

PARAMETER ESTIMATION OF A PROBABILISTIC AUTOMATA MODEL
OF DNA REPLICATION

BY

KUAN CHAN WU

B.S., Chung Cheng Institute of Technology, Taiwan, 1963.

A MASTER'S REPORT


submitted in partial fulfillment of the
requirements of the degree

MASTER OF SCIENCE

Department of Computer Science

Kansas State University
Manhattan, Kansas
1976.

Approved:


Major Professor

LD
2668
R4
1976
W82
C.2
Document

201

TABLE OF CONTENTS

	Page
Acknowledgement	iii
List of Tables	iv
List of Figures	v
Introduction	1
Background	2
I. Formal grammars and probabilistic grammars	2
II. Automata, Finite Automata and Type 3 Languages, Probabilistic Automata	6
III. Parameter Estimation and Nonlinear Programming	14
Solution Synthesis	19
1. Model of Meiosis (P-finite state automata)	19
2. Development of Equations	21
3. Experimental Data	22
4. Selection of Estimation Technique	23
Results	25
1. Estimates from Nonlinear Program	25
2. Sensitivity to Initial Guesses	28
3. Normalization	29
4. Simplification	33
5. Comparison of Running Times	36
Conclusion	38
Appendix	40
References	68

ILLEGIBLE

**THE FOLLOWING
DOCUMENT (S) IS
ILLEGIBLE DUE
TO THE
PRINTING ON
THE ORIGINAL
BEING CUT OFF**

ILLEGIBLE

Acknowledgement

I wish to express my sincere gratitude to my major professor, Dr. Fred J. Maryanski for his guidance, creative suggestions and encouragement during the writing of this report.

LIST OF TABLES

	Page
Table 1	26
Table 2	31
Table 3	34

LIST OF FIGURES

Figure		Page
1	State Diagram of The Finite Automaton	8
2	Diagram of States	12
3	Aviemore Model	20

Introduction

Automata theory is the study of the dynamic behavior of discrete parameter information systems. Within this area, we define not only problems that deal with digital computers but also problems associated with such topics as describing the behavior of nerve networks, the representation of the properties of languages etc. (1)

A Probabilistic Automaton is an automaton in which the characteristics of the mapping can be described only in a probabilistic rather than a deterministic manner. (2)

The purpose of this report is to estimate parameters of a probabilistic Finite-state Automaton model to fit as closely as possible. The model DNA (Deoxynucleic Acid) Meiosis, with the sample data taken from experiments conducted by Mortimer (3).

Background

For this study we will need the following basic properties.

I. Formal grammars and probabilistic grammars.

Definition I - 1

Grammar: A formal grammar is an ordered quadruple

$$G = (V_n, V_t, X_0, F)$$

where

V_n - is a finite collection of variables or nonterminal symbols.

V_t - is a finite alphabet of terminal symbols
such that $V_n \cap V_t = \phi$

X_0 - is the initial letter (start sample) $X_0 \in V_n$.

F - is a finite set of ordered pairs (P, Q)
elements (P, Q) of F are called rewriting
rules or productions and are written $P \rightarrow Q$,

$$P \in (V_n \cup V_t)^+, Q \in (V_n \cup V_t)^*$$

P is the premise of the rule.

Q is the consequence of the rule.

$$(V_n \cup V_t)^+ = (V_n \cup V_t)^* - \lambda$$

λ is the empty string.

Note:

$(V_n \cup V_t)^*$ means any string of zero or more nonterminal or terminal symbols.