exposition.

Give definitions: ← note that this is ~~not~~ meant as an explication of an intuitive term

1) ~~Abstraction~~ abs: set of objects in spacetime that are

absss: | to be considered "the same"

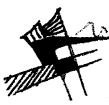
2) n tuple: ordered set of objects

ntp, ntpst.

3) structure: ordered set of relative coords.

str. strs, strst, strsts

Give several examples of each. Show by examples, that an abs may be a large # of objects in a config. - that an ntuple, may have, for its elements, any object, no matter how complex.



In 1), "Abstraction" is not a good word to use.

~~perhaps~~ n-gram would be O.K.

Note that these dots are temporary

1 example learning.
— how a single example may suggest soln. to search process which gives new gp. that solves prob.

2 scient. problems.

For talk: Make up examples

Heuristic Approach: from th. beginning: Start with simplest induction problems from everyday life:

*1) E.G. A bus has ~~come~~ here every day at 8 a.m. Give proble[m] How language helps Show just what gps. are formed. / Give example of some legitimate but useless gps. / i.e. when tested, they prove useless. Perhaps discuss ad-hoc gps. Bring out pt. that problem in Science is to form ~~[illegible]~~ gps that have ~~[illegible]~~ reasonable proby. of being good — that usually testing their goodness is routine. Show how this is true in Math as well as physics, and soc. sci.

That this is main prob. of most science

*2) ~~More comp~~ Induction in everyday life of more complexity, involving com~~plex~~ abs. — linear eqns. — E.G. Man with bk. hat in T.V. is "bad": That ~~certain~~ you don't like certain foods. Psychoan. abs. resulting in neurosis. e.g. Birds.

Gen. problem: that gn. a device (adding machine) T.M. will learn to predict output (Say: I will give

*3) Simple arith. learning on intuitive examples intuitively, ~~then~~ Then I will go into more detail on how machine ~~[illegible]~~ mite implement these abss.

Explain closeness idea: [2 abss. are "close" if by U of one ~~is~~ often → by U of other] ~~B~~

2 kinds of closeness — on 2 levels
#) 2 objects close if in same gp.
2 gps are close if by U of 1 implies by U of other.

Some kinds of closeness: a) physical closeness (e.g. continuous space problems, and ~~Euclid~~ geom. metrics how extrap. to some ~~[illegible]~~ discrete cases)

b) isomorphism and how an isomorphism defines idea of a str. and a ntpst. (Th. arg. of th. str.)

c) Some simple kinds of closeness of strs.

d) " " " " " ntpsts.

Give examples from everyday life and from Math.

e.g. indiv. members are close,

e) How to generate new strs. and ntpsts and ngms from old — What this means in intuitive problems.

How this T.M. program mite be used to teach children

How T.M. operation corresps. to devising of word in lang. i.e. a) testing of U. of idea. b) Inventing word c) That ~~concepts~~ hypotheses using few words ~~[illegible]~~ (simplicity) have hyest aprip.

Start with 1 dim. ngm. and str. : extend to 2 dims.

≡ permutation, repetition, ommission.

Tues
Topics for talk: July 10, 1956, Dartmouth

1) 2 kinds of sc. problems:
a) If α happends, what will follow? (Pure sc.)
b) If I want β to happend, what must I do to bring it about (Engrng - ≡ inverse of (a))

2) a) Example of Bus at 9 A.M. Show how one groups phenomena for prediction.
b) Example of Linearity of equ.
c) Example of Isomorphism between Electric ~~Filters, mechanical~~ acoustic filters.

3) Some examples of poor gps.
a) Bus without 9 A.M. in description.
b)

4) That most imp. work in sc. is devising categories – that correl. of cause and effect is routine and need not be done very well. Poor statistics will slow down predictions, and decrease accuracy somewhat, but poor gps. will make prediction very poor.

5) Some methods of forming gps. directly.
a) Same object except for time, or space displacement.
b) Same except for small continuous distortion.
c) " " " small ommission, or addition.
d) Gave rise to "same" effects [chemicals, operational similarity]
e) Have similar cause [not so common]
f) Have " structure (isomorphism) electrical → acoustic analogs

6) Methods of combining gps. to form new gps. of hy ~~[illegible]~~
a) Boolean mult: e.g. red apples
b) Boolean addition (not too useful)
c) Combining structures and ordered sets of elements.
d) concatenation.

7) General Scientific scheme is Occam's razor. I.E. The simplest hypothesis is most likely apriori. That this means simplicity in terms of existing lang.
a) Take old words, and concepts in lang., combine to form hypotheses, or new tentative words and concepts
b) Test hypothesis, concepts, words for usefulness in prediction.
c) Give useful concepts ~~[illegible]~~ names, so that they have hy combination proby. in next time around.

Topics for talk.

8) Scheme to be used in prediction:
   a) start with categories.
   b) test categories for U in prediction.
   c) Combine " of hy U to find new **trial** categories.
   d) test new cats. for hy U in prediction.
   e) etc.

9) An example: Printed English.
   a) Use of ngrams in prediction.
   * b) Formation of longer ngrams by combining smaller ones
   i.e. concatenation of "words". [imp. method of cat. formation]
   c) sentences: Th. Question, Answer predictor.

How old is J.? : J is 10 yrs. old.
" " " B? : B " very "
" " " A? : A " 7 months " .

Gen. form : How Old is α?: α is β old.
Idea of structure as permutation, repetition, ommission.

S_1(How, old, is, α, β) = How old is α; α is β old.
S_1(α, ♂, 7, ∀, ∈) = α ♂ 7 ∀ ∀ 7 ∈ ♂
S_1(1, 2, 3, 4, 5) = 1 2 3 4 4 3 5 2
↑ structure. ↑ n tuple

Ntuple set. (How, old, is, α_i, β_i)

some α_i, β_i pairs:
J, 10 yrs
B, very
A, 7 mo.

c) We now have (Str.), (Ntpsts)
we can make new cats. by ~~devising new strs~~
① devising new strs. from old
② devising new ntpsts. from old.

10) Another example in 2 dims:
Is Rather general prediction problem: Can be thot of as device that watches a black box's input, output, then can, itself, reproduce same output for same input.

a) Generalized Arithmetic.

= 1011 / 1[0]11 , = 01[0]0 / 0100 , = [1]001 / 1001 , = 1101 / 110[?]

① Use of monograms of low U in prediction.
② " " digms ☐ ☐ , ☐ , ☐ is very useful.

(b) Put diagram of flow Chart on Board

topics for talk

10) b) ~1011 , ~01⓪0 , ~①001 , ~1101
0[1]00 , 1011 , 0110 , 001[3]

① ~~By digrams~~ No digrams work for both = and ~
② ~~Maxim~~ Some trigrams work part of th. time.
~~i.e.~~ e.g. = 1/□ , ~ 1/□ , = □/0 , ~ □/1 , etc.

③ Tetragrams work all th. time.
e.g. 1 0 / 1 □ , 0 □ / ● 0 , 1 1 / 0 □ , 0 0 / 0 □ etc.

④ ~~Have to use strs.~~
Intuitive feeling that ~~e.g.~~ e.g. = □ □ 1/□ should be imp.

⑤ Method of getting this using strs.
ⓐ Define strs. as 7 dim. perm, concatenation, repetition, omission.
Notations : [0, 0; 0, 1] · (α, β) ≡ [α | β]
[0, 0; 1, 0] · (α, β) ≡ [β / α]

[1 | 2] · (α, β) ≡ [α | β]
[2 / 1] · (α, β) ≡ [β / α]
[2 | 1] · (α, β) ≡ [β | α]

ⓑ start with monograms, useful digms. i.e.
cart. product of □/0 , □/1 , 0/□ , 1/□ , = , ~ , 0 , 1 with itself.
ⓒ Mult ↑ by s_1 ≡ [1|2] : we get 128 ngms (ambig)
only 32 occure, only 24 contain inter. sqw. ~~Th~~
~~this~~ of these 24, 8 are useful. i.e. 1 in 3,
which is very good.
see 18.12 for these useful trigrams. ←
~~these~~ these trigrams = $S_1 \cdot N_1$ where N_1 ≡
we may modify S_1 and try slite variations:
see P 22 for examples.
Th. last 3 (S_6, S_7, ~~&~~ S_8) are of hy U ≡
we get 8 trigrams of hy U (see bottom of P22)

---

similarly we can get = □ □ 1 , etc.

Topics for Talk.

11) Present progress:
a) up to normal binary addition.
b) expect signif. modification of methods in future
~~[illegible]~~

(12) Future
a) will try to stick to intuitive xfmns.
b) ~~[illegible]~~ Much Math, and then teach simple lang. — say Q - A game.
c) Originality will appear soon in ~~[illegible]~~ Math operations e.g. literal soln. of alg. equs: numerical solns.; ~~[illegible]~~ numerical interp., extrap. of functions (say sine x to several cycles!) differentiation; Integration, etc. [see PX730 for list]

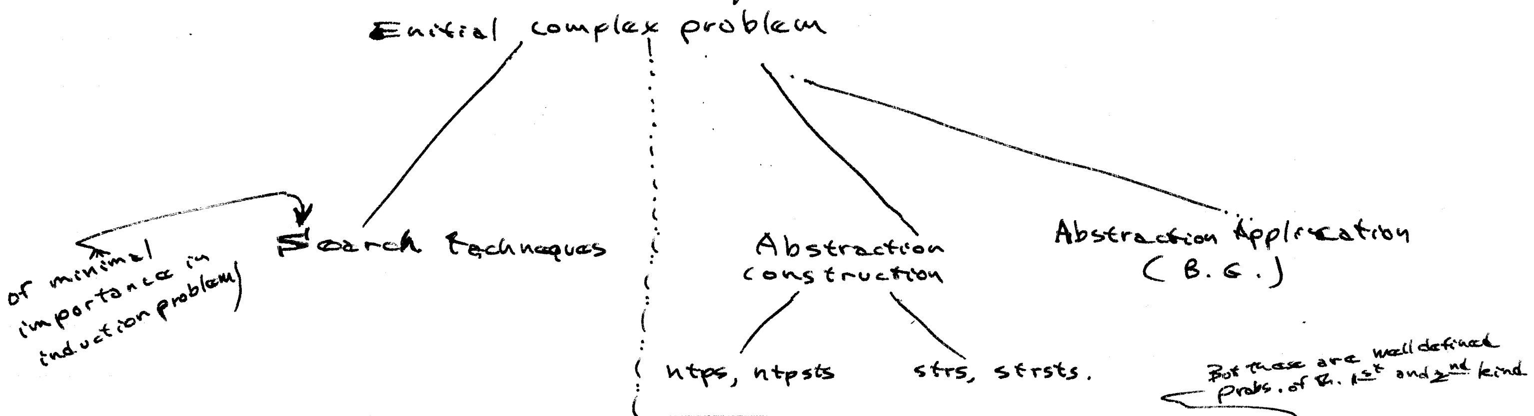
13) ~~[illegible]~~ Criterion of real learning:
When one can feed T.M. learning seq. ~ to that for human, and have it learn ~~and when~~ providing one hasn't inserted any new ad-hock devices for a long time.
This may be rather soon.
I.E. Draw General expected history of T.M.
— Many ~~ad hock device~~ new devices at beginning of Tng. seq. — then they will become less and less frequent — and perhaps stop — and still T.M. will be able to learn new things

SN

[SN]'s

1. A much more Heuristic approach is needed. Th. presentation should make th. problem to be solved clear at each point, and it should be clear that what I am proposing is an attempt to solve it. Perhaps more examples of what I am trying to solve — at each point — would be helpful.

Breakdown of T.M. type problems:



Abs. construction may involve some searches occasionally; these can be done by 1) A Black box 2) An actual exhaustive search.

For induction problem: The followed points should be made clear in this order: ① That th. main problem is abs. construction of hypotheses, and that abss. ↔ sets of objects. (Give examples — in particular, how ngms → sets)

② Some simple ways of constructing sets ⓐ compact ngms.

③ Factoring old, useful sets into sub-elements (i.e. strs, ntps), recombining sub-elements into new trial sets, or new sub elements.

④ Testing the new trial sets, new sub-elements

⑤ Repetition of 3, then 4 then 5 (Note this program loop)

⑥ Dev. of standard methods of use of Abss (B.G.)

A more detailed order of presentation:

0) Th. idea of frequency counting of an element →

1) Th. idea of an ngram — give examples from a) Printed English, b) th. arith. problem with digrams, c) how this is an example of th. gen. epist. post. that the causes of an event tend to be close to it in time and space. → as an example, perhaps of th. epist. law: "Things tend to remain th. way they were" (Not too good an example — one is just assuming th. statistics remain th. same — which one does for any other abs. anyway.

Un ordered Remarks

methods of generating gps.

How Grouping methods/ are a deterministic system, ~ to an axiom system.
As such, they all have limitations.
That ~~certain~~ any gping method will be restricted to a certain universe of phenomena that it can work with. (=predict).
That we will concern ourselves with an attempt to obtain a small set of methods that are sufficient to cover the phenomena of the universes we are familiar with.

[SN] on use of blank☐ for examples is poor: will not solve A:B as R:D — i.e. pick another ntp. from a gf. for completion.

On "Closeness": 2 kinds of closeness.
1) On low, phenomological level. 2 phenomena are ~ if it is useful to gp. them for prediction purposes. So we have ~ ity with resp. to various "properties"
2) 2 gps are "close" if hy U of one suggests hy U of other. Here we have only one property of interest — i.e. hy U — we may find others. i.e. feasibility in combining them in certain ways.
3) It is probable that we will need hyer order methods of grouping.

---

i.e. we have sets of phenomena, sets of sets of phenomena, and sets of sets of sets of phenomena ... etc. This can be dealt with in a machine that mixes levels of abstraction — as, I think, humans often do. Give examples of sets, sets of sets

---

One can simple treat th. "sets" problem as any other T.M. induction problem — or perhaps th. "set of sets" prob. in this way.

Th. concept of "Number." This can be done in a simple, Russle way: I.E. th. set of all triples bears th. betweenness relation to th. set of all pairs, and th. set of all quadruples.

insert ③ [redacted]

f. A simple type of [redacted] mathematical learning.
The organism is presented

Partial Ordering of operations: enitial — ngms from enitial strs gn.

1) [redacted] digit Freq. count
2) digm. freq. counts
3) trigram " counts
4) Tetragram " "
⋮
ngm. " "

1) a) Creation of ntpsts by cart. products of ngms of hy U [redacted]
[redacted] b) enitial strs.
2) Mult. of strs. by ntpsts to form ngmsts.
3) Use of neighboring [redacted] strs.

1) combining hyu strs. to obtain new hy oprip strs.
a) by trying "neigborihgs" strs
b) multiplying strs and strsts.
times strs and strsts to obtain new strs
e.g. $\boxed{1\,|\,2} \times \left(\boxed{\frac{2}{1}}, \boxed{1\,|\,2}\right) = \boxed{\frac{2}{1}}\boxed{3\,|\,4}$ or $\boxed{\frac{2}{1}}\boxed{3\,|\,4}$

2) combining ngms, ntpsts to obtain new ntpsts, ngms.
a) cart. product.
b) function c) others.

Use of testing. as black box, in which one puts gps. and their params., and output is U and other params.

Another xform to create new ntpsts from old:
the "function" idea. It is a comb. of 2 relations!
i.e. if (a_i, b_i) and (c_j, d_j) are 2 ntpsts,
then they combine to form
the ntpst $(a_i, b_i, d_j) = (a_i, c_j, d_j)$ ~~where~~
for all i,j where $b_i = c_j$.

There is the generalized idea of inversion, which implies a search. It isn't clear that it is trivial.

Search Problems: imp.

There ~~abstraction pro~~ are 2 problems in search
1) ~~what~~ Gn. a present situation [present states plus final goal] what should be next sub-goal?

2) Given a present state and a sub-goal, what xfmn should be used next?

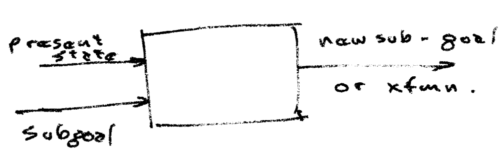
---

1) The present state, and th. final goal, are described interms of characters (abstractions). The next sub-goal will be described in terms of characters. What we want, is an operator on th. present state and th. final goal, that gives th. next sub-goal. This operator will be a function of the characters of th. present state and th. final goal.

2) Here we assume that we want no more sub-goals betw ~~...~~ our present state and th. sub-goal we have just chosen.

[SN] We divide our operations into a) choice of sub-goal, b) attempt to reach sub-goal c) decision after each few steps whether to continue toward same sub-goal, or devise new sub-goal. One may routinely apply th. procedure of c) to all steps — usually, th. same sub-goal will be indicated at each pt. in th. process (?).

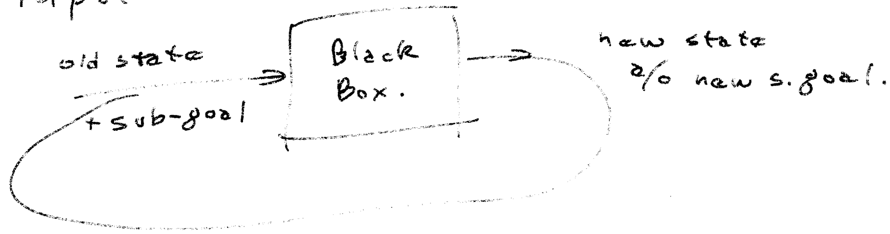
Elementalization: From init. state and final goal, to devise sub-goals. To then try to achieve sub-goals.
At each point, one puts present state and sub-goal into box: at output is either next xfmn on present state, or a closer sub-goal.
How th. box works: Each present state and sub-goal is ~~...~~ gn. characters. The output of th. box is a) a sub-goal or b) a xfmn.



At any rate, th. output of th. box is an xfmn to be applied to th. ~~...~~ present state + subgoal.

Search problems.

Sometimes th. present state is changed, sometimes th. goal is changed. Th. output then is fed-back to th. input



Black box contains 1) characterizer of inputs
2) Set of xfmns that may be applied to input.
3) ~~Search~~ Device for keeping score — i.e. which xfmns have most success with which input characters.

~~Device may just~~

4) When device reaches a sub-goal, its next sub-goal becomes th. one that was selected for it before.

5) The device may simply keep on ~~...~~ finding new xfmns and sub-goals. To put a stop to this, each sub-goal should have a time limit. When this time limit is exceeded, ~~..~~ an alternate sub-goal or xfmn is tried.

6) Minsk also suggests a time limit for th. pursuit of each goal or sub-goal — then going on to th. next, most probable goal or subgoal. ——— i.e. this could be analysed in some detail.

1) Least □ is bad.
2) str. system is bad — ~~bots~~ parenthesis may be O.K.

x, x,, x,

□ + □ = 2□

B + (B) = 2B

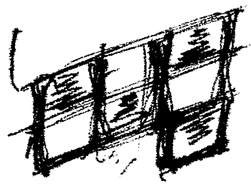
x, x, x

p-case no. (= old cases)

case no. (= old cond)

case no. of ① α β = that of ② α β = that of ② α β, ≠ case no.

(x, 2

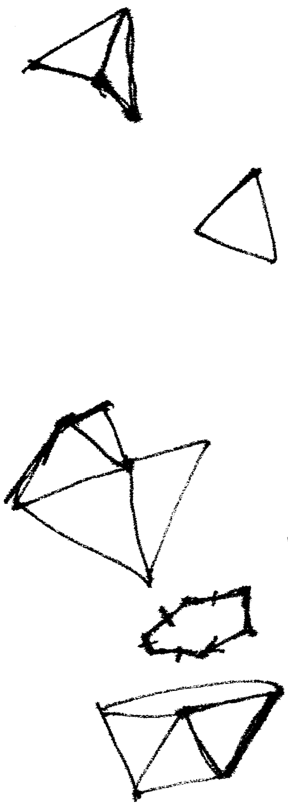


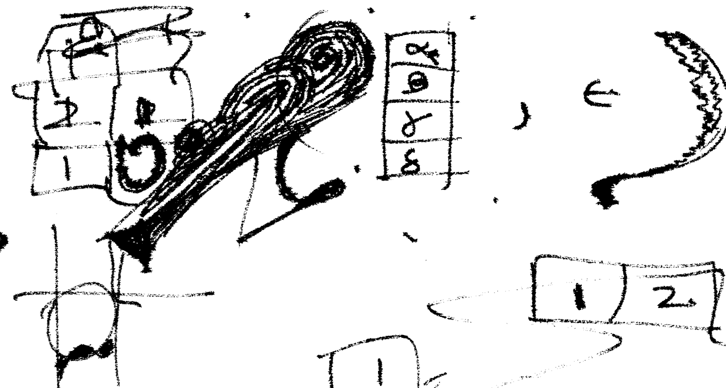
| 1 | 2 | ... | | | ... | (x, , B + B = 2B)

(A+B)

| ① | + | ① | = | 2 | ① |

| 1 | + | 1 | = | 2 | 1 |





α, β, γ, δ , ∈

→ strs are ngms — also

We may have mixed strs and ngms.

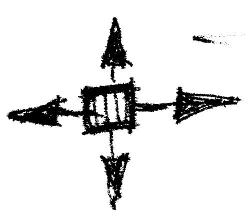
| 1 | 2 | a "pure" str;

① ② ⇒ ①

a "mixed" str, would be

① α ② β = ② α

this notation mixes methods space and problem space: what is good.



What about shrinking ~~of~~ components of an ngp. to make them fit the spaces of a str.?

Don't forget !!

Fa board at Berger's

Blanket + cover to Minsk.

~~2~~ pants, shirt in basement.

Laundry to Minsk.

T. Brush.

Stuff at Math dept.

Books on Math Shelf.

Send 2nd set of change of add. cards with All 4 X'ed.