

DISCRETE COMPUTATION:
THEORY AND OPEN PROBLEMS

Notes for the lectures by

Albert R. Meyer

Preceptorial Introduction to Computer Science for
Mathematicians

American Mathematical Society

San Francisco

January, 1974

This work was supported in part by the National
Science Foundation under research grant GJ-34671.

DISCRETE COMPUTATION
THEORY AND OPEN PROBLEMS

Notes for the lectures by

Albert R. Meyer

Preceptorial Introduction to Computer Science for
Mathematicians

American Mathematical Society

San Francisco

January, 1974

This work was supported in part by the National
Science Foundation under research grant GS-34671.

LECTURE I

1

MULTIPLICATION IN BINARY

$$\begin{array}{r} U = 101100 \\ V = \underline{111111} \\ 101100 \\ 101100 \\ \cdot \\ \cdot \\ \cdot \end{array}$$

$$U \times V = 101011010100$$

2

α = the add and shift multiplication algorithm

$T_{\alpha}(U,V)$ = time (number of basic operations on digits) to multiply U and V by method α .

$$T_{\alpha}(n) = \max \{ T_{\alpha}(U,V) \mid l(U) = l(V) = n \}$$

Remark: $T_{\alpha}(n) = O(n^2)$

3

Recursive algorithm for multiplication

$$U = \begin{array}{|c|c|} \hline U_1 & U_2 \\ \hline \end{array} = U_1 \cdot 2^{n/2} + U_2$$

$$V = \begin{array}{|c|c|} \hline V_1 & V_2 \\ \hline \end{array} = V_1 \cdot 2^{n/2} + V_2$$

$$U \cdot V = U_1 V_1 \cdot 2^n + (U_1 V_2 + U_2 V_1) \cdot 2^{n/2} + U_2 V_2$$

①

②

③

④

4

p = recursive algorithm

$$T_p(n) = 4T_p(n/2) + (\text{time to add and shift length } n \text{ numbers})$$

$$= 4T_p(n/2) + O(n)$$

$$= 4^k T_p(n/2^k) + O(n)$$

$$= O(4^{\log_2 n}) = O(n^2)$$

5

β = better recursive algorithm using only three half length multiplications

$$\textcircled{1} \quad (u_1 + u_2) \cdot (v_1 + v_2)$$

$$\textcircled{2} \quad u_1 v_1$$

$$\textcircled{3} \quad u_2 v_2$$

$$u \cdot v = \textcircled{2} \cdot 2^n + (\textcircled{1} - \textcircled{2} - \textcircled{3}) \cdot 2^{n/2} + \textcircled{3}$$

6

$$T_\beta(n) = 3T_\beta(n/2) + O(n)$$

$$= O(3^{\log_2 n})$$

$$= O(n^{\log_2 3})$$

$$\approx O(n^{1.6})$$

7

Best upper bound known for multiplication:

$$O(n \log n \cdot \log \log n)$$

by Strassen and Schönhage.

Question: What is the fastest possible way to multiply?

Need there even be one?

Might have algorithms β_1, β_2, \dots

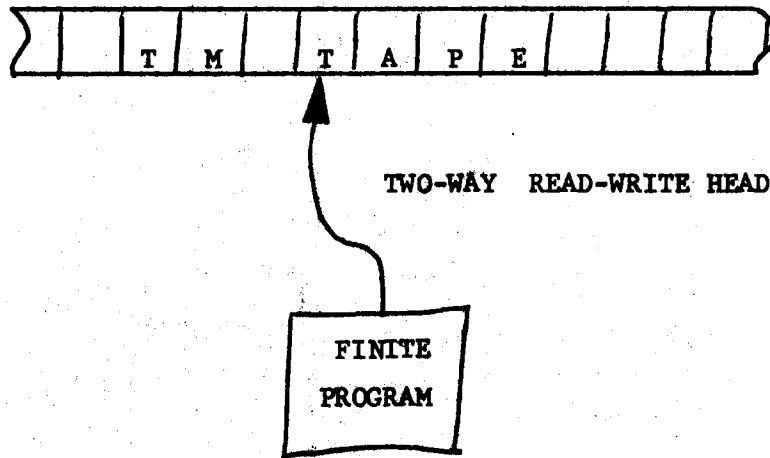
$$T_{\beta_1}(n) = n \log n$$

$$T_{\beta_2}(n) = n \sqrt{\log n}$$

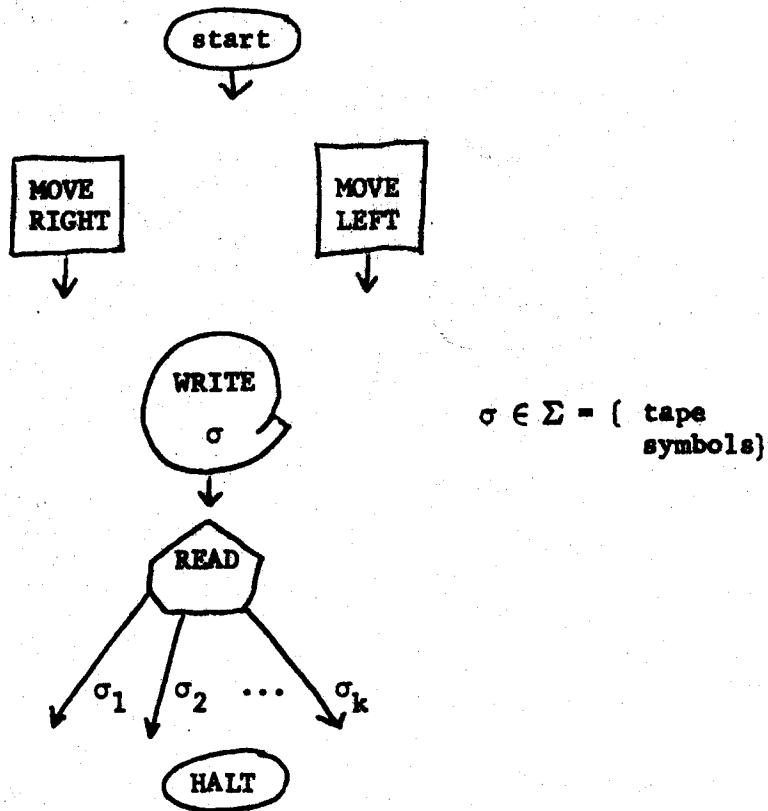
$$T_{\beta_i}(n) = n (\log n)^{\frac{1}{i}}$$

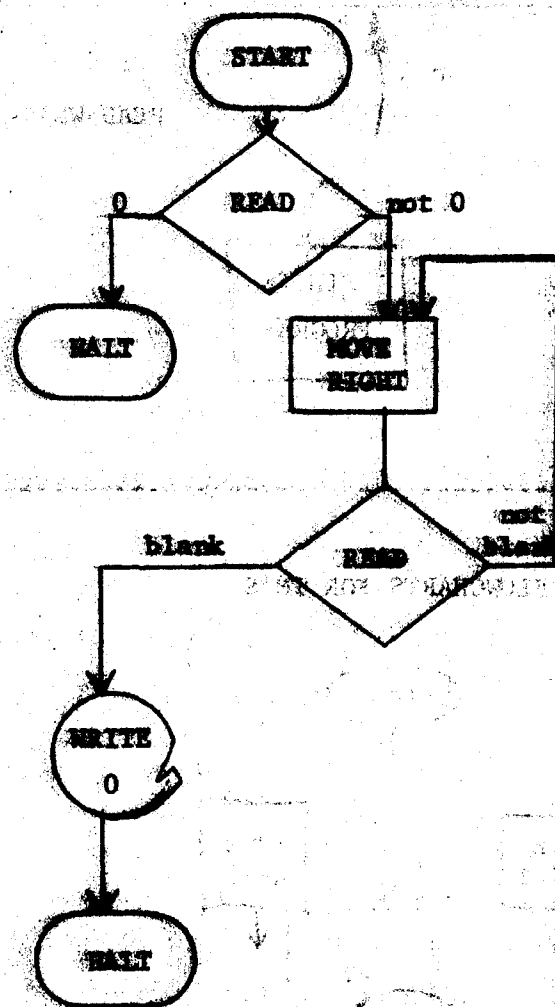
Then there is no fastest one.

Turing Machine



FLOWCHARTS FOR TM'S



FLOWCHART FOR $F(x) = 2x$ (Input x as integer in binary notation.)

\mathcal{M} a T.M.,
 x an input word.

$T_{\mathcal{M}}(x) =$ number of instructions executed by \mathcal{M}
 on x if \mathcal{M} halts; ∞ if \mathcal{M} doesn't halt
 on x .

$S_{\mathcal{M}}(x) =$ number of tape squares visited by
 head of \mathcal{M} with input x if \mathcal{M} halts;
 ∞ if \mathcal{M} does not halt.

$\varphi_{\mathcal{M}}(x) =$ output of \mathcal{M} on x , if any;
 ∞ if no output.

$T =$ time $S =$ space $\varphi =$ function

Church's Thesis:

The effectively (mechanically) computable
 functions and the Turing machine computable
 functions are the same.

Extended Church's Thesis:

If a function is computable in time T on
 any reasonable computer model, then it is
 computable in time \leq polynomial (T) on a
 Turing machine.

14

Infinitely-often Speed-up Theorem:

(M. BLUM). Let $t: N \rightarrow N$ be any computable function. Then there is a computable function $C_t: N \rightarrow \{0,1\}$ such that

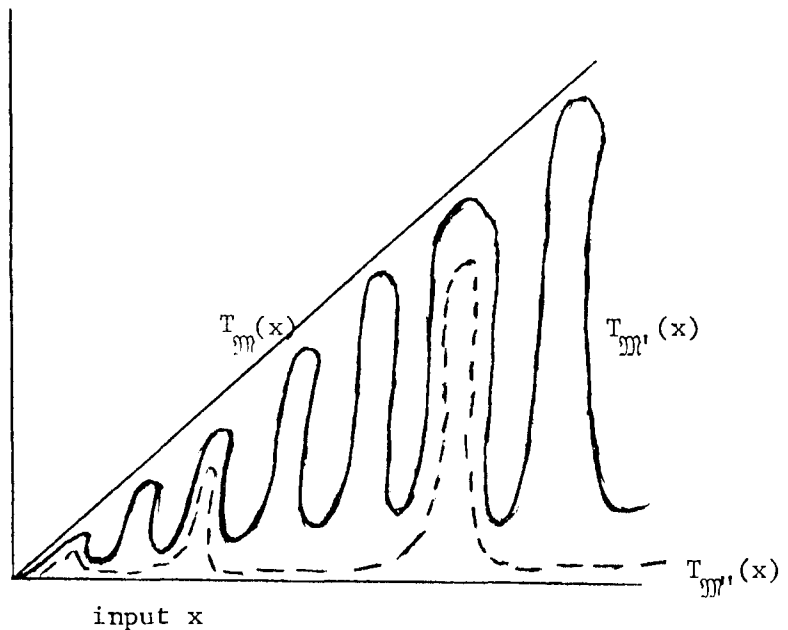
given any \mathfrak{M} computing C_t one can construct an \mathfrak{M}' also computing C_t with the property that

$$T_{\mathfrak{M}}(x) > t(x) \text{ and } T_{\mathfrak{M}'}(x) < \text{constant}$$

for infinitely many $x \in N$.

15

numbers of steps



\mathfrak{M}' is faster than \mathfrak{M} infinitely often,
 \mathfrak{M}' is faster than \mathfrak{M}' infinitely often,
etc.

16

Let M_0, M_1, \dots, M_i be an orderly list of
of all Turing machines (say in order of the
size of their flowcharts).

Let φ_i abbreviate φ_{M_i} ,

" T_i " T_{M_i} .

Universal Machine Theorem:

$\varphi_i(x)$, regarded as a function of both
 i and x , is computable.

16a

PADDING LEMMA:

Given any program, one can "pad" it with
instructions which it never uses. Thus,
we obtain

a new program with the same behavior
as the old one.

More formally,

17

LEMMA: For any T.M. \mathcal{M}_e there is an infinite

set $PAD(e) \subset \mathbb{N}$ such that

(1) for any $e' \in PAD(e)$

$$\varphi_e = \varphi_{e'}, \quad T_e = T_{e'}, \quad \text{and} \quad S_e = S_{e'}$$

and (2) there is a T.M. which recognizes elements of $PAD(e)$ in constant time.

(Think of $PAD(e)$ being binary numbers of the

form

e	#	irrelevant
-----	---	------------

)

18

Proof of L.G. Speed-up Thm:

Let

$$C_x(x) \stackrel{\text{df.}}{=} \begin{cases} 1 - \varphi_x(x) & \text{if } T_x(x) \leq t(x), \\ 0 & \text{otherwise.} \end{cases}$$

$$I \rightarrow y \stackrel{\text{df.}}{=} \begin{cases} 0 & \text{if } y \geq 1, \\ 1 & \text{if } y = 0. \end{cases}$$

18a

(1) C_t is computable (implicit in the Universal Machine Thm.)

(2) If $\varphi_{e'} = C_t$, then $T_{e'}(e') > t(e')$

and $C_t(e') = 0$ (by def. of C_t).

(3) Say $\varphi_e = C_t$. Then for any $e' \in \text{PAD}(e)$,

$t(e') < T_{e'}(e') - T_e(e')$

and $C_t(e') = 0$.

So speed-up \mathcal{R} by always testing if the

input is in $\text{PAD}(e)$, and if so immediately

print output 0.

Def. $\text{Time}(t) = \{ \varphi_i : \mathbb{N} \rightarrow \mathbb{N} \mid T_i(x) \leq t(x) \}$

almost everywhere)

$\text{Space}(t) = \dots S_i \dots$

Lemma. $C_t \notin \text{Time}(t)$ for any computable t .

Remark: Time to compute C_t depends on time to compute t .

Convention: $n = \text{length}(x) = l(x) \approx \log_2 x$. Thus,

$\text{Time}(2^n) = \text{Time}(2^{l(x)}) = \text{Time}(x)$,

$\text{Time}(2^{2n}) = \text{Time}(x^2)$, etc.

Def. A computable $t: \mathbb{N} \rightarrow \mathbb{N}$ is time-honest iff $t \in \text{Time}(t^3)$ and $t(x) \geq l(x)$.

Cor. (Compression Theorem, Hartmanis-Stearns)

For any time-honest t , $C_t \in \text{Time}(t^4) - \text{Time}(t)$.

Remark: Lots of time-honest fcn's.

$\lceil \log_2 x \rceil, x, 2^x, 2^{2^x}$ } closed under +, ·, exp, composition.

21

Is more time better than less?

Is

$$\text{Time}(t^4) - \text{Time}(t) \neq \emptyset$$

for all computable t ?

NOT NECESSARILY!

22

Gap Theorem: (Trachtenbrot, Borodin) For any computable g , there exist arbitrarily large computable t such that

$$\text{Time}(t) = \text{Time}(g \circ t)$$

22a

Proof of Gap Theorem.

Given g , define

$t(x)$ = the least z such that

$$(\text{Time})[T_1(x) < z \text{ or } T_1(x) > g(z)].$$

Honesty Theorem (McCreight, Meyer)

For every computable t , there is a time-honest t' such that

$$\text{Time}(t) = \text{Time}(t')$$

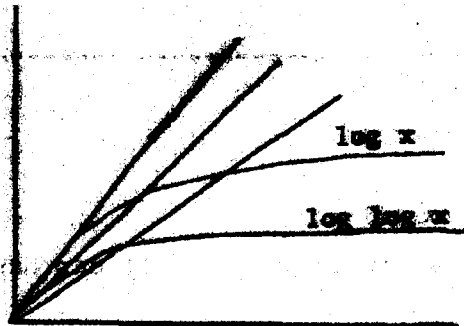
Summary:

For arbitrarily large t, t' it can happen that

$$\text{Time}(2^{2^t}) = \text{Time}(t) = \text{Time}(t') \subseteq \text{Time}((t')^4)$$

↑ ↑ ↑
 GAP HONESTY COMPRESSION

RECORDING PICTURE



Lines(f) = radial lines almost everywhere under f

Compression: For any line $L \neq 0$,

$$\text{Lines}(2L) \stackrel{=}{\neq} \text{Lines}(L)$$

Monesty: For any function f , there is a line L ,

$$\text{Lines}(f) = \text{Lines}(L).$$

Cap: $\text{Lines}(t) = \text{Lines}(2^t) = \{\text{zero line}\}$

for $t = \text{loglog}$.

Def. Let f be a computable function. A sequence t_1, t_2, \dots of functions is a (space)

complexity sequence for f iff

(1) If $\varphi_e = f$, then $S_e \geq t_i$ almost every
where for some i ,

and (2) For every i , there is a $\varphi_e = f$
such that

$$t_i \geq S_e \text{ almost everywhere.}$$

Def. A sequence of functions

p_1, p_2, \dots is an r.e. complexity sequence
(for space) iff

(1) $p_{i+1} \leq \left\lceil \frac{1}{2} \cdot p_i \right\rceil$ for all i ,

(2) for each i there is a j such that,

$$p_i = S_j$$

and (3) $p_i(x)$ is a computable function of i and x .

LECTURE II

1

Σ = finite set called the alphabet or vocabulary,

an element $\sigma \in \Sigma$ is called a letter.

Σ^* = set of all finite sequence of letters,

an element $x \in \Sigma^*$ is called a word.

2

Binary operation concatenation, written " . "

on Σ^* :

$x \cdot y = xy$ = word x followed by word y .

Example: $001 \cdot 01 = 00101$

$l(x)$ = length (number of occurrences of letter) of the word x .

$l(001) = 3$

$l(x \cdot y) = l(x) + l(y)$.

3

$\lambda \in \Sigma^*$ acts as an identity element under concatenation.

$$\lambda \cdot x = x \cdot \lambda = x \quad \text{for all } x \in \Sigma^*,$$

$$l(\lambda) = 0.$$

Remark 1: $\langle \Sigma^*, \cdot \rangle$ is the free monoid generated by Σ with identity λ .

Remark 2: Remark 1 is irrelevant.

Remark 3: λ is introduced as a technical convenience and could be eliminated in what follows at the expense of some minor awkwardness.

4

A set $L \subset \Sigma^*$ is called a language. Extending concatenation to languages in the usual way:

$L \cdot M$, also written $LM \stackrel{\text{def.}}{=}$

$$\{x \cdot y \mid x \in L \text{ and } y \in M\}$$

Example: $\{0\} \cdot \{0,1\} = \{00,01\}$

$$\{0,00\} \cdot \{1,01\} = \{01,001,0001\}$$

$\{0,1,\lambda\} \cdot \{0,1,\lambda\} \cdot \{0,1,\lambda\} =$ all binary words of length ≤ 3 (including λ).

5

For $x \in \Sigma^*$, $n \in \mathbb{N}$,

$$x^n = \underbrace{x \cdot x \cdots x}_n$$

$$x^0 = \lambda$$

Example: $(01)^3 = 010101$ Similarly for $A \subset \Sigma^*$

$$A^n = \underbrace{A \cdot A \cdots A}_n$$

$$A^0 = \{\lambda\}$$

6

Example: $(0,1)^4 =$ all binary words of lengthexactly 4. $(0,1,\lambda)^4 =$ all binary words of length ≤ 4 .
 $((((0,1)^2)^2)^2)^2 = (0,1)^{2^4} =$ all binary words of
length 16.

7

Important example:

$$\begin{aligned}
(01)^n &= \underbrace{0101\dots 01}_{2n} = \\
&= \{0,1\}^{2n} - (1 \cdot \{0,1,\lambda\}^{2n} \cup \\
&\quad \{0,1,\lambda\}^{2n} \cdot 0 \cup \\
&\quad \{0,1,\lambda\}^{2n} \cdot \{00,11\} \cdot \{0,1,\lambda\}^{2n})
\end{aligned}$$

= all binary words of length $2n$ which do not

(1) start wrong

or (2) end wrong

or (3) move wrong (contain a forbidden
local pattern)

8

Problem: Given two expressions involving letters

in Σ, λ , and operations

" . " concatenation

" \cup " union

" 2 " squaring

" \cap " intersection

" - " set difference

is there a way to tell if they describe the same
language?

YES!

BUT NO GOOD WAY!!

Lemma. An expression containing n operation symbols describes a subset of

$$(\Sigma \cup \lambda)^{2^n}$$

Proof. By induction on n :

If $n=0$, the expression must consist of a single letter or λ .

If E is an expression containing $n+1$ operations, then E is of the form

$$E_1 \cdot E_2$$

$$E_1 \cup E_2$$

$$(E_1)^2$$

$$E_1 \cap E_2$$

$$E_1 - E_2$$

where E_1, E_2 are expressions containing $\leq n$ operation symbols. Proof follows immediately.

12 For any expression E, let

$\mathcal{L}(E) \subset \Sigma^*$ be the language described by E.

Remark: Formally, $E_1 = E_2$ means that E_1 and E_2 are identical expressions. E_1 and E_2 are equivalent (written $E_1 \equiv E_2$) iff

$$\mathcal{L}(E_1) = \mathcal{L}(E_2).$$

13

$$E_1 \equiv E_2 \text{ iff } (E_1 - E_2) \cup (E_2 - E_1) = \emptyset$$

Hence sufficient to test whether an expression describes the empty set.

14

To test if $\mathcal{L}(E) = \emptyset$, convert E to a list of the words in $\mathcal{L}(E)$ beginning at the "innermost" subexpressions of E and working out.

See if the list is empty when you finish.

Difficulty: The list for

$$(\dots(((0 \cup 1)^2)^2)\dots)^2$$

n

contains

15

$2^n \cdot 2^{2^n}$ bits

16

Theorem 1. There is (for any finite Σ) a constant $k > 0$ and a Turing machine \mathcal{M} such that

- (1) \mathcal{M} accepts an input w iff w is a well-formed expression and $\mathcal{L}(w) = \emptyset$.
- (2) $T_{\mathcal{M}}(n) \stackrel{\text{df.}}{=} \max\{T_{\mathcal{M}}(x) \mid \mathcal{L}(x) \neq \emptyset, |x| \leq n\}$

$$\leq 2^{2^{kn}}$$

17

Theorem 2. There is a finite Σ and a constant $k > 1$ such that

if \mathcal{M} is any T.M. accepting precisely the expressions over Σ describing the empty set, then

$$\forall n \quad T_{\mathcal{M}}(n) > 2^{k\sqrt{n}}$$

for infinitely many n . (That is, $\{E \mid \mathcal{L}(E) = \emptyset\} \notin \text{Time}(2^{\sqrt{n}})$)

To prove Theorem 2:

- (i) Define a relation on languages

$$L_1 < L_2$$

with intuitive meaning that L_1 is easy to decide given L_2 .

- (ii) Show that for any $L \in \text{Time}(2^n)$

$$L < \{E \mid \mathcal{L}(E) = \emptyset\}.$$

- (iii) Deduce from the Compression Theorem that there is an $L \in \text{Time}(2^n)$ which is hard to decide.
- (iv) Conclude that $\{E \mid \mathcal{L}(E) = \emptyset\}$ is hard to decide.

Def. For $L_1 \subset \Sigma_1^*$, $L_2 \subset \Sigma_2^*$ we say $L_1 < L_2$

(L_1 is polynomial time reducible to L_2) iff

there exists a function $f: \Sigma_1^* \rightarrow \Sigma_2^*$

- (1) f is computable in time bounded by a polynomial in the length of its argument ($f \in \text{Time}(p \circ \ell)$ where $p: \mathbb{N} \rightarrow \mathbb{N}$ is a polynomial and $\ell: \Sigma_1^* \rightarrow \mathbb{N}$ is the length function),
- (2) $x \in L_1 \Leftrightarrow f(x) \in L_2$ for all $x \in \Sigma_1^*$.

20

Lemma. Let $t: N \rightarrow N$ be nondecreasing, and $t(n) \geq n$.

If $L_1 < L_2$ and $L_2 \in \text{Time}(t(n))$, then $L_1 \in \text{Time}(t(p(n)))$ for some polynomial $p: N \rightarrow N$

Contrapositive. If $L_1 \notin \text{Time}(2^{n/4})$ and $L_1 < L_2$, then $\exists k > 1$ such that $L_2 \notin \text{Time}(2^{\sqrt[k]{n}})$

21

Because $2^{n/4}$ is time-honest, the compression theorem implies $\exists L_1$ such that

$$L_1 \in \text{Time}(2^n) - \text{Time}(2^{n/4}).$$

Thm. 2 follows immediately from the preceding contrapositive if we show

22

Main Construction for Theorem 2.

Lemma. For any $L \in \text{Time}(2^n)$, there is an alphabet Σ such that

$$L < \{E \mid E \text{ is an expression over } \Sigma \text{ and } \mathcal{L}(E) = \emptyset\}.$$

Choose any $L \in \text{Time}(2^n)$

Say $L \subseteq A$ for some alphabet Σ .

Let M be a Turing machine which

(i) halts on any input of length n in $\leq 2^n$ steps, and

(ii) halts accepting a symbol "1" on its tape iff the input is in L .

for any $L \in \text{Time}(2^n)$, there is an alphabet Σ such that

$L = \{ x \mid x \text{ is an expression over } \Sigma \text{ and}$

$$L(x) = 1 \}$$

24

Let Q be the states (boxes in the flowchart) of

\mathcal{M} ,

let W be the tape symbols of \mathcal{M} including $b \in W$ for the blank tape symbol, let $\#$ be still another symbol.

$\Sigma \stackrel{\text{def.}}{=} Q \cup W \cup \{\#\}$.

25

For $x \in \Delta^*$, $l(x) = n$,

$\text{Comp}(x) \in \Sigma^*$ is to be:

$\# b^{2^n} \cdot \text{start} \cdot x b^{2^n} \#(\text{tape after one step})\# \dots$

$\dots \#(\text{tape after } k \text{ steps}) \#(\text{tape after } k+1 \text{ steps})\# \dots$

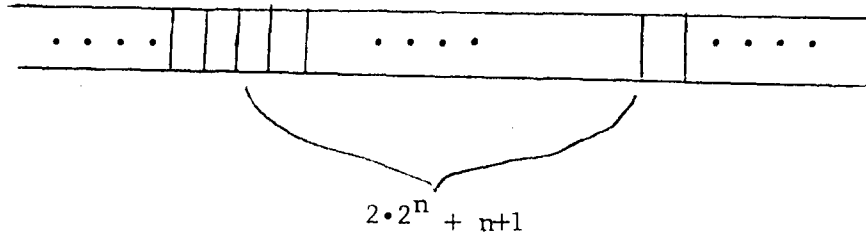
$\dots \# \text{tape } \text{halt} \text{ symbols } \#$

Exactly $2 \cdot 2^n + n + 1$ symbols between successive $\#$'s.

$l(\text{Comp}(x)) \leq 2^{3(n+1)} \stackrel{\text{df.}}{=} N$

26

Comp(x) has the property that any four consecutive letters determine the letter $2 \cdot 2^n + n$ to their right:



Let $F = \{(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5) \mid \sigma_5 \text{ is not the letter determined by } \sigma_1 \sigma_2 \sigma_3 \sigma_4\}$. This follows from the fact that at any step the next move of \mathfrak{M} is determined by the state and the tape symbol being scanned.

27

Comp(x) = (starts right) \cap
 (ends right) \cap
 $((\Sigma \cup \lambda)^N - (\text{moves wrong}))$

starts right: $\#b^{2^n} \cdot \text{start} \cdot x \cdot b^{2^n} \cdot \# \cdot (\Sigma \cup \lambda)^N$

ends right: $(\Sigma \cup \lambda)^N \cdot \text{halt} \cdot (\Sigma \cup \lambda)^N \cdot \#$

28

moves wrong:

$(\Sigma \cup \lambda)^N \cdot \left(\bigcup_F (\sigma_1 \sigma_2 \sigma_3 \sigma_4 \Sigma^{2 \cdot 2^n + n - 1} \cdot \sigma_5) \right) \cdot (\Sigma \cup \lambda)^N$

29

Let $\text{Rejects}(x) =$

$$(\Sigma \cup \lambda)^N \cdot \text{halt} \cdot (\Sigma - \{1\}) \cdot (\Sigma \cup \lambda)^N.$$

Then

$x \in L \Leftrightarrow M$ halts reading a 1

$$\Leftrightarrow \text{Comp}(x) \cap \text{Rejects}(x) = \emptyset.$$

But expressions for $\text{Comp}(x)$ and $\text{Rejects}(x)$

can be constructed in polynomial time in

$l(x)$, so $L < \{E \text{ over } \Sigma \mid \mathcal{L}(E) = \emptyset\}$.

Q.E.D.

30

Remarks: (1) Thm. 2 holds for expressions

using only " \cdot ", " \cup ", " 2 " and letters 0,1.

(2) If we allow " $\{0,1\}^*$ " to be used in expressions Stockmeyer has shown that

$$\{E \text{ with } \{0,1\}^* \mid \mathcal{L}(E) = \emptyset\} \in \text{Time } \left. \begin{matrix} 2^{2^{\dots^2}} \\ \dots \\ 2^2 \end{matrix} \right\} n$$

$$\text{but } \notin \text{Time } \left. \begin{matrix} 2^n \\ \dots \\ 2^2 \end{matrix} \right\} \in \cdot \log_2 n$$

for some fixed $\epsilon > 0$.

(3) If we allow only " \cup ", " \cdot ", the

equivalence problem is complete in \mathcal{NP}

(discussed in Karp's lecture).

31

Remark: Most decidable theories studied in mathematical logic require exponential time or worse. (An important exception being the propositional calculus, for which lower bounds larger than a polynomial are unknown.)

32

Open problems:

- (1) Can the satisfiable formulas of the propositional calculus be recognized in polynomial time? (This is the $P = NP$ question of Cook and Karp).
 - (2) Can a multi-tape Turing machine multiply integers (in binary notation) in linear time?
-

33

- (3) What is the relation between time and space?

Known: $S_M(n) \leq T_M(n) \leq c S_M(n)$

($c > 1$ depends on M)

Open: If $L \in \text{Time}(2^n)$ is $L \in \text{Space}(n)$?

- (4) Is $\text{Space}(n) = \text{Nondeterministic Space}(n)$?
(The LBA problem of Myhill)
-

- (5) Are linear time 3 tape T.M.'s more powerful than linear time 2-tape T.M.'s?
- (6) Can the primes (represented in binary) be recognized in linear time?
Can the context-free languages?
- (7) Can two $n \times n$ matrices be multiplied in proportional to $n^{2.8}$ arithmetic operations?
($n^{2.9}$ is known to be possible.)

References:Abstract Complexity

1. Borodin, A. Computational Complexity: Theory and Practice, in Currents in the Theory of Computing, A. Aho, ed., Prentice-Hall, Englewood Cliffs, N.J., 1973, pp.35-89.

2. Hartmanis, J. and J. Hopcroft, An overview of the theory of computational complexity, Jour. ACM, 18, 3 (1971), pp.444-475.

Fast arithmetic

1. Knuth, D. The Art of Computer Programming: Vol. 2, Seminumerical Algorithms, Addison-Wesley, Reading, Mass., 1969.

MIT/LCS/TM-39

**DISCRETE COMPUTATION:
THEORY AND OPEN PROBLEMS**

Albert R. Meyer

January 1974