My Notes

(0)	J. Bigelow	Aug 15
(1, 2)	Minsky	Aug 13
1, 2, 3, 4	Birnstein on Chess.	Aug 8
1 2 3	McCarthy	Aug 7
1, 2	Samuels	Aug 6
1	T. More.	July 24
1 2 3	Ashby	July 23

Mon July 23
1956

Ashby : Mon
July 23, 1956

a

Linear Sensory Manifolds: saturation levels imp., ∴ knowledge of what basis is, is very imp.. In Sound, this may be very imp.. Actually, this saturation and the basis vectors can be discovered by J.N.D.'s in the vector space.

There is some q. about whether Piaget's expts. on abstraction formation were correctly performed. My impression is that he would often use questions to get responses — like "which pile has more marbles in it?" Rather than choice expts. like with a rat — e.g. Rat learns to chose red square rather than black △. Then how does he chose between red △ and black square?

Center at Geneva. "Int. center for Genetic Epistemology."
Mandelbrot; Piaget.

Storage of sets (or hyper order sets) as operators that tell one whether an object is a member or not. I doubt if one must run into any serious ∞ difficulties.

Problem of homeostat

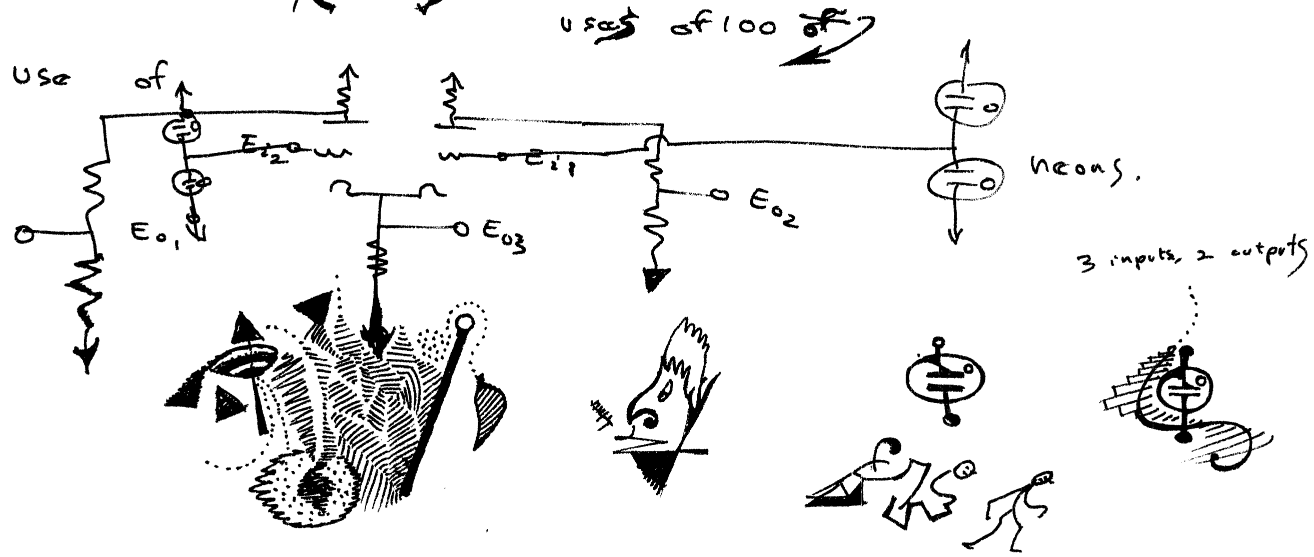
1) No memory of previous solns. i.e. learns to handle A, then learns B → then going back to A takes as much time as before

2) time to go to = lib. is too long.

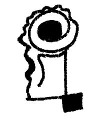
~~Apparently in the Ashby's Homeostat, there is not even localization of error when the system is unstable the next config. is completely random.~~

1) kinds of kicks? not decided
2) kinds of un randomness tried? (row localization)
3) ~~How~~ What criterion ~~controls~~ of stability? not decided.
4) Condensers used to slow down tubes.

Ashby (Cont.)



Apparently Ashby hasn't gotten into the abstraction problem.



W.R. Ashby : Barnwood House
Gloucester
England

Ashby is interested in bringing some rather/imp. simple ideas to th. psychologists et. al, that they have had little contact with.

Ashby is much facinated by th. fact that a table of random nos. can give one all one wants — i.e. that selection is th. only problem.

McCarthy brought out following: that one could predict a time series ~~[illegible]~~ by a systematic search process, in th. foll. way: To each sequence of binary digits, associate a computing machine, in some "simple, natural," way. Try various sequs. until you / Find a machine whose output from t=0 to t=n, are th. first n members of th. time series. Then th. output of th. computer at t=n+1, will be th. desired prediction.

~~Try various se~~ Th. computer would also have to be able to predict other sequences as well, before it could be trusted on this n+1 th case. Essentially, it would have to have th. same "trng. seq." that an intell. machine would have to have.

This idea may, indeed, work.

At any Rate, McCarthy is convinced that th. only real problem is th. search problem — i.e. how to speed up th. search. Ashby feels th. same way, but has fewer ideas

Ashby (cont.)

Ashby wants to start slowly like organic evolution, and work up to more complex phenomena, putting in as few ad-hoc things as possl., and only doing so at each step, when it seems absolutely necessary. His main interest is still in physiology, and humans.

One thing that this study project has made clear: that Minsky, McCarthy, Selfridge ~~[illegible]~~ and myself are all aiming for th. same goal. Ashby is, too, but he would, at present, be satisfied with more modest returns. Simon may be serious about th. ultimate goal — and perhaps Newell. Shannon doesn't yet see anything worth working on that is promising. ~~Mana~~ with Trench More, its hard to say. He will change — he is probably not too hopeful yet, but probably has picked up much from th. spirit of th. group — apparently More has come from rather negative surroundings, and they (and other surroundings) tend to influence him strongly.

July 24, 1956

~~[struck]~~ Trench More

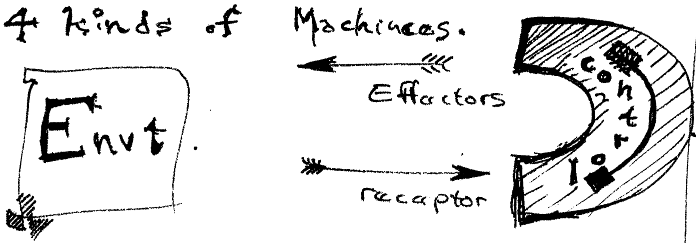
Levels of complexity

Assembly program takes
(Clear and Add) A9 symbolic address

Program which takes macro program and translates it into μ program.

---

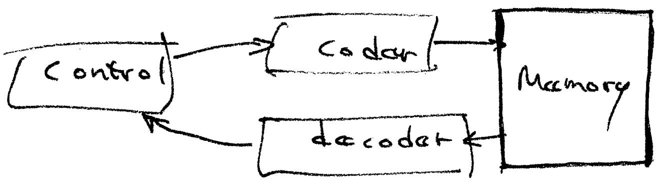
4 kinds of Machines.

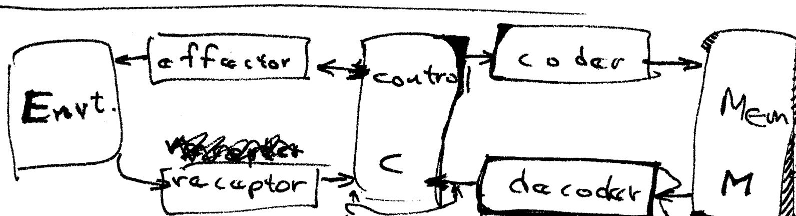
(1) 

---

(2) Envt → Receptor [C control] → coder → Memory

(3) Memory → Decoder [C control] → Effector → Envt.

(4) 



outputs of receptor and decoder are of ~ kinds.

How to divide intelligence between C and M. C may be a sort of "hyer order" type of operation than M

Sign: object signifies mostly something other than itself

---

4 o'clock P.M. ≈ awake. pulse ~ 72
was my beta (≈ 80) betw 2 and 4 somewhere

McCarthy: Problem of "simplicity":

1) That in many cases a bunch of defs., aren't particularly good.
2) That in calculus, ∫ is a good word
3) That perhaps ∃ a def. of simplicity ∋ a very hy % of "simple" expressions are very useful.

Trench plans to list immediate connotations of stimulus (associations) along with it in same memory space.

"G IV wanted to know if Scott was R. author of W."

Here "author of W" and "Scott" are not substitutable; since we would then get G IV wanted to know if Scott was Scott.

Also "Does Pegasus exist?"

In "Nature" ~ April, March, May on odors — British worker

Mon
Aug 6, 1956 Samuels: (IBM). 1

Introduction of intermediate goals.

4 moves minimum. | 4K words storage. (by speed).

At each time only 1 board position for each board position is stored.

Board state = 4 words

possl moves = 5 words.

9 words (= 4+5) for each stage of depth.

We are not memory, but we are time limited.

Move storage method ~ to that of strachey in England.

~~...~~ 16 moves in depth (occasionally).

4 moves ~~depth~~	=	15 sec
~~3~~ 5	—	3 min
6	—	30 min
7	—	5 hrs.

About 10 possl. moves each ~~...~~ position.

Position scoring.

Score = aA + bB + cC + dD + eE + fF

A = piece advantage. (scoring of kings). = ~~k~~ $(P_A - P_B)(40 - P_A - P_B)$ ← gives more wt. to piece advantage ~~in~~ near end of game.

King = 2 pieces.

B = Advancement (toward kings)

C = blocking ~~of~~ opponents pieces from moving.

D = pinning " " " " (They would lose them if they moved).

E = recapture potential for pinning

F = denyal of occupancy of certain squares.

G = center control

M = Move (turns out to be ≡ to 2 others)

These seem, empirically to be imp.

At present a, b, c ... are fixed — are changed 1/10 way thru game.

These don't have to be "indip", but it saves time to have them indip.

About 1/2 of time is spent on position eval, 1/2 on updating

It appears that good eval. method is not imp. since it played rather well once or twice, when th. "piece advantage" evaluator was seriously wrong — almost irrelevant.

Samuels says he is ~~[illegible]~~ entirely interested in the learning problem, not so much in checkers.

He ~~[illegible]~~ runs jump situations out further.

Machine once draw E. Lasker! (tho perhaps in a weak game)

Considers putting old game positions on tape, ~~[illegible]~~ and results of move ~~that~~ was tried — sequencially on tape.

704 can "And" or "or" ~~[illegible]~~ 2 35 bit words.

Tues Aug 7, 1956 McCarthy:

What about use of true ‖ computer for chess and checkers.

704 uses 24 μsec/instruction.

Th. 2 armed bandit problem: We have 2 1 armed bandits — each with its own payoff proby, P_1 ; P_2. We don't know P_1 and P_2, but we are gn. th. aprip $f(P_1, P_2)$. The Q. is how to play. Play depends on no. of trials expected. If we have many trials, we may work them ≈ly for a while, to get info, then work th. best one only.

There is an optimum constant time for deciding how good a position is. ∴ if we play very deep, its time should be negligable compared to th. amt. of time spent updating positions. Th. plausable move program (if such is expedient) is more imp., however, in determining total time required.

"plausable move program" is an operator that has as input, th. board state, and th. pl. moves as outputs. — Contrast with "Evaluation procedure", which is simply a board state functional.

Th. moral is, that even in chess (not only checkers), if one does play, say 10 or 20 moves into th. future, that th. time required for th. final eval. is negligable, compared to th. ~~updating~~ time required for updating. It still, however, is probably worth while to spend much time on final position eval., since it is equivalent to a factor of ~100 (say for 2 moves deep) in ~~extra depth~~ computation time. At any rate, th. time per move, for a very large computer, will be determined by th. time nec. to select plausible moves. At any rate a good eval scheme will just add that much to th. effectiveness of play. Th. point is, that th. time to be spent on end-position eval, th. time spent on determining plausable moves, and n, th. move depth, are all to be selected indipendantly.

Th. time per move = $((n T_p) + T_{Ev.}) K^n$ (No) see next pg

n = depth, T_p = Time spent in determining plausible moves
K = mean no. of alternatives ... final posit.

McCarthy

So indipendently we want to reduce k (?)
reduce T_p and T_{eval}, increase n — in line with ↑ in efficacy
and reducing total time.

T_{eval} has ~~a~~ method of determining its goodness.

n is indep of everything.

Th. problem of how large to make T_p and k isn't clear.

It appears that T_p and $T_{Ev.}$ ~~should~~ may be of th. same order of magnitude.

It would seem a lot easier to invent a pseudo chess with simpler moves. T.M. would play human players that have as much training in the game as it has.

Th. time per move $= (k^n + k^{n-1} + k^{n-2} \ldots\ldots) T_p + k^n T_{Ev.}$

$$= \frac{k^{n+1}}{k-1} \approx k^n$$

$k\Sigma k^n - \Sigma k^n = k^{n+1}$

McCarthy on Artificial intelligence.

1) Program must improve itself
2) How machine can write programs for itself.

Probably McC is content to solve the th. prediction problem by building a min. machine. — which is a problem of th. 2nd kind.

Logical problems and proofs:

Syntactical methods : rules for xfmg chains into new chains

Semantic " : use of examples to suggest theorems or use of counter examples to prove counter theorems.

Bigelow says that for recognizing signals in noise, there is no lower threshold of perception. That for sine in noise (sound) with enuf cases, th. observer will get average > 0.

McC thinks chess imp. because of small ad-hoc instructions that it is good to know how to program (?) This isn't too clear to me. At any rate, McC is still strong for chess.

Samuals doesn't feel chess has much over checkers, but will go along with McC. Samuals doesn't seem too happy about how to get learning out of checkers, either.

Selfridge doesn't think chess is so great, but will go along with McC. Also S. thinks some statistics are nec. for learning — contrary to McC.

Rochester isn't to hot on chess, — ~~would like~~ but will go along with McC if there is nothing better. Nat's pet idea is th. proby. problems in Feller — (which I think is too hard to start on — tho eventually my machine should be able to do it). Nat also seems to like Marv's ideas, but these latter aren't concretely formulated enuf for Nat to start on them yet (perhaps).

Shannon is interested in chess, but isn't too enthused over its relevance to the art. int. problem.

Minsky is very much against chess, also would like to get more active interest in his own approach, Which is Geometry, with figure drawing and a kind of model.

Bigalow Is tolerant of chess, and any other approach, is not so optimistic about how good T.M.'s can be in a few yrs. Is usually rite on most questions, but hasn't really moved to any particular side.

~~Wed~~ Aug 8, 1956 Alex Bernstein on Chess. 1

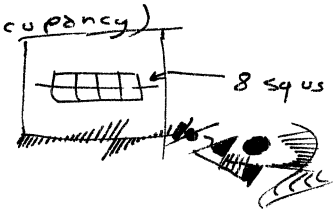
Expects to get ~5 alternatives.

Position eval: - 4 criteria: applied to present position.

1) Mobility. total no. of moves, and captures available.

2) Control of area. (or occupancy)

a) central squs



8 squs are "center"

b) non-central squs.

~~Can~~ squ is controled, contested or occupied.

Each piece or pawn has index telling how many attacks or defending pieces.

3) ~~B~~ Values of pieces

P = 10
~~Kt~~ = 29
B = 30
R = 50
Q = 90

undefended - 10%

undef unattacked - 50%

(SN) These criteria will evaluate initial position — determines whether to attack or defend. Will be used in attenuated for for end position evaluation.

A pawn is weak if it can't be defended by ~~a~~ another pawn

Burnstein probably plays much good chess - at any rate, he is pretty much trying to explicate his and others' intuitive ideas about ~~a~~ what good chess is.

4) Sqs. around the king, should be controled.

General technique from Nimsikovich "My System".

Learning: 4 vector components

a) Uses Samual system ~~to~~ of trying to reweight so that the winning player's moves would have been made.

b) Looks at (losing) opponent's play to see in which way his moves appear to be weak, and then the machine modifies its own strategy to take advantage of this weak point. (e.g. if the opponent isn't as vigilant about control of center as he mite be, machine should take advantage of this)

Barnstein's approach, if successful, will be a very good example of explication of ~~some rather~~ concepts not too close to machine lang.. In itself, this is a good ~~...~~ study problem for art. intell. — it is a "technical skill" that is imp. and useful. Perhaps th. invention of "Fortran" is as good in this respect, as a "study problem".

Unfortunately, that ~~is~~ the ~~hope~~ that programing of various ad-hock conditions, would be imp. seems unlikely

There still may be some pt. to chess, in th. sense of boiling down many ad-hock conditions into fewer "general principals." — However I don't think th. work is proceeding in this direction. In fact th. "engrng work" of the "programmer" is more likely to be close to this, than ~~to~~ what is regarded as th. chess problem proper! In fact Nimsowich seems to have done all of this kind of work that Bernstien's problem will make use of.

This "Grouping of data" in groups that are useful for prediction, would also be directly applicable to th. mathine improving its play thru learning. The 4 ideas upon which ~~...~~ B's parameters are based are of this basic "grouping" type.

The point, here, is not to invent an additional component of th. "4 vector", but devise some principals for inventing new components. Again, th. work seems to be proceeding in a direction not at all related to this.

An acceptable criterion that th. people working on chess, who probably agree on in action, if not in word, would be in the finding of many very good, very easy to evaluate, components for th. evaluation vector. There is no work on methods of ~~...~~ guessing at poss. good components.

McC. seems to be interested in expressing all of th. ad-hock rules, in more unified form. This is o.k. if th. machine itself were to do this condensation. If McC. does this condensation, th. Machine will play very good Chess but we will not have advanced toward soln. of th. A.I. prob

Th. problem of "condensation" of a bunch of ad-hock rules is a tuff prob. and imp. one, but th. point is not to solve any particular such problem, but to know how to solve a general one. That one should chose chess to work with as an example mite be debatable. Th. main point, is that chess isn't being looked upon in this way.

McC says that problem of searching — ~~[illegible]~~ using tests like "mobility" that have hy correlation with ~~[illegible]~~ desirability to look at certain possys — is imp. (as it is). ~~However, it~~ He also admits that it is th. ability of a machine to guess new characters (like "mobility") to ~~[illegible]~~ try — that is imp. in intelligence. However he feels that making th. problem more concrete is imp. — so one can work on it.

upper bound of 50 bits/sec input rate
V.S. hy rate of memorization thru words.

Bigelow and McC 8/20

Upon further proding McC decided that it was th. "speed of search" problem that he was interested in — in chess. This is essentially his very early ~~[illegible]~~ position, near th. beg. of th. summer. ~~[illegible]~~ He says that even in addition to good "characters" that describe one's situation in th. search problem, something more is needed in th. analysis. I expressed th. opinion that in addition to getting good "characters", th. rest was a trivial problem in statistics. McC. denyed that statistics have anything to do with it. Another key to his orientation, is in his statement that Wald's sequencial analysis only ↓ ~~[illegible]~~ inspection time by a factor ~ 3, at best.

If this arguement tends to tie McC down to th. "search" aspects of chess, then maybe this isn't too bad, and th. arguement was very well worth having.

It is <u>very</u> questionable whether chess is a partic. good example of a "search" problem. It may be easier to get financial support for ~~[illegible]~~ th. search problem <u>in this form</u>.

One of these days, I shall have to look into th. "search problem." There may, indeed, be more to it than good "character" generation.

As an aside: J. Bigelow said that th. chess problem was essentially trivial and that he would get an existing computer to play better than a human, if someone would give him $20k for th. work. I think that he said, or implied, that its relevance to th. art. intell. prob. was rather small. Bigelow sounds like he mite be really <u>very</u> good at expressing intuitive ideas in explicite form. If so, his claim about his ability to "do" chess is probably essentially correct.

Apparently th. "search" problem's importance was almost forgotten by McC. in his approach to chess and "not seeing forest because of trees." He must, apparently, be kept pointed toward problems that are relevant to art. int., since he can become easilly distracted.

Mon Aug 13, 1956

Minsky's approach to A. I. problem: (1

Make machine to prove theorems in ordinary plane geometry. Statements of theorem are in modified English. Th. machine makes up proofs by drawing diagrams, to suggest various sub-goal theorems.

Th. truth of a sub-goal theorem (before it is proved) can be found to a hy ° of probable accuracy, by looking at th. diagrams: e.g. if a sub-theorem says ∠A = ∠B, we can "see" if they are = on th. diagram. Actually, no diagram need be drawn. Th. machine can take an arbitrary ("random") set of coörds to draw th. diagram, and compute ∠s and intersection points by analytic geometry.

Since this may be done with an arb. ° of accuracy, the probability of a "theorem" being "apparently" true, when it is not really true in the general case, can be made vanishingly small.

"Heuristic devices" will be incorporated in the machine ab initio and we will see just how far the machine can get from Euclid.

An example of an "heuristic device" — "When one has to prove (the sum of) 2 ∠s are = to a third, draw the 1st or 2nd within the third so as to subtract from it." Also, after construction lines are drawn, the resultant figure can be evaluated for "goodness" in the sense of probable relevancy to th. soln.

Basically, I think his primary idea is to study heuristic devices. — Eventually, to make generalizations about them, so that one can get <u>methods</u> of inventing new ones.

There is probably th. hope that th. heuristic devices for inventing heuristic devices, will be ~ to th. heuristic devices themselves. This may be true if the th. primary problems worked on are sufficiently ~ to th. problem of inventing heuristic devices.

At any rate, an understanding of heuristic devices, is what M. expects to get from this study.

Th. main Q., is whether this is a partic. good, ~~way~~ ~~the~~ direct way to make such a study: i.e., would chess, ~~be~~ or checkers, or th. Simon-Newell ~~machine~~ ~~be better?~~ or th. More ~~[illegible]~~ machine, better?

Aug 15, 1956

Julian Bigelow: He is extremely wary of <speaking vaguely, with th. hope of being able to translate this vague talk into machine language>. (as is also McCarthy)

On th. other hand Minsky, Rochester and perhaps More, are less wary of this, and M, in particular, is not afraid to describe things at a fairly intuitive level, then become more and more concrete.

J. B. would rather work by making a small specific machine, than a bigger, more complex one, etc.

I think that this can be interpreted in both J.B. and J.Mc. as an inability to know what is translatable into machine talk, and what isn't. J.B. ~~expressed his~~ said that occasionally he would write papers that would turn out to be meaningless, on further thot.

If these analyses are much to th. point, there seems to be little likelyhood of J.B. or J.Mc. ever doing anything worth while. Mc's concern with programing, and impatience with "Substanceless Burble" seems to bear out above concept.

J.B. wants to work on proj. geom. to start. Wants to feed as many intuitive ideas on properties of lines (s), etc., into computer, as ~~possib.~~ he thinks are nec. for th. task. This will be done by metric properties ~~of~~ th. objects, somehow. Th. machine will be able, as a basic operation, to tell how close 2 pts. are. In speaking to him about this I wasn't able to get a very clear idea as to what he had in mind. Perhaps it isn't clear enuf in his own mind.

He mentioned his not being able to see any advantage of my approach to T.M. v.s. his. that he should find them comparable is interesting — tho. I don't understand his ~~math~~ idea well enuf to say what this means

This programing in order to learn how to program most anything, is what Oettinger did in England, and is ruffly what he is doing here on M.T.

f.c

ABC DEF GHI JHL MNO
1 2 3 4 5 6

PRSTUV WXY
7 8 9

TR · 68274

Lebedev : ← Computers and Automation ≈ 2 mo. ago.

Thurs Morn : 1 hr.

1) Division into a) Search
* b) Induction

2) Machine efficiency ≈ no. of trials × sample size in past

3) Problems : Prop. calculus (These are "Search problems"). i.e. inversion problems
- Plane Geometry
- Arithmetic
- Combinatorics
- Integration : $\int_a^b$
- Design of sw. ckts.
- Spacial construction (Blocks, etc.)
- Computer programming.

Partially well defined kind of problem where we know A > B.