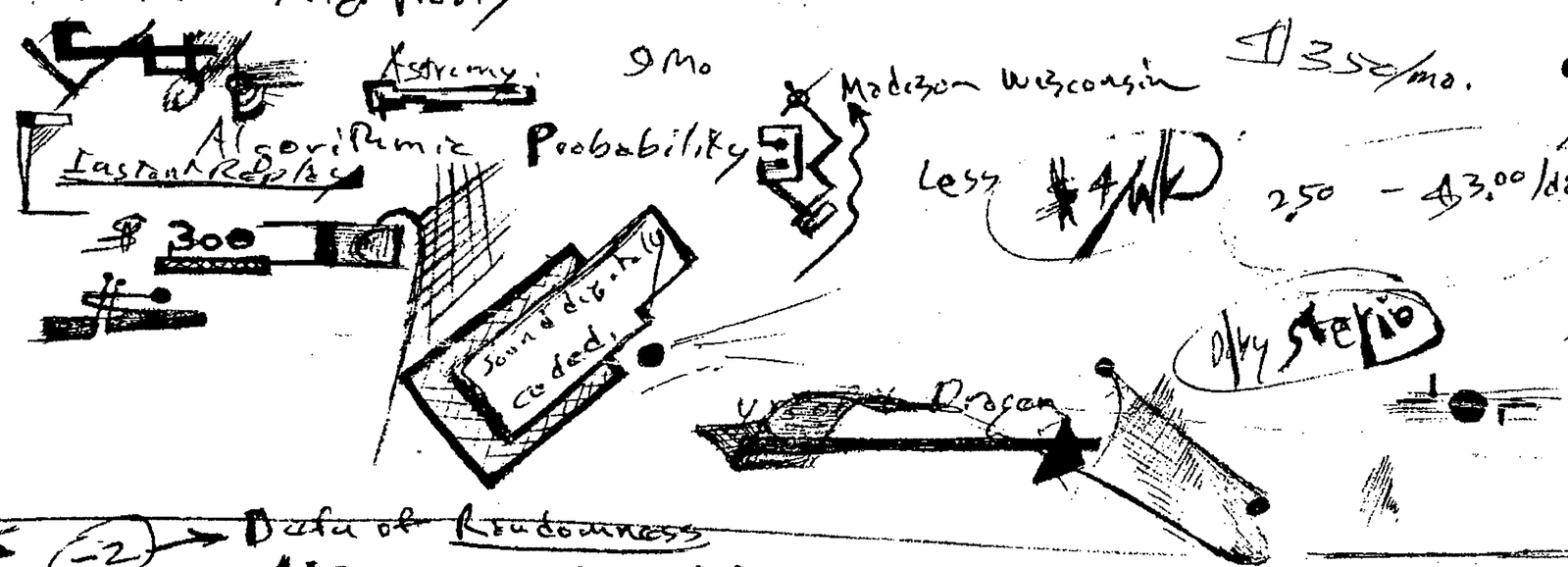
92385 TM MAIN Pts. of Alg. Probty



rental of video cassettes?

Sach Somerville

what was this about? phen. conv. w. Mine Bate?

(2) 1

(-2) → Data of Randomness

Algorithmic Probability.

(-1) How it can be defined in terms of a Univ. Machine or Univ. algm.

What are some of its important characteristics)

(Later: Kolmog's ~1930 "Euclidean" of probty! — Enabld proby dist → proby dist. — But not RW data → proby dist.

→ 0) PM goes from set of data to Pd.

1) Unlike the frequency concept of probability, it enables us to get reasonable probability estimates for very small samples, and even gives ~~[illegible]~~ ~~few as 1 or cases.~~ ~~It also gives~~ estimates for the probability of events that have never occured befor.

2) It has the Completeness property: ABC ~~[illegible]~~ from ~~reasonable~~ sample of the output of a stochastic source, it will be able to discover the ~~[illegible]~~ true probabilities associated with the generator, ~~with certain~~ It will ~~be able~~ do this with a relatively small sample.

3) ~~Associated with~~ Any complete probability evaluation method must be incomputable. Conversely any computable probability evaluation method must be incomplete. ~~[illegible]~~ ~~[illegible]~~ It is inherently impossible to put a useful upper bound on the error of any incomplete probability evaluation method (or any complete probty method) In using computable approximations to PM, the size of the error is of necessity unknown.

Here we can use Alg. complexity theory to give a system that gives us best proby values possible in a given amt. of computing time.

4) An advantage of PM over other probability evaluation methods, is that in approximating PM, we always know ~~[illegible]~~ what computations to do in order to attempt to improve the

25.00
zenith 2FL-8892
1880
complete

accuracy of our probability estimate — tho we never can know how much (if at all) our " " has improved.

5) PM is a ~~Bayes~~ completely Bayesian Method. The apriori probability on all possible objects, is obtained by considering a Universal Turing Machine (or any other Universal process), that can produce all possible objects to be considered, as its possible outputs.

If the machine has random input, each possible output will have some probability of being generated by the machine. This probability is that wich we use for the Bayesian apriori probability.

The unavailability of this apriori probdistribution was (up to now) the only thing preventing Bayesian Statistics from being a complete solution to all problems in statistics. Algorithmic Probability gives us the needed distribution

ALP gives a reasonable apriori distribution based on what we have found useful in the past

6) The ~~Bayes probs~~ conditional probabilities obtained via PM are relatively insensitive to the choice of the universal ~~reference TM~~ Turing Machine used for reference

This sensitivity is further reduced if ~~the~~ a large amount of data used for the probability estimate.

7) One objection to the use of Alg. probty is ~~that~~ not that it is incomputable, but that it takes too long to compute — i.e., its Computational complexity is too high.

However, by use of suitable training sequences a/o the construction of concept nets, the computational complexity can be reduced to no more than any other probability evaluation method of equivalent "accuracy"

8) The study of concept nets and training sequences: This enables us to construct very intelligent machines — but it also enables us to understand the learning process in humans and find ways to teach them most effectively.

N1385 The MAIN PTS of Alg Probty (cont from 92385)

(9) That just about all problems can be expressed as INV (Inversion) or OZ (Timelimited optimization) problems. In particular — that the problem of inductive inference is an OZ problem. That Levin's search method gives a near minimum solution time for both kinds of problems, if all of one's knowledge about the problem is included in a suitable conditional probability distribution relating the problem to possible solutions — or alternatively all one's knowledge is in the form of a probability distribution that tells the probabilities for various choices in the maze of choices that one has in searching for a solution.

(10) The concept of conceptual jump size: Given a certain specified knowledge in a machine — both heuristic and problem-specific knowledge — this tells how much time it would take for that machine to solve a particular problem in a particular way — either an INV or an OZ problem. Makes it possible to design the sequence of concept nets we actually running any problems

(11) While the size of error in probability estimate can never be known with any useful accuracy, Alg. probty gives us a way to spend our time (3) in an optimum way so that at any time our probability estimate is about the best possible. A 2nd probability estimate is about the best possible. It's been sort of added, but not very well.

(12) That Alg probty gives us a v.g. understanding of Randomness.