What is...machine learning in mathematics - part 4?

Or: AI and automated proofs

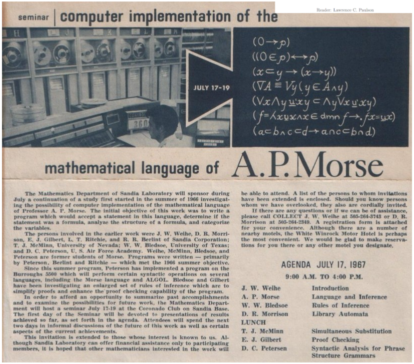
No humans, please!

Taken from:

HISTORY OF INTERACTIVE THEOREM PROVING

John Harrison, Josef Urban and Freek Wiedijk

Reader: Lawrence C. Paulson



seminar | computer implementation of the mathematical language of A.P.Morse

JULY 17-19

The Mathematics Department of Sandia Laboratory will sponsor during July a continuation of a study first started in the summer of 1966 investigating the possibility of computer implementation of the mathematical language of Professor A. P. Morse. The initial objective of this work was to write a program which would accept a statement in this language, determine if the statement was a formula, analyze the structure of a formula, and categorize the variables.

The persons involved in the earlier work were J. W. Weihe, D. R. Morrison, E. J. Gilbert, L. T. Ritchie, and R. R. Berlint of Sandia Corporation; T. J. McMinn, University of Nevada; W. W. Bledsoe, University of Texas; and D. C. Peterson, U. S. Air Force Academy. Weihe, McMinn, Bledsoe, and Peterson are former students of Morse. Programs were written — primarily by Peterson, Berlint and Ritchie — which met the 1966 summer objective.

Since this summer program, Peterson has implemented a program on the Burroughs 5500 which will perform certain syntactic operations on several languages, including the Morse language and ALGOL. Bledsoe and Gilbert have been investigating an enlarged set of rules of inference which are to simplify proofs and enhance the proof checking capability of the program.

In order to afford an opportunity to summarize past accomplishments and to examine the possibilities for future work, the Mathematics Department will host a seminar July 17-19 at the Coronado Club on Sandia Base. The first day of the Seminar will be devoted to presentations of results achieved so far, as set forth in the agenda. Attendees will spend the next two days in informal discussions of the future of this work as well as certain aspects of the current achievements.

This invitation is extended to those whose interest is known to us. Although Sandia Laboratory can offer financial assistance only to participating members, it is hoped that other mathematicians interested in the work will be able to attend. A list of the persons to whom invitations have been extended is enclosed. Should you know persons whom we have overlooked, they also are cordially invited.

If there are any questions or if we can be of assistance, please call COLLECT J. W. Weihe at 505-264-3743 or D. R. Morrison at 505-264-2349. A registration form is attached for your convenience. Although there are a number of nearby motels, the White Winrock Motor Hotel is perhaps the most convenient. We would be glad to make reservations for you there or any other motel you designate.

AGENDA JULY 17, 1967

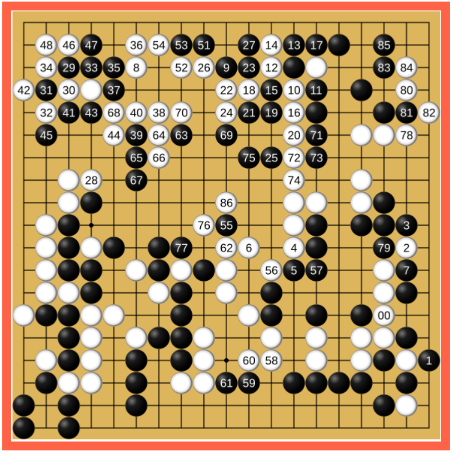
9:00 A.M. TO 4:00 P.M.

J. W. Weihe	Introduction
A. P. Morse	Language and Inference
W. W. Bledsoe	Rules of Inference
D. R. Morrison	Library Automata
LUNCH	
T. J. McMinn	Simultaneous Substitution
E. J. Gilbert	Proof Checking
D. C. Petersen	Syntactic Analysis for Phrase Structure Grammars

Figure 1: Proof-checking project for Morse's 'Set Theory'

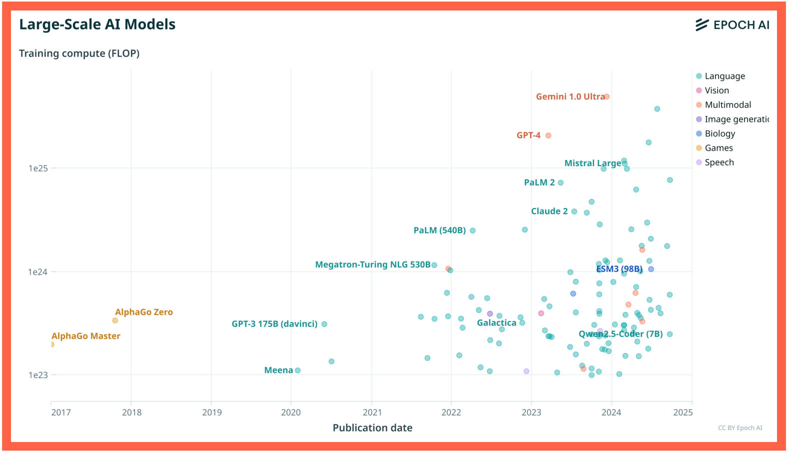
- **Assume** for this video that "mathematics = proving theorems"
- **Task** Automated theorem proving – no humans needed (this started in the 1950s)
- **Problem** The state of the arts is not all that great

Here we Go



- Crucial for automated theorem proving are formal systems
- Go is often interpreted as a miniature formal system
- Idea Maybe modify Go neural networks (NN) for automated theorem proving

GPT and large language models



- The first successful Go NN were convolutional neural network (CNN)
- Enter, the transformer More modern versions use transformers (GPTs)
- Idea Use GPT for automated theorem proving

Enter, the theorem

OpenAI's GPT-f performance on closing proofs is

$\approx 56\%$

(in the Metamath environment, allowing 32 repeats)

- Quote This work is motivated by the possibility that a major limitation of automated theorem provers compared to humans – the generation of original mathematical terms – might be addressable via generation from language models
- Problem They mostly generated proofs are still quite "boring"

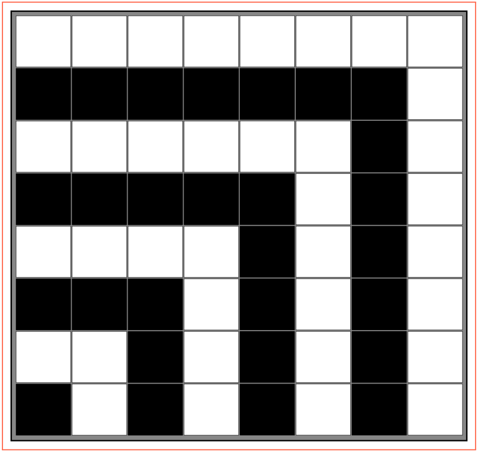
Table 3: Metamath theorems use by our Ring Algebra synthetic generators. Theorems are available in the Matmath Proof Explorer.

Theorem	Weight	Description
`eqcomd`	1	Commutative law for class equality.
`int-addcomd`	1	Addition commutativity.
`int-addassocd`	1	Addition associativity.
`int-mulcomd`	1	Multiplication commutativity.
`int-mulassocd`	1	Multiplication associativity.
`int-leftdistd`	3	Left distribution of multiplication over addition.
`int-rightdistd`	3	Right distribution of multiplication over addition.
`int-sqdefd`	5	Definition of the square.
`muladdd2`	5	Product of two sums

Examples of equalities produced by the generator:

$$ABBA(AB)^2 + (C + A) = A + (ABBA)^2 + C$$
$$(AA)^2 = A^2AA$$
$$((BA + CA)^2)^2 = (BA + CA)^2(BAAB + ACCA + BAAC + ABCA)$$
$$((A + B)^2)^2(A + A) = ((A + B)^2(AB + AB + AA + BB) + (A + B)^2(AB + AB + AA + BB))A$$

Make it shorter



- **Above** A very efficient proof of $\sum odd = square$
- They found **shorter** – and accepted by the math community – proofs
- **Problem** We are still far from "new proofs of new theorems"

Thank you for your attention!

I hope that was of some help.