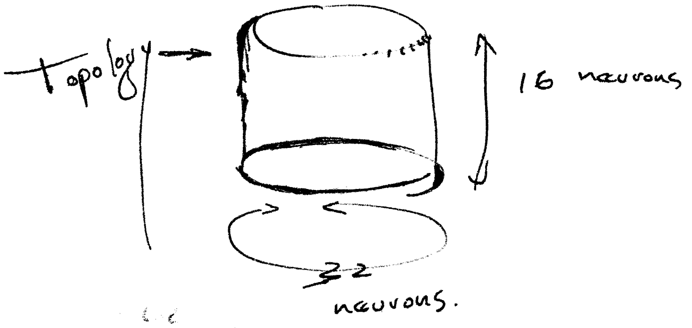
Second Model **704**   John Holland, with Hebb.

both exitory, inhibitory stims: internal ~~[illegible]~~ activity of cell exitory. Betw. cells — inhibitory.

Output was betw. −1, +1. with ~ 30 steps.

512 neurons used. Each neuron has a frequency of firing assoc. with it. Also distance bias put in.

Random connections except within distance of 8.



4 areas chosen to act as receptors.

connections ~~[illegible]~~ become exitory when ≈ cross corr. ~~[illegible]~~ >1
inhib when " " " ~~[illegible]~~ ≈ 0.

There was development of localized, connected ~~[illegible]~~ regions. ~~[illegible]~~ groups had 50-75 neurons in them.

McCarthy

a_i ~~[illegible]~~ constants. ← "Smulyan Formal System"

x_i vars.

L_i Lables. (predicates)

⟶ implication

w ff. ≡ well formed formula $A_1 \to A_2 \to A_3$

where Th A's are strings without arrows.

Axioms : Some set of w ff's.

Rules of inference

1) Detachment

~~A~~ A

$A \to \underline{X}$

gives $\underline{X}$

A is always a string that doesnt involve →

$\underline{X}$ " " " may " ".

2) substitution : any string in Th. A's may be sub.

for any x

~~A set o~~

Let $S(L_i)$ be Th. set of wf formulas in Th. a's ("words") $\ni$ only $L_i A$ is a theorem.

Metatheorem: Th. set of theorems in any logical lang. is $S(L_i)$ for some elementary formal system.

Problems: Trig. identities.

McC wants machine to keep record of trials that were used and score kept — This will govern order of trials

Then, I think he wants to go to hyer and hyer order computers; ~~[illegible]~~ Each Level writes trial programs for Th. next lower level; derives usefulness of each heirarchical program on Th. basis of experience. This gives machine that will improve itself.

1) how to perform induction.
2) Th. min. machine and info packing approach
3) How ~~solve to Th.~~ solu. to Th. induction prob. gives solu. to search problem
   a) by suitable stat. a's.
   b) by using some of Th. abs's. of induction, for search.
4) A M.C. method of realization.
5) Way to divide induction problem into 2 parts. — a) Abs. construction
   b) B.G.

There seems to be big hassle about ~~whether~~ terminology of "programs" at various levels.

McC is much interested in the Smolyan system because it can refer to itself — presumably so that the lower order machines can be used to improve the hyer.

One idea ~~of~~ suggested method of proof is, say in finite group theory — to try sub-theorems on a specific group.

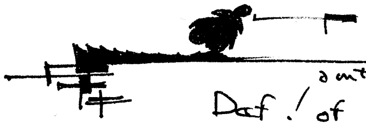
1) Try to find methods to improve proof search procedures.

~~Use~~ i.e. Make machines' task to improve own programs to improve own programs ....
It must also have some <u>other</u> task. —
McC believes it should be the Smolyan formal systems, because they are simple, and yet very general. This is <u>very</u> improbable, since all the info. the machine could possibly gain, would be info about these Smolyan formal systems. ~~It is~~ improbable that such rules about program improvement would be useful in, say, solving ~~formal logic pr~~ Trig. identities.

McC Mentions methods of deciding upon sub-goals.

Minsky

Samuel's checker player: IBM 704 – uses all 36 bits of each word – plays all sequences 5 moves out. Evaluates moves for best evaluation on ~~e~~ end of 5 moves.



Def. amt. of knowledge

Defs. of info: how intuitive idea is simply clue as to what mite be useful. – that "true" def. is dependant upon usefulness in some situation.

M's Goal: to define self-awareness and see how much a machine can gain by introspection.

Actually what he is saying:

1) Psych. Theorys and Physics Theorys are much different in content, method. This is because Psych is more complex – in th. sense that ~~observable behavior~~ observable macro behavior in physics is not distant from micro behavior. One can easilly combine atomic laws to get ~~many~~ reasonably relyable macro laws. In psych, however, ~~human~~ Animal behavior is a very complex ~~[illegible]~~ function of atomic behavior. So one make semi-macro behavior entities in Psychology. These entities, like "Id, Goal, reenforcemen" are entirely different in their laws of combination, from atoms.

2) It is .·. reasonable that man should ~~[illegible]~~ consider mind behavior and ~~body~~ physics ~~behavior~~ ~~diff~~ as essentially different, ~~[illegible]~~ since they have entirely different kinds of theories for their behavior.

3) This "dualist" state should exist untill science advances to the point where ~~[illegible]~~ the bridge betw. Human behavior and atomic behavior is made. At this point, scientists should become monists. Untill then, they may be monists by faith – ie. they may think that ~~[illegible]~~ monism is likely, but they have not yet proved it.

3.2

Selfridge

says he is interested in Th. learning process. While his general statements are good, his specific statements and projects don't seem to be very good. → Realizes that pattern learning in itself is not imp., but methods of learning is.

→ This was also very true of his west coast comp. conf. paper.

Tabula Rasa

1) T. R. ~~[illegible]~~ (≡ Blank sheet of paper)

2) R. M. Reasonable machine (Game player)
   Has
   a) Lang. : about world.

S. mentioned desire to build a T.M. befor "others" did; since he trusts himself and not "them".

Imp. words used in project by many

"Matric" ~~mean~~ ~~the~~

"Sub Goal"

Talks about what sentences are close to gn. sentence.

1) trys substitution of ~ words for other words.

2) " " " " phrases " " " phrases.

What is a reasonable lang. for a game:

Consider 5-in-a-row.

[Rouse Ball Math. rec. also Kratchez] says 1st player can always win, ~~but and then~~

∴ 4-in-a-row ~~etc~~ is a useful word.

3 " " " " " " " .

Also it is bad to let him get (5 in a row)
(4 " " ")

S. feels that deductive logic is rather useless.

Subsitution	Substitution
Datatchment	Datachment
Chaining	Probably chaining
~~Mara~~ S and N.	More : Intends to extend to further, hyperlogical operations.

Some basic philosophical orientations.

1) Randomness is unimp. except as simple programing method (sometimes).

2) Math. logic is unimp. ~~as~~ except as an inversion problem

3) For intelligent behavior either

   a) Universe must be special to fit machine

   b) or machine must be gn. info about universe.

This suggests that McC's machine that ~~works~~ was brought up only on Smulyan systems is unlikely to do well in other fields. — unless by coincidence — I.e. physical world is not ∪ to logic.

4) Cannonical methods

   a) Factoring. (str., ntpsts)

   b) Evaluation of factors.

   c) Inverses as a method to get strs, ntpsts

   d) Some strs, ntpsts. to start with.

5) Occam's Razor — just how it is tautological — the idea of simplicity [look at old lecture notes]

6) Idea of working on many humanly solvable problems likely, to find out what they have in common, rather than initial intense work on any 1 problem. Emphasize how human works problem, rather than any easy trick ways ~~to work~~ problem that may be peculiar ~~to that sort of~~ problems. This has been (for my own ~~orientation~~) ~~a~~ great overemphasis of ~~special problems~~ peculiar difficulties peculiar to th. partic. probs. worked on rather than ~~probs~~ diffs. peculiar to all probs. Use of computers is, I think, usually going in more deeply than is good at th. present time — getting little ~~info~~ of th. rite kind of info, per arg expanded. This ~~is~~ is also a strong arg. v.s. chess.

7) Recent approach via evolution and non-redundant coding

How this evolved into Neural net approach.

8) Some specific Machines.

   a) Th. hyororder reenf. Machine from "Notes on M",

   b) Th. Math. machine.

   c) Th. evolutionary Machin

   d) " devel. of th. random net machine.

neural nets

non-redundant coding for evolution T.M.

Th. reenf. Machine in "Notes on Minsk";

Method of eval. of efficacy of prediction.

Soln. of B.G.

Blinking Gambler?

Organization

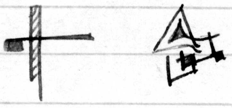
M.C. Methods.

1) Gen. Phil. orientation
- a) Randomness unimp. – is artifact ~~as~~ assoc. with any unconcious search procedure
- b) Computer must have much info on world built in
- c) Emphasis on learning in general, ~~rather~~ than peculiarities of individual problems. 1. Meatron Samuels 2.
- d) Math logic isn't too imp.
- e) Use of definitions – ~~as~~ rigorization of intuitive def. as good approx. for useful sc. object

how this applies to "info." "learning" "thinking"

Occam's Razor into search for good trials

2) Some learning problems — Problem of 1st or 2nd kind. ← Good Terms
- a) The search problem (how to get quick solns)
- b) Time series prediction
  - ① Continuous
  - ② Double digital.
- c) Chess, other games.



3) What is learning?
- a) Use problem of 2nd kind as example.
- b) Division into problem of ① Expected G of a gn. try ② Experimental problem.

What are probs. of interest?

1) "search" Prob. of 1st kind.
2) " " " 2nd kind. (optimization)
3) Prediction : Time series [or multi-dim]
   - a) continuous
   - b) semi continuous
   - c) doublely digital.
4) Evolution : as sort of problem of 2nd kind.

Go into self-reprod. machines and "low-redundance" coding.

Occam's razor

Sc. method: 1) Start with basic set of entities, methods of combining

2) **Each** entity, method, has params. assoc. with it.

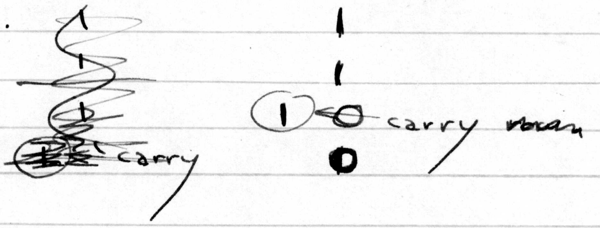
3) **Try** weighted combinations in ~~science~~ problem; weights depend on params. reeval params. of components in view of efficacy in problem Solving.

4) Using new entities and methods and new evals. of old ones, try solving new problems.

5) Continue.

Explain a "gp" as used in prediction.

Give example from Math T.M. of what prob. is for =, ~, $\beta\Sigma$, $\beta\Pi$, +.



Use $\boxed{?}$

some initial gps.
" " ~~ntpsts~~ ngms
" " ntpsts
" " strs. , strsts

How to get new entities from old.

$\beta\Sigma$, $\beta\Pi$, Cart Π (for ntpsts), convolution for strs,

---

The concept of factoring:

$$\begin{pmatrix}1 & B & 1\\ 1 & B & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}, \begin{pmatrix}1 & B & B & 1\\ 1 & B & B & 1\end{pmatrix}$$

$$= \left[(1,1),(1,B,1),(1,B,B,1)\right] \times \begin{pmatrix}1\\1\end{pmatrix}.$$

$$(1B1) \times \left[\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix}\right] =$$

$$\begin{pmatrix}1 & B & 1\\ 1 & B & 1\end{pmatrix}, \begin{pmatrix}1 & B & 0\\ 1 & B & 0\end{pmatrix}, \begin{pmatrix}0 & B & 0\\ 0 & B & 0\end{pmatrix}, \begin{pmatrix}0 & B & 1\\ 0 & B & 1\end{pmatrix}$$

How analogy can be created by str, ntpst.
The use of definition by inversion.

M.C. methods:

1. simple search. ~~non~~-(exhaustive) [$\frac{n}{2}$ trials]
2. " " (non- ") [n "] Ashby
3. weighted " (non ") [n "]
4. weighted " (exhaustive) ?
5. by λ non-linearly weighted search, exhaustive
6. Ideal, ordered exhaustive. [first moment]

only powers do this.

Probably M.C. search worked on by Op. res. theory.

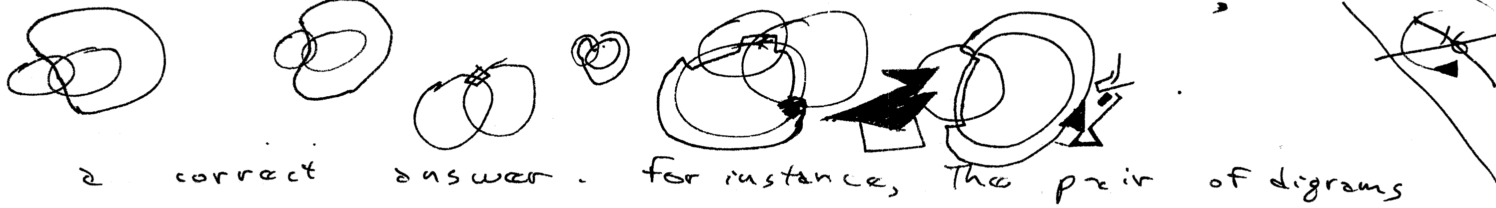
March 56

Additional Note on my approach:

1) Explain concept of "simplicity" and ~~un~~ad-hockness, giving various examples.

2) try to show that Ockam's razor is a trivial statement - i.e. one couldn't <u>possibly</u> try all complex combinations of old good ideas, because there are too many.

3) Application to problem of how many params. to use in curve fitting, and fallacy of dividing data in 2 - one part for param. determination, and 1 part for checking.



a correct answer. for instance, the pair of digrams $\underset{\square}{\boxed{1}}$ and $\underset{\square}{\boxed{0}}$ would give probabilities of 50% for both 0 and 1. This question bears a strong resemblance to ~~than prob~~ that of the 2 prize fiters with 80% and 70% wins, respectively.

(SN) What about use of th. digram pair $\underset{\square}{\boxed{1}}$, $\boxed{0}\boxed{1}$? Is this legal? I think it is "legal", ~~and that any gp. set can be used;~~ can one assign → prediction params. to single gps. or can one just do so for pairs of gps? Note that within this system th. digram $\overset{|}{\boxed{0}}$ is very useful in prediction, and it makes many other digms useful — say $\boxed{1}\,1$. If $\overset{|}{\boxed{0}}$ is used, $\overset{|}{\boxed{1}}$ becomes no more useful than $\boxed{1}\,1$. As long as there are only th. 2 possibilities of 0 or 1, $\overset{|}{\boxed{0}}$ is of much more U than $\overset{|}{\boxed{1}}$.

2 × 8 × 3 | what about $\overset{o}{\boxed{1}}$ v.s. $\boxed{0}\,0$? Shouldn't this be as good as $\overset{o}{\boxed{1}}$ v.s. $\overset{o}{\boxed{0}}$ ← which is th. same as asking what probty. distrib. one gets for $\overset{\bullet}{\square}$. One cannot ask an equivalent Q for.

This is perhaps why one feels that $\overset{o}{\boxed{1}}$ and $\overset{o}{\boxed{0}}$ ought to be imp. Also, $\overset{o}{\boxed{0}}$ is of hyer aprip than $\overset{o}{\boxed{1}}$, (tho not much)

Paying Th. weatherman.

$f(q_k)$ If we want to pay $f(q_k)$

p_k where k is Th. weather that actually occured.

$$\Sigma\, p_k\, f(q_k) + \lambda \Sigma\, q_k$$

$$p_k\, f'(q_k) + \lambda = 0$$
$$f'(q_k) = \frac{-\lambda}{p_k}$$

if we want $q_k = p_k$. $f'(p_k) = -\frac{\lambda}{p_k}$

If we use

payoff of $f_j(q_1 \cdots q_k)$
if j^th^ event occures.

Then $\sum_j p_j\, f_j(\vec{q}) + \lambda \Sigma\, q_j$ maximized.

$\frac{\partial}{\partial q_r}$: $\sum_{\sigma} p_j \frac{\partial f_j}{\partial r}(\vec{q}) + \lambda = 0$

we want f's $\ni$ $\vec{q} = \vec{p}$ ← "true probies".

so $\sum_j q_j \frac{\partial f_j}{\partial r}(\vec{q}) + \lambda = 0$.

soln. is $f_j = \frac{\partial}{\partial q_j} F(\vec{q})$. where $f(\vec{q})$ is homogeneous of 1st °.

Expected payoff $= \sum_j p_j \frac{\partial f}{\partial p_j}$ which $= f$. (By Eulers rule)

$$\frac{\partial^2 f}{\partial q_i \partial q_j} x_i x_j > 0$$

so f must be simply convex.

This has direct intuitive significance.

~~Are there 2 different cost of observation functions for Th. weatherman~~, are there 2 different $f(\vec{q})$'s $\ni$ his ~~predictions~~ expenditures on different kinds of observations will be Th. same?

Also, if Th. payoff funct. for Th. weatherman is constrained (say >0), what is soln. ?

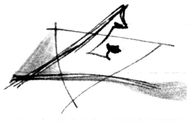
For Simon, Newell.

1) Improvement of program by comparing it with their own intuitive processes.

2) They **should** compare their own program with More's in as great generality as possl.

McCarthy.

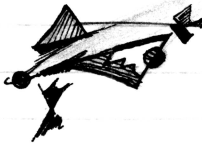
Objectives of project:



To ~~devise~~ program machine to ac

McC is still involved with concept of simplicity, quickness of machine as criterion of intelligence, rather than goodness of soln. He is thinking of "well defined" problems.

McC. still doesn't think probability of much use.



Few advantages listed for various possl. problems.

Design problems.

Inventing possible Theorems

That any of proposed proof machines

Certain probability problems: Involve taking certain verbal descriptions of ~~numbers~~ integers, and translating them into integers. E.G. "All integers <5, ≥1" would give 5. Here we try a set of xfms. on th. verbal expression, in attempts to get it into ~~an~~ one of a set of expressions that has a known no. assoc. with it.

This, then, is "a problem of th. 1st kind"

Th. method of Soln. is very ~ to that of Simon-Newell-More with McC, Minsky refinement in particular.

NOTE: We have, then a kind of machine that can be designed to work very many types of problems of th. 1st kind. That a large class of proby. probs can be put in this form is imp. It mite be well to make a list of problems that can be put into this form. Altho these machines are not very "creative" they will work most problems much more rapidly than most humans who work on them.

General note on Neurological Models of behavior proposed by psychologists: That they tend to emphasize the "law of Use" [I.E. if an animal does a certain act many times, he will be more likely to do it in th. future.] This view does not take reinforcement — or results of this action — into account. Wiener's theory of th. affective tone mechanism is about th. next level of sophistication.

I think Hebb's theory was a simple "law of Use" theory.

[this is confirmed on th. bottom of p4 and top of 5]

☐ ☐ ☐

8.1

M

Ⓐ ⊃ B

B

G = "Good"?

~~Good~~ a) G eval. for statements

b) → " " " terms

~~if we have a)~~

a term is G if it can be easily used generate G things.

A term can be G in 3 ways (at least).

1) simplifying expression of other existing G terms.

2) simplifying expression of many old theorems

3) Expressing old statements using this term, enables one to prove theorems quicker, easier.

ΣN) 2 and 3 may usually ~~be~~ imply one another.

8.2

A statement is G if: (It is a suitable goal or sub-goal)

1) It is th. theorem to be proved (G = ∞)

2) " " "close" to a G statement.

[closeness has several different meanings]

a) Logical equivalence.(?) [trouble is that if 1 is G, the other may not be G, because of its unwealdyness.

b) Assymmetrical closeness. A is close to B if B can be somewhat easily proved from A.

c) That 2 theorems (combined with some other fixed statements) can be both used to prove (using standard machine methods) Fixed theorem T, in about th. same no. of steps.

Another way to say this, is that these 2 theorems are about the same (upward) distance from T on th. part. ord. Lattice. If 2 theorems are th. same dist. from many T_i's, then they have, in a sense, th. same "coördinates".

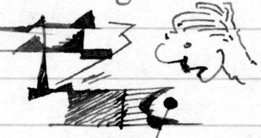
Actually, th. idea of G for th. "free" computer is signif. different from that for th. computer that has to prove a particular Theorem. In th. "bound" case, the G of an intermediate result is related to th. closeness to th. terminal expression.

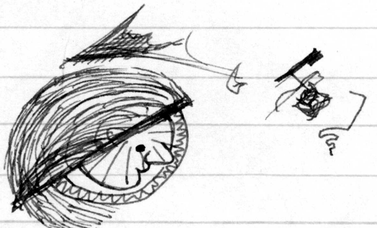
In th. "free" case th. G is defined in terms of "simplicity", "brevity", (using "simple terms"). Also by statements from th. operator ~~about~~ how G an expression

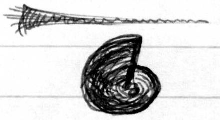
Thurs June 28, 1956

[SN] ~~II~~ Make small gp. study to determine ways to get all members of gp. ~~to~~ motivated {by proper words, cues, encouragements}.

Minsky:







Machine for proving theorems

1) Input is logical statement

2) First process is standard "logical clean-up" mechanism, that is used at various stages of proof.

$\sim\sim A \to A$

$\sim A \cdot A \to I$.

3) statement then goes to "Characterizer" that ~~gives~~ puts statement into 1 or more of several important categories. Produces string of nos. to give apri categories. Th. categorizations methods should be improved as time goes on.

2 kinds of categories

a) Statements about statements themselves eg.: ~~..~~ how many letters, how many expressions how many levels, how many different letters.

b) Which theorems are like it? (memory search) e.g. same hypothesis, or same conclusion, or

large parts that are R. same.

c) look for key words ~~like~~ in statements, like "probability".

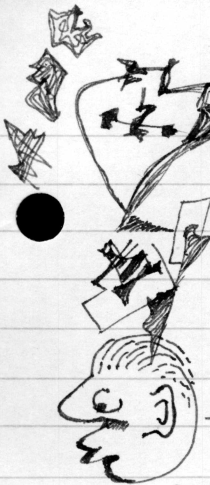
d) To what examples does this statement apply

e) What statements can it prove? (if certain things are added to it).

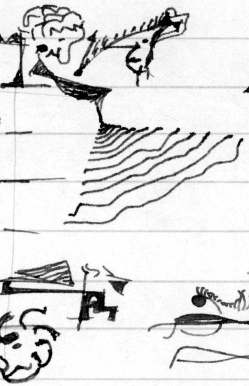
---

4) Method box. M_1, M_2 . . .
Are various proof methods.
A method xfms a statement or set of statements into a new " " " " ".

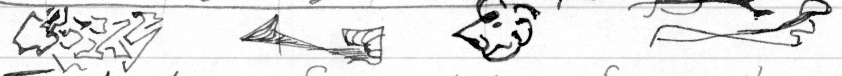
5) Success memory: Decides when a method has been useful in proving a theorem. We then score various characters of theorem assoc. ~~..~~ with that Method.



	C_1	C_2	C_3	C_4
M_1	3	9	8	0
M_2	0	7	0	8
M_3	4	1	1	1
M_4	4	1	0	1
M_5	3	2	2	0



Keep of record of corr. ~~state~~ between characters, methods. We may want some joint correlations.



6) Evaluator of state of system: whether theorem has been proved.

7) A device that tries to find subgoals — i.e. "states of system" Say T_1 is initial state, T_2 is desired final. V tries to find a T_3 such that it is easy to go from T_1 to T_3 and from T_3 to T_2. This is done by looking over various pairs T_1, T_3 in the success memory.

8) V evaluator — Looks at present situation, and desired final situation, ~~and~~ gives ~~an~~ order as to what method to use next.

Milner

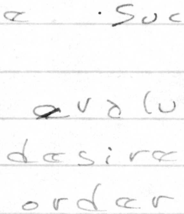
Associationism of 18th, 19th cent. as of Hebb.

How something in brain can become "representative" of a certain type of stymulus.

General interstim., and localized cross inhibition.

Golgi cells give of burst of pulses for single pulse input.

"priming substance" lasts for ~ 1 second.

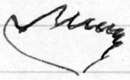


Assumption, that if A fires B many times, that this is facilitory for that connection.

No empirical evidence for this.

Milner is using this

Milner.



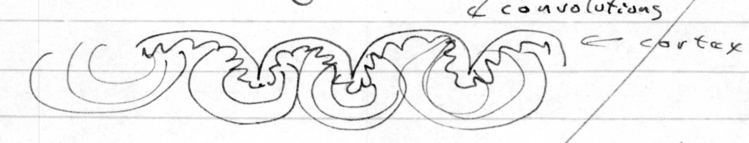
~ 4:30 AM Awake.

lite lunch over ~2:30 AM.

Fire A, then B ~~produces~~
gives A' gives B' in time or space
Now A and B together / several times,
~~an~~ then if A then A' + B' almost — also
" B " " " .

This is called association.

A seeming str. of the brain



cortex is mapped, to some extent, onto various strs. of the brain: e.g. thalamus, cerebellum are known for certain to have this mapping.

McCullock says that there have been many regions ~~found~~ that / partly map topologically onto the ~~brain~~ cortex. (If they map onto the cortex it is ~~imp~~ probable that they do so topologically)

This would be mildly like my sec. reenf. T.M. — with the cortex being the main mechanism, and the rest of the brain being the sec. reenf. computer.

Problem of association ~~as~~ in this simple temporal or spacial form seems not too imp. Marv. suggest same assoc. ~~mechanisms + reenf. would be useful, so working on~~

This is A basic hypothesis of naive Association learning theory: [illegible]

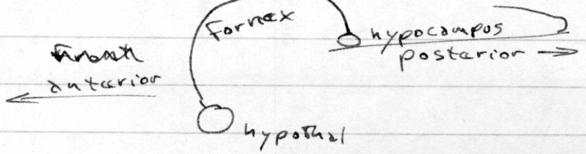
→ these / neural mechanisms of learning mite also be imp.

Reenforcement "centers" in rats: Milner (assoc. with Hebb at McGill)

Olds was early worker in this. Some work (not by Milner) done on cats, monkeys. Sensitive regions:

anterior hypothalamus, fornex, hypocampus. Best on / ant hypothal.

Sensitive region slitely different in different animals.



(Rat acts as tho stimulation (.2 to 2 volts, 60 ~) was a drug to a drug addict.

They will do or learn almost anything to get

Milner

that stimulation. They will, if nec., die from lack of H_2O or food, if they can be allowed to press stim. button instead. They quickly learn mazes if this stim. is reward.

Electrodes are easy to place. One asst. of Milner has ~ 100% success in placing them. Wires are from top, from loose spring or rubber band. — Is small impediment to animal.

In humans, accidental or purposeful cuting out of th. hypocampus, made subject unable to learn new things, tho he retained old learning! This is extremely suggestive, that th. hypocampus is ~~[illegible]~~ along or at one end of any imp. reenforcement pathway. I don't know which way th. pulses travel along th. hypothal - fornex - hypocampus circuit.

→ Extremely eager maze behavior. Some tendency not to learn about dead ends in maze, and repeated making of same mistakes. — much worse than normal behavior.

R.F. (wireless) control, with antenna used (not by Milner) for some stimulation work (I don't know if in this partic. type of expt.). Trouble with ~~[illegible]~~ very non uniform distrib. of R.F. field. This would be elim. by using 60~ or hyer frequ. induction field, with an Fe core coil inside th. rat's brain, for a pickup. But Milner says the wires leading to th. animal are no cause of trouble, anyway.

Milner.

Assoc learning, even with reenf. is only a small part of Th. no. of ~~useful~~ imp. epist. tricks that are nec. in prob. solving. However, it is still imp.. A good idea would be to list Th. various epist.

Golgi type 2 are inhib. cells.

Afferent = input neurons

Milner theory: widespread inhib. from a gn. neuron.

Probabilistically, there will be some non-inhibitory loops formed

Selfridge suggests cross-corr. as a funct. of dist. — that for very close electrodes, cross-corr would be small. Th: Paramidal cells have large, recognizeable spikes,

tricks that are nec., and try to impliment them with neurological models. Geometric closeness is easy to model, but substitution of variables in an expression is difficult. It would be

Milner: that his cell assemblies are on for a duration ~~and~~ activity shifts from one to another in jumps.

fine to get a "basis" set of epist. concepts. These can be put in the motor, as well as prediction mode.

One has a ~~motor~~ sensory situation, and certain motor responses have hyer than normal probability of firing because of this situation. ~~This was~~ I want this to corresp. directly to hy aprip. If such motor responses turn out to be ~~reenforced~~ "rite" or "Good", then they are reenforced — which corresponds to an increase in statistical count. Statistically, ~~anoran~~ occurence of only a few or one time, of cases of hy aprip., [illegible] does give signif. conformation, of goodness of prediction method.

If A caused B with reward, then we want A' to cause B' where ${}_AR_B$ and ${}_{A'}R_{B'}$

and R has been found to be a useful relation.

Th. basic Q. in neural nets is not how to get various kinds of "association", but it is "assuming I can get any kind of definable association that I want, how can I get reasonably useful behavior out

4Milner.

of this?" This q. is an epist. one, and is indip. of neural mechanisms. It tells what kinds of "Assocn." mite be almost <u>nec.</u> to get observed behavior.

One simple kind of conditioning learning: temporal sequence:

$A_{sensory} \rightarrow B_{motor} \rightarrow$ reward.

$A'_{sensory}$, $A_{sensory}$ often occure together temporally,

$A'_{sensory} \rightarrow B_{motor}$ then tends to be <u>tried</u>.

(This is, or mite, be call "substitute stimulus").

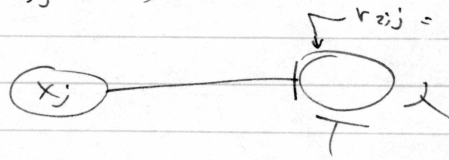
<u>Milner</u> : ~200 or 300 synapses per cell observed

Suppose one has a certain sensory config → motor config. connection. To figure out a way in which another sensory config would tend to give another motor config. ∋ th. second pair were "close" to th. first.

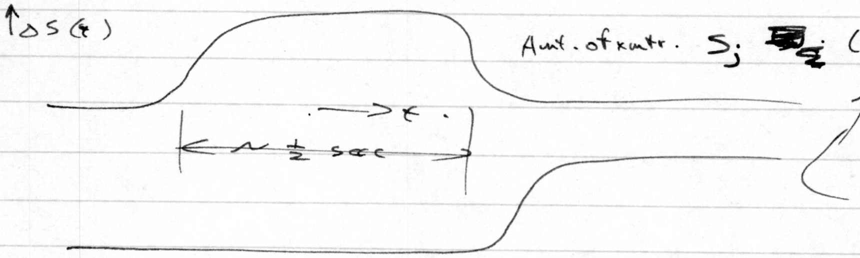
N. Rochester. on how to simulate Milner's nets.

$$\overset{w_j(t)}{\cancel{\Delta(t)}} = k \sum_i x_i(t-1)\, r_{i,j}(t-1)$$

change in amt. of primer (xmtr)



Time of action of cell assembly ~ 10 Δt.



Amt. of xmtr. $s_j(t) = (1 - \frac{1}{m}) s_j(t-1) + w_j(t)$

So $s_j(t)$ is smoothed $\Delta s_j(t)$.

low pass filter of time const. m.

fatigue: $d_j(t) = g_2(x_j(t-1), d_j(t-1))$

x_j = freq. of j at time t.

d_j is increasing if x_j is > certain level

if freq./x_j is above certain level, d_j increases each Δt, by a const. amt.

if below, d_j decreases each time.

Increasing fatigue tends to ↓ frequ., so this tends to some extent to keep x_j at a certain level.

$$x_j(t) = g_3(s_j(t), d_j(t), \text{inhibitor}, x_j(t-1))$$

Inhibition assumed to but <u>unconditionally</u>, and absolute and is special/short range effect.

← this makes Δx_j an imp. contributor to learning. also $x_j(t-1)$

Synapse magnitude $r_{i,j}(t) . = \delta_4 (r_{ij}(t-1), x_i(t), x_j(t), \sum_{i,j} x_{ij})$ Seat of long term learning. If both x_i and x_j are large, r_{ij} tends to ↑. This is periodically renormalized, so that $\sum_{ij} r_{ij}$ is kept ≈ some constant.

A neuron give out its local inhib. if its freq. is > ϵ, but whether ϵ should $= 0$, or be > 0 isn't yet decided.

Hebb wants special mechanisms both short and long term memory.

For selection of "best" motor action, simply ↑ all exciter freqs., untill a motor pattern is triggered.

Make list of sensory stims. that are "close" in ℜ. sense. of having same desired motor response.

Thurs O.K. : Sun. O.K. Sat O.K. [Mon: L. must be at meeting in N.Y.]