Fuzzy Star Functions, Probabilistic Automata, and Their Approximation by Nonprobabilistic Automata*

A. Paz†

Department of Electrical Engineering and Computer Sciences,
Electronics Research Laboratory,
University of California, Berkeley, California

Received October 16, 1967

ABSTRACT

Let \( \Sigma^* \) be the set of all tapes over \( \Sigma \) and let \( \psi \) be a function \( \psi: \Sigma^* \rightarrow [0, 1] \). The problem of approximating such functions (which can be defined, in particular, by probabilistic automata) by nonprobabilistic automata is investigated, in several aspects.

INTRODUCTION

Let \( \Sigma \) be a finite alphabet and let \( \Sigma^* \) be the set of all words over \( \Sigma \). Suppose that a preassigned function \( f: \Sigma^* \rightarrow [0, 1] \) is given and consider the following problems:

To construct a physical device (black box) which is capable of reading words fed into it and such that after a word \( x \) is read, will produce an output \( q(x) \) such that:

1. \( q(x) \) will provide us with full information as to the exact value \( f(x) \), for all \( x \in \Sigma^* \).
2. \( q(x) \) will provide us with full information as to whether \( f(x) > \lambda \) or \( f(x) \leq \lambda \) for a given real number \( \lambda, 0 \leq \lambda < 1 \) and for all \( x \in \Sigma^* \).
3. \( q(x) \) will provide us with enough information so that we shall be able to approximate \( f(x) \) within any preassigned \( \epsilon \).
4. \( q(x) \) will provide us with enough information so that we shall be able to approximate within any preassigned \( \epsilon \) whether \( f(x) > \lambda \) or \( f(x) \leq \lambda \) for given \( \lambda, 0 \leq \lambda < 1 \).

The solution to the above problems may play an important role in domains such as pattern recognition, communication of information and analysis of noise processes in

* The research reported herein was supported by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force under AFOSR Grant No. AF-AFOSR-639-66.
† On leave from the Technion Israel Institute of Technology.
sequential network. The study of functions $f$ of the above form but over general spaces (not necessarily $\Sigma^*$), has been introduced by Zadeh [7] in the recent few years with a similar purpose, functions over $\Sigma^*$ which may result from noise processes have been studied by several authors (see, for example, the paper of Booth [2]) while the study of a special kind of functions of the above type, namely, functions induced by probabilistic automata have been studied by Rabin [3], Paz [4] and others.

It seems to the author that the solution to the problems (1) and (2) above is involved, in general, with too stringent requirements and the following considerations will make this point clear.

In connection with the second problem it was proved by Rabin ([3], Theorem 2) that even the functions $f$ over $\Sigma^*$ induced by probabilistic automata are a noncountable set, but the set of Turing machines (which is the most powerful set of sequential devices having a finite number of internal states), is a countable set. Rabin has also proved that if a function induced by a probabilistic machine is such that there is a number $\delta > 0$ with $|f(x) - \lambda| > \delta$ for all $x \in \Sigma^*$ then $f$ can be realized by a finite state automaton, but the above condition (that $|f(x) - \lambda| > \delta$) seems to hold only for degenerate cases. It should be pointed out here that in addition to the above mentioned theorem of Rabin there are other theorems on probabilistic automata which seem to imply that the nature of the cutpoint $\lambda$ and its relation to the range of the function $f(x)$ has much to do with the equivalence or nonequivalence of probabilistic automata to nonprobabilistic automata. Such theorems are the following.

**Theorem 1 (Paz, [4]).** There exists a probabilistic automata $A$ such that the set of words $T(A, \lambda)$ (e.g., the set of all words $x$ such that $f(x) > \lambda$ where $f(x)$ is the function induced by $A$) is regular iff $\lambda$ is a rational number.

**Theorem 2 (Paz, [4]).** There exists a probabilistic automaton $A$ such that $T(A, \lambda)$, with a given rational number $\lambda$, is not a regular set.

**Theorem 3 (Salomaa, [5]).** For any probabilistic automaton $A$ with a single input letter, there are only finitely many numbers $\lambda$ such that $T(A, \lambda)$ is not regular.

In the light of the above considerations we come to the conclusion that we need a new and weaker criterion for comparison between probabilistic automata and deterministic automata, a criterion which will be able to overcome the cardinality gap between the two kind of automata and which will loosen the stringent quality of the cut-point $\lambda$.

We shall therefore consider only the problems (3) and (4) introducing the concept of $\epsilon$-approximation of functions from $\Sigma^*$ into the interval $[0, 1]$, by sequential machines (an attempt to introduce $\epsilon$-approximation has already been done by the author in Ref. [4] but in a much narrower sense). It is worth mentioning however that also
Problems (1) and (2) have been considered by several authors ([3], [4], [6]) in some aspects. The main results of our paper are given below.

1. A characterization of functions for which the Problem (3) is solvable by using finite automata as the approximating device.

2. The same for Problem (4).

3. Problem (4) is shown to be nonequivalent to Problem (3).

4. A noncountable class of functions is shown to be approximable by finite automata whose realization is given explicitly.

5. Functions induced by probabilistic automata are shown to have a property which is necessary, but not sufficient for the existence of a solution to Problems (3) and (4).

6. It is shown that there are functions induced by probabilistic automata which are not approximable by finite automata in the sense of Problems (3) and (4). Therefore the class of probabilistic automata is stronger than the class of non-probabilistic automata, this being a consequence of the intrinsic nature of the probabilistic automata and not of the actual properties of the cut-point \( \lambda \) or of the relation of \( \lambda \) to the range of the function \( f(x) \).

7. Some open problems are posed.

1. Definitions and Notations

It is assumed that the reader is familiar with automata theory and probabilistic automata theory. A detailed account on these topics can be found in Refs. [3], [7]-[10], etc. Nevertheless, for the sake of completeness, we recapitulate here the basic definitions and introduce some notations which we shall need.

Let \( \Sigma \) be a finite nonempty alphabet. The elements of \( \Sigma \) are called symbols. Finite sequences of symbols are called words. The set of all words over \( \Sigma \) is denoted by \( \Sigma^* \), the empty word is denoted by \( \Lambda \). If \( x \) and \( y \) are words, \( xy \) denotes the concatenation of \( x \) and \( y \) (\( x \Lambda = \Lambda x = x \)), \( l(x) \) denotes the length of \( x \), and \( y \) is a \( k \)-suffix (prefix) of \( x \) if \( x = uy(x = yu) \) for some (possible empty) word \( u \) and \( l(y) = k \).

Definition 1.1. Following Zadeh [1] we define a fuzzy star function (f.s.f.) \( \psi \), as a function which associates with each tape \( x \) in \( \Sigma^* \) a real number in the interval \([0, 1]\).

Definition 1.2. A fuzzy star acceptor (f.s.ac.) is a f.s.f. together with a cut point \( \lambda \) (a real number), \( 0 \leq \lambda < 1 \).
DEFINITION 1.3. If \((\psi, \lambda)\) is a f.s. ac. then \(T((\psi, \lambda))\) (the set of tapes accepted or defined by \((\psi, \lambda)\)) is
\[
T((\psi, \lambda)) = \{x \mid x \in \Sigma^*; \psi(x) > \lambda\}.
\]

We shall now define and make distinctions between probabilistic, deterministic, automata acceptors and machines.

DEFINITION 1.4. A finite automaton (f.a.) over the alphabet \(\Sigma\) is a system \(A = (S, M, s_0)\) where \(S\) is a finite nonempty set (the internal states of \(A\)), \(s_0\) is an element of \(S\) (the initial state) and \(M\) is a function from \(S \times \Sigma\) into \(S\) (the transition function).

Let \(M\) be the extension \(M : S \times \Sigma^* \rightarrow S\) of \(M(M(s_i, A) = s_i\) for all \(s_i \in S\)). For the purpose of distinguishing an f.a. \(A = (S, M, s_0)\) we shall use the notation \(A(s_i, x) = M(s_i, x)\).

DEFINITION 1.5. A finite acceptor (f. ac.) is an f.a. together with a subset \(F\) of \(S\) (the set of final states). The set of words defined or accepted by an f. ac. \(A\) is the set \(T(A) = \{x \mid A(s_0, x) \in F\}\).

DEFINITION 1.6. If a set of words \(U\) is equal to \(T(A)\) for some f. ac. \(A\), then \(U\) is called a regular set.

Given a set of words, \(U \bar{U}\) denotes the complement of \(U\) in \(\Sigma^*\). The bar notation will be used for this purpose only in connection with sets of words.

Other nonprobabilistic devices of a more complex structure (push down automata, linear bounded, etc.) can be defined in the same form and the reader is referred to Ref. [9] for their definition. When referring to machines (instead of automata) we refer to devices which, at each instant of time, receive an input and yield an output.

NOTATION. \(P_n\) denotes the set of all \(n\)-dimensional probabilistic vectors.

DEFINITION 1.7. A probabilistic automaton is a system \(A = (S, \pi, \{A(\sigma)\}; F)\) where \(S\) is a finite set (the set of states of \(A\)), \(\pi\) is an element in \(P_n\), \(n\) being the number of elements in \(S\) (\(|S| = n\)), which represents the "initial distribution" of \(A\). \(\{A(\sigma)\}\) is a set of \(|\Sigma|\) stochastic matrices of order \(n\) (the "transition matrices" of \(A\)) such that \(A(\sigma) = [a_{ij}(\sigma)]\) and \(a_{ij}(\sigma)\) is the probability that the automaton will enter the state \(s_j\) beginning from state \(s_i\), and after scanning the symbol \(\sigma\). Finally, \(F\) is a subset of \(S\) (the set of final states).

NOTATION. \(A(x) = [a_{ij}(x)]\) denotes the matrix \(A(\sigma_1) \cdots A(\sigma_k)\) where \(x = \sigma_1, \ldots, \sigma_k\), and \(\pi(x)\) denotes the vector \(\pi A(x)\). It follows that \(\pi(xy) = \pi(x) A(y)\).
It is easily seen that $a_{ij}(x)$ is the probability that the automaton will enter the state $s_i$ beginning from state $s_j$ and after scanning the word $x$, and $\pi(x)$ is the final distribution over the states, beginning with initial distribution $\pi$ and after scanning the word $x$.

It is assumed throughout that the values $a_{ij}(x)$ are computable.

Let $\eta^F$ be an $n$-dimensional column vector $\eta^F = (\eta_i^F)$ such that

$$
\eta_i^F = \begin{cases} 
1 & \text{if } s_i \in F \\
0 & \text{otherwise,}
\end{cases}
$$

then $\eta^F(x) = A(x) \eta^F$ denotes a column vector the $i$th entry of which is the probability of entering a state in $F$ when beginning in state $s_i$ and after scanning the word $x$.

Thus $p(x)$, the probability of entering a state in $F$ when beginning with initial distribution $\pi$ and after scanning the word $x$ is

$$
p(x) = \pi A(x) \eta^F = \pi(x) \eta^F = \pi \eta^F(x)
$$

more generally

$$
p(xy) = \pi(x) \eta^F(y)
$$

**Notation.** $p(A, x)$ is the value $p(x)$ above, for a word $x$ as related to a given p.a. $A$.

**Definition 1.8.** A probabilistic acceptor (p. ac.) is a p.a. $A$ together with a cut point $\lambda$, $0 \leq \lambda < 1$ (it is clear that any p. ac. gives rise to an f.s. ac.). The set of words defined or accepted by a p. ac. $(A, \lambda)$ is the set $T(A, \lambda) = \{x \text{ } | \text{ } p(A, x) > \lambda, x \in \Sigma^*\}$.

**Definition 1.9.** Two acceptors are equivalent if they accept the same set of words.

**Definition 1.10.** An f.a. $B$ is equivalent to a p.a. $A$ (to an f.s.f. $\psi$) if there is a function $\varphi$ from the set of states $S$ of $B$ into the interval $[0, 1]$ such that for all $x \in \Sigma^*$

$$
\varphi(B(s_0, x)) = p(A, x),
$$

Note that if $B$ is a f.a. which is equivalent to a p.a. $A$ (f.s.f. $\psi$) then, for any cut-point $\lambda$, it is possible to transform $B$ into an f. ac. which is equivalent to the p. ac. $(A, \lambda)$ [f.s. ac. $(\psi, \lambda)$].

**2. $\epsilon$-Approximating by Nonprobabilistic Devices**

**Definition 2.1.** A Turing automaton over the alphabet $\Sigma$ is a system $A = (S, M, s_0)$ where $S$ is a finite nonempty set (the internal states of $A$) $s_0$ is an element of $S$ (the initial state) and $M$ is a function from $S \times \Sigma$ into
$S \times \Sigma \times \{-1, 0, 1\}$ (the transition function). The operation of a Turing machine is described in the following way: An infinite tape having an input word on it is fed into the machine. The machine begins scanning the input word on the tape, starting at the leftmost input symbol in its initial state and reading a symbol at a time. If the scanned symbol is $\sigma_i$ and the machine is in state $s_j$ then it replaces $\sigma_i$ by another symbol $\sigma_k$, changes its state into a new state $s_t$ and moves the tape either left or right or keeps it stationary depending on the value of $m$ in $M(s_j, \sigma_i) = (s_t, \sigma_k, m)$ ($m = 1$ means left move, $m = -1$ means right move, $m = 0$ means no move).

**Definition 2.2.** A linear bounded automaton is the same as a Turing automaton but its input tape is finite. The length of the input tape of a linear bounded automaton is a linear function of the length of the input word printed on the tape.

Any type of nonprobabilistic automata having finitely many states and capable of reading tapes will be called a nonprobabilistic (np.) device.

**Definition 2.3.** An np. device $B$ $\epsilon$-approximates an f.s.f. $\psi$ if there is a function $\varphi$ from the states of $B$ into the interval $[0, 1]$ such that for all $x \in \Sigma^*$ we have that

$$| \psi(x) - \varphi(B(s_0, x)) | \leq \epsilon.$$ 

If for given $\psi$ and $\epsilon$ there is $B$ satisfying the above properties, then $\psi$ is $\epsilon$-approximable.

The motivation of this criterion of comparison is made evident by the following proposition and corollary.

**Definition 2.4.** A Turing automaton (or a linear bounded automaton) $B$ $\epsilon$-computes the f.s.f. $\psi$ if $B$ when fed with a tape $x$ will perform a finite number of operations, at the end of which a number $\xi$ will be printed on the tape, such that $| \xi - \psi(x) | \leq \epsilon$.

**Remark.** We assume here that the symbols in $\Sigma$ are numerals so that a word $x = \sigma_1, ..., \sigma_k$ represents the number $\sigma_1, ..., \sigma_k$.

**Proposition 2.1.** Given an f.s.f. $\psi$ and $\epsilon$, $\psi$ is $\epsilon$-approximable by a Turing automaton (or linear bounded automaton) $B$ iff $\psi$ is $\frac{1}{\epsilon}$ $\epsilon$-computable by $B$.

The proof of this proposition follows easily from the definitions and is left to the reader.

**Corollary 2.2.** The f.s.f. induced by p.a.'s are $\epsilon$-approximable by Turing automata.

It is clear that Turing automata, having infinite tape, can compute the f.s.f.'s induced by p.a.'s for any tape $x$ within any arbitrary given $\epsilon$, but this may not be true for linear bounded automata. Note, however, that the set of f.s.f.'s induced by p.a.'s is not a countable set, but the set of Turing automata is countable.
3. $\varepsilon$-Approximating by F.A. Characterization

We proceed now to characterize f.s.f.'s which are $\varepsilon$-approximable by finite automata.

**Definition 3.1.** Given an f.s.f. $\psi$ and $\varepsilon > 0$, an $\varepsilon$-cover induced by $\psi$ is a finite set $\{C_i\}_{i=0}^k$ where $C_i$ are sets of points in the interval $[0, 1]$ satisfying the following requirements.

1. $\bigcup_{i=0}^k C_i = \{\xi \mid \psi(x) = \xi, x \in \Sigma^*\}$.
2. $\xi_1, \xi_2 \in C_i \Rightarrow |\xi_1 - \xi_2| \leq \varepsilon, \ i = 0, 1, 2 \ldots k$.
3. Let $C_{iz}$ for $z \in \Sigma^*$ be defined as

$$C_{iz} = \{\xi \mid \psi(xz) = \xi, \psi(x) \in C_i\}.$$ 

Then for any $i$ and any $z$ there is $j$ such that $C_{iz} \subseteq C_j$.

We are now able to prove the following:

**Theorem 3.1.** Given an f.s.f. $\psi$ and $\varepsilon > 0$, $\psi$ is $\varepsilon$-approximable by a f.a. $B$ iff there is an $2\varepsilon$-cover induced by $\psi$.

**Proof.** Suppose there is an $\varepsilon$-cover induced by $\psi$, define the f.a. $B$ as follows:

The states of $B$ are $C_0 \ldots C_k$ (the elements of the $\varepsilon$-cover). Let $C_0$ be the first set such that $\psi(A) \in C_0$, then the initial state of $B$ is $C_0$.

The transition function of $B$ is defined by the relation

$$B(C_i, \sigma) = C_j \quad \text{if} \quad C_i \sigma \subseteq C_j$$

and $j$ is the smallest index satisfying this relation.

Finally, set

$$\varphi(C_i) := \frac{1}{2} \left[ \sup_{\xi \in C_i} \xi + \inf_{\xi \in C_i} \xi \right]$$

We prove first, by induction, that for any $x \in \Sigma^*$, $\psi(x) \in B(s_0, x)$.

(i) For $x = A$ the statement follows from the definition of $B$.

(ii) Let $x$ be a tape with $l(x) = t$ and assume $\psi(x) \in B(s_0, x) = C_i$. Then $\psi(x\sigma) \in C_i \sigma = B(s_0, x\sigma)$ by the definitions of $C_i \sigma$ and $B$. Our statement is thus proved.

We have therefore that, for any $x \in \Sigma^*$,

$$|\psi(x) - \varphi(B(s_0, x))| = |\psi(x) - \frac{1}{2} \left[ \sup_{\xi \in C_i} \xi + \inf_{\xi \in C_i} \xi \right]| \leq \varepsilon$$
using the fact that \( \psi(x) \in C_i \) and the second property of the \( \epsilon \)-cover. Assume now that \( \psi \) is \( \epsilon \)-approximable by an np. a. \( B \). Let the states of \( B \) be \( s_0, s_1, ..., s_k \), and define the sets

\[
C_i = \{ \xi \mid \psi(x) = \xi, B(s_0, x) = s_i \} \quad i = 0, 1, ..., k.
\]

It is easily verified that the set \( \{C_i\} \) thus defined is a 2-\( \epsilon \) cover as required.

4. APPLICATIONS

**DEFINITION 4.1.** A f.s.f. \( \psi \) is quasi-definite if for any \( \epsilon \) there is \( k(\epsilon) \) such that for any \( x \) with \( l(x) \geq k(\epsilon) \) the inequality \( |\psi(x) - \psi(y)| \leq \epsilon \) holds, where \( y \) is the \( k(\epsilon) \)-suffix of \( x \).

Quasi definite p.a. have been introduced by the author elsewhere [4] by a similar definition. In that paper, a decision procedure has been given for ascertaining whether a given p.a. is quasi-definite, and some other properties of those machines have been studied. A theorem similar to the theorem to be proved now has been proved in that paper but for p. ac.

Using Theorem 3.1 we shall now prove the following:

**THEOREM 4.1.** Any quasi-definite f.s.f. \( \psi \) is \( \epsilon \)-approximable by an f.a. for any given \( \epsilon \).

**Proof.** Given \( \psi \) and \( \epsilon \) we define the following \( \epsilon \)-cover induced by \( \psi \).

Let \( y_1 \cdots y_t \) be all the words such that \( l(y_i) < k(\epsilon) \) where \( k(\epsilon) \) is as in Definition 4.1.

Let \( z_1 \cdots z_q \) be all the words with \( l(z_i) = k(\epsilon) \). Define the sets \( C_i \) as follows:

\[
C_i = \{ \xi \mid \psi(y_i) = \xi \}, \quad i = 1, 2, ..., t,
\]

\[
C_{t+i} = \{ \xi \mid \psi(xz_i) = \xi, x \in \Sigma^* \}, \quad i = 1, 2, ..., q.
\]

It is clear that

\[
\bigcup_{i=1}^{t+q} C_i = \{ \xi \mid \psi(x) = \xi, x \in \Sigma^* \}.
\]

If \( \psi(u) \) and \( \psi(v) \) are in the same set \( C_{t+i} \) (the sets \( C_i \) for \( i \leq t \) are one-sets and therefore are out of consideration), then \( u = u_k z_i \) and \( v = v_k z_i \) so that

\[
|\psi(u_k z_i) - \psi(z_i)| \leq \frac{1}{2} \epsilon \quad \text{and} \quad |\psi(v_k z_i) - \psi(z_i)| \leq \frac{1}{2} \epsilon
\]

by the quasidefinite property, with the result that

\[
|\psi(u) - \psi(v)| = |\psi(u_k z_i) - \psi(v_k z_i)| \leq \epsilon.
\]
Finally, using the definitions we have that

\[ C_{t+1}w = \{ \xi : \psi(\xi w) = \xi, \psi(u) \in C_{t+1} \} = \{ \xi : \psi(xz, xw) = \xi, x \in \Sigma^* \} = \{ \xi : \psi(xz_1x_2) = \xi, x \in \Sigma^* \} \subseteq C_{t+1}, \]

where \( z_j \) is the \( k(\frac{1}{2} \epsilon) \) suffix of \( zbw \).

The set \( \{ C_i \}_{i=0}^{t+1} \) is thus seen to be an \( \epsilon \)-cover, and by Theorem 3.1 the proof is complete.

**Remark.** A definite f.a. can be defined as an automaton \( B \) such that there is an integer \( k \) with the property that \( B(s_0, xy) = B(s_0, y) \) for all \( y \) with \( l(y) \geq k \) and all \( x \).

It is clear from the proof of the previous theorem, that if \( A \) is a quasi-definite automaton then \( A \) is \( \epsilon \)-approximable by a definite automaton \( B \). The converse is also true and the simple proof of this fact is omitted. One may now ask whether there are f.s.f.'s which are \( \epsilon \)-approximable by np. a.'s but are not quasi-definite. That this is possible is shown by the following:

**Example 4.1.** Consider the f.s.f. \( \psi \) induced by the p.a. \( A \) defined as follows:

\[ \Sigma = \{ a, b \}, \quad S = \{ s_0, s_1 \}, \]

the initial vector is \( \pi = (1, 0) \), the final vector is \( \eta = (0) \), the transition matrices are

\[ P(a) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad P(b) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \]

Thus \( \psi(x) \) is the 1.1 entry in the matrix \( P(x) \). It is easily seen that this function \( \psi \) is not quasi-definite, for \( \psi(x) = \psi(bk) \) where \( k \) is the number of \( b \)-s in the word \( x \) (e.g., \( \psi(ba^n) = \psi(b) \) for any \( n \) and therefore for any \( n \); \( \psi(ba^n) - \psi(a^n) \mid = \frac{1}{2} \), in contradiction to Definition 4.1.

To show that \( \psi \) is \( \epsilon \)-approximable for any \( \epsilon \), consider the following cover induced by \( \psi \).

Divide the interval \([0, 1]\) by \( k - 1 \) points \( \xi_1, \xi_2, \ldots, \xi_{k-1} \) (\( \xi_0 = 0, \xi_k = 1 \)) so that \( \xi_i - \xi_{i-1} = \frac{1}{2} \epsilon \). \( i = 1, 2, \ldots, k - 1 \) and \( 1 - \xi_{k-1} \leq \frac{1}{2} \epsilon \).

Define the cover \( \{ C_i \}_{i=0}^{k-2} \) as follows

\[ C_i = \{ \xi : \xi_i \leq \xi \leq \xi_{i+2} \} \cap \{ \xi : \psi(x) = \xi, x \in \Sigma^* \}, \quad i = 0, 1, \ldots, k - 2. \]

The Properties (1) and (2) of an \( \epsilon \)-cover are satisfied by the very definition of the \( C_i \) - s. As for Property (3) let \( \psi(x) \) and \( \psi(y) \) be in the same set \( C_i \). Then \( \psi(xa) \) and \( \psi(ya) \) are also in the same set \( C_i \). If

\[ \pi(x) = (u, 1 - u) \quad \text{and} \quad \pi(y) = (v, 1 - v), \]
then
\[ u - v \leq \epsilon \text{ (for } u = \psi(x) \text{ and } v = \psi(y) \text{ with } \psi(x), \psi(y) \in C_i) \]
and
\[
\begin{align*}
\pi(xb) &= \left(\frac{1}{2} u + 1 - u, \frac{1}{2} u \right) = (1 - \frac{1}{2} u, \frac{1}{2} u), \\
\pi(yb) &= \left(\frac{1}{2} v + 1 - v, \frac{1}{2} v \right) = (1 - \frac{1}{2} v, \frac{1}{2} v),
\end{align*}
\]
so that
\[
|\psi(xb) - \psi(yb)| = |(1 - \frac{1}{2} u) - (1 - \frac{1}{2} v)|
= |\frac{1}{2} (v - u)| \leq \frac{1}{2} \epsilon.
\]
By induction we have that, for all \( z \), either
\[
\psi(xz) = \psi(x) \quad \text{and} \quad \psi(yz) = \psi(y)
\]
or
\[
|\psi(xz) - \psi(yz)| \leq \frac{1}{2} \epsilon
\]
provided that
\[
|\psi(x) - \psi(y)| \leq \epsilon.
\]
This implies that for any \( x \) and \( i \), either \( C_i x = C_i \) or the maximal distance between any two points in \( C_i x \) is not greater than \( \frac{1}{2} \epsilon \). Now let \( \xi < \eta \) be two points in \( C_\epsilon z \), suppose
\[ \xi_j \leq \xi \leq \xi_{j+1}; \]
then, because
\[ |\xi - \eta| \leq \frac{1}{2} \epsilon, \]
we have that
\[ \xi_j \leq \eta \leq \xi_{j+2} \]
and therefore \( \xi \) and \( \eta \) are both in \( C_j \). Thus \( C_\epsilon z \subseteq C_j \) for some \( j \) and the function \( \psi \) is \( \epsilon \)-approximable.

5. THE \( \bar{P}_\epsilon \) RELATION

**Definition 5.1.** Given an f.s.f. \( \psi \) and \( \epsilon > 0 \), a \( \bar{P}_\epsilon \) relation induced by \( \psi \) is a relation over \( \Sigma^* \) having the following properties:

1. \( \bar{P}_\epsilon \) is symmetric and reflexive;
2. \( \bar{P}_\epsilon \) is right invariant, i.e., \( xP_\epsilon y \Rightarrow xzP_\epsilon yz \) for \( x, y, z \in \Sigma^* \);
3. \( xP_\epsilon y \Rightarrow |\psi(x) - \psi(y)| \leq \epsilon. \)
Definition 5.2. A relation over a set $X$ is of finite index if there is an integer $k$ such that in any subset of $k + 1$ elements of $X$ there are at least two relatives.

Lemma 5.1. Consider the relation $P$, defined by

$$xP,y \iff (z) | \psi(xz) - \psi(yz) | \leq \epsilon$$

for given $\psi$, and $\epsilon$. Any relation $P$, as above is a refinement of $\tilde{P}_\epsilon$.

Proof. The proof is obvious.

Lemma 5.2. For given f.s.f. $\psi$ and $\epsilon$, if $\psi$ is $\epsilon$-approximable by an f.a. then there is a $P_{2\epsilon}$ relation of finite index induced by $\psi$.

Proof. Consider the machine $B$ which $\epsilon$-approximates $P$ and, as in the proof of Theorem 3.1, let $\{C_i\}_{i=0}^k$ be the $\epsilon$-cover defined by $B$. Thus $\psi(x)$ and $\psi(y)$ are in the same $C_i$ if $B(s_0, x) = B(s_0, y) = s_i$.

Define now the relation $P_\epsilon$ over $\Sigma^*$ by the definition

$$xP_\epsilon,y \iff B(s_0, x) = B(s_0, y).$$

$P_\epsilon$ is clearly symmetric, reflexive, right invariant, and of finite index, by its very definition. As for condition (3) in Definition 5.1, it is also satisfied, for $xP_\epsilon y$ implies that $\psi(x)$ and $\psi(y)$ are in the same $C_i$ and, as in the proof of Theorem 3.1, this implies that

$$\psi(x) - \psi(y) \leq 2\epsilon.$$

Combining Lemma 5.1 and Lemma 5.2 we have the following:

Proposition 5.3. Given an f.s.f. $\psi$ and $\epsilon$, if $\psi$ is $\epsilon$-approximable by an f.a., then the explicit relation $P_{2\epsilon}$ defined in Lemma 5.1 is of finite index.

Consider the set

$$P_n = \{ \xi = (\xi_1, ..., \xi_n) | \xi_i \geq 0, \sum_{i=1}^n \xi_i = 1 \}.$$ 

Let $U_\epsilon$ be any set $U_\epsilon \subseteq P_n$ having the following property: for any pair of vectors $\xi$ and $\eta$ in $U_\epsilon$,

$$\sum_{i=1}^n | \xi_i - \eta_i | \geq \epsilon$$

($\epsilon$ is a given positive real number).
LEMMA 5.4 (Rabin). If \( P_n \) and \( U_\varepsilon \) are as above, then \( U_\varepsilon \) is a finite set containing at most \( k(\varepsilon) \) elements, where

\[
k(\varepsilon) = \left(1 + \frac{2}{\varepsilon}\right)^{n-1} \quad \text{for} \quad n > 1.
\]

**Proof.** The proof of this lemma is implicit in the proof of Theorem 3 in [3].

**Remark.** The bound above is not sharp and from a practical point of view, it would be desirable to have a sharp bound. This is, however, an open problem.

We are now able to prove the following:

**THEOREM 5.5.** Given a p.a. \( A \) and \( \varepsilon \), the relation \( \bar{P}_\varepsilon \) induced by \( A \) is of finite index \( k \) with

\[
k \leq \left(1 + \frac{1}{\varepsilon}\right)^{n-1}
\]

(\( n \) here is the number of states of the p.a. \( A \) inducing the relation \( \bar{P}_\varepsilon \)).

**Proof.** Let \( x_1 \cdots x_k \) be a set of tapes which are pairwise nonrelatives by \( \bar{P}_\varepsilon \). This implies that for every \( 1 \leq i < j \leq k \) there is a tape \( y \) such that

\[
| p(x_i, y) - p(x_j, y) | \geq \varepsilon
\]

or

\[
| \pi(x_i, y) - \pi(x_j, y) | = | (\pi(x_i) - \pi(x_j)) \eta(y) | \geq \varepsilon.
\]

Now

\[
\sum_{t=1}^{n} (\pi_t(x_i) - \pi_t(x_j)) = \sum_{t=1}^{n} \pi_t(x_i) - \sum_{t=1}^{n} \pi_t(x_j) = 1 - 1 = 0.
\]

Therefore

\[
\Sigma_t^+(\pi_t(x_i) - \pi_t(x_j)) = -\Sigma_t^-(\pi_t(x_i) - \pi_t(x_j)),
\]

where \( \Sigma_t^+ \) and \( \Sigma_t^- \) are summations over indices \( t \) for which \( \pi_t(x_i) - \pi_t(x_j) \) is non-negative or negative, respectively.

Combining (5.1) and (5.2) we have that

\[
\varepsilon \leq | \Sigma_t^+(\pi_t(x_i) - \pi_t(x_j)) \eta_t(y) | = | \Sigma_t^+(\pi_t(x_i) - \pi_t(x_j)) \max_t \eta_t(y) + \Sigma_t^-(\pi_t(x_i) - \pi_t(x_j)) \min_t \eta_t(y) | \leq \Sigma_t^+(\pi_t(x_i) - \pi_t(x_j)) (\max_t \eta_t(y) - \min_t \eta_t(y))
\]

where \( n \) here is the number of states of the p.a. \( A \) inducing the relation \( \bar{P}_\varepsilon \).
using again (5.2) we have

\[ 2\epsilon \leq \sum_{i=1}^{n} |\pi_i(x_i) - \pi_i(x_j)| \]

and this inequality implies by Rabin's lemma that the set of vectors

\[ \pi(x_1), \pi(x_2) \cdots \pi(x_k) \]

is finite with

\[ k \leq \left(1 + \frac{1}{\epsilon}\right)^{n-1} \].

6. Fuzzy Star Acceptors and Probabilistic Acceptors

**Definition 6.1.** Let \( A \) be an f.s. ac., \( B \) an f. ac., \( T(A) \) and \( T(B) \) the languages defined by \( A \) and \( B \), respectively. \( T(B) \) \( \epsilon \)-approximate \( T(A) \) if

\[ (T(B) \setminus T(A)) \cup (T(B) \setminus T(A)) \subseteq \{x | x \in \Sigma^*, |\psi(x)| \leq \epsilon\} \]

where \( T(A) \) and \( T(B) \) are the compliments of \( T(A) \) and \( T(B) \), respectively, in \( \Sigma^* \).

**Proposition 6.1.** Let \( \psi \) be an f.s.f. and \( B \) an f.a. If \( B \) \( \epsilon \)-approximates \( \psi \) then, for any \( \lambda \), \( B \) can be transformed into an f. ac. which \( \epsilon \)-approximates the f.s. ac. \( \psi, \lambda \).

**Proof.** Let the final state of \( B \) be the states such that \( \phi(s_i) > \lambda \). If \( x \in T(B) \) then \( B(s_0, x) = s_i \) with \( \phi(s_i) > \lambda \) and, because

\[ |\psi(x) - \phi(s_i)| \leq \epsilon, \]

we have that

\[ \psi(x) > \lambda - \epsilon. \]

If \( x \notin T(B) \) then \( B(s_0, x) = s_j \) with \( \phi(s_j) \leq \lambda \) and therefore \( \psi(x) \leq \lambda + \epsilon. \) Now \( x \in T(A) \) implies that \( \psi(x) > \lambda \) and \( x \in T(B) \) implies that \( \psi(x) \leq \lambda \) and the result follows.

The following proposition is a converse of Proposition 6.1 and is related to Theorem 5 in Rabin [3].

**Proposition 6.2.** Let \( \psi \) be a f.s.f. such that for any \( \lambda \), \( (\psi, \lambda) \) is an f.s. ac. which is \( \epsilon \)-approximable by some f. ac. \( B \). Then there is a f.a. \( B \) which \( 2 \epsilon \)-approximates \( \psi \).
Proof. Divide the interval \([0, 1]\) into \(k\) equal parts by \(k - 1\) points \(\lambda_0, \lambda_2, \ldots, \lambda_{k-1}\) \((\lambda_0 = 0, \lambda_k = 1)\) such that \(\lambda_i - \lambda_{i-1} \leq \epsilon\) \(\forall i = 1, 2, \ldots, k\), and let \(B_{\lambda_i}, i = 0, 1, \ldots, k - 1\) be the corresponding \(\epsilon\)-approximating acceptors for \((\psi, \lambda_i)\). Define the machine \(B\) as follows. \(B = (S, s_0, M)\) with

\[
S = \{(s_{i_1}(\lambda_0), s_{i_2}(\lambda_1), \ldots, s_{i_k}(\lambda_{k-1})) \mid s_{i_k}(\lambda_k) \in S_{i_k}\},
\]

\[
s_0 = (s_0(\lambda_0), s_0(\lambda_1), \ldots, s_0(\lambda_{k-1})),
\]

\[
M((s_{i_1}(\lambda_0), s_{i_2}(\lambda_1), \ldots, s_{i_k}(\lambda_{k-1})), \sigma) = (M_{\lambda_1}(s_{i_1}, \sigma), M_{\lambda_2}(s_{i_2}, \sigma), \ldots, M_{\lambda_k}(s_{i_k}, \sigma)),
\]

with

\[
B_{\lambda_i} = (S_{\lambda_i}, s_0(\lambda_i), M_{\lambda_i}, F_{\lambda_i})
\]

Set

\[
\varphi(s) = \varphi((s_{i_1}(\lambda_0), s_{i_2}(\lambda_1), \ldots, s_{i_k}(\lambda_{k-1})) = \max_j \{\lambda_j \mid s_{i_j-1}(\lambda_j) \in F_{\lambda_j}\}
\]

so that

\[
\varphi(B(s_0, x)) = \lambda_j \quad \text{implies that} \quad x \in T(B_{\lambda_j}) \quad \text{and} \quad x \notin T(B_{\lambda_{j+1}}).
\]

then

\[
\varphi(B(s_0, x)) = \lambda_j
\]

implies that

\[
\psi(x) > \lambda_j - \epsilon
\]

and

\[
\psi(x) \leq \lambda_{j+1} + \epsilon \leq \lambda_j + 3\epsilon.
\]

Thus

\[
|\varphi(B(s_0, x)) - \psi(x)| \leq 2\epsilon.
\]

Q.E.D.

Remark. It follows from the above propositions that \(\epsilon\)-approximation of an f.s.f. by an f.a. is possible if, and only if, \(\epsilon\)-approximation of the f.s. ac. derived from that function, with any given \(\lambda\), is possible.

We shall however show further, by an example, that there are f.s.f.'s which are not \(\epsilon\)-approximable by f.a.'s but the derived f.s. ac. with some \(\lambda\) is \(\epsilon\)-approximable by a f. ac.

Corollary 6.3. The p. ac.'s are \(\epsilon\)-approximable by Turing acceptors.

Proof. By Corollary 2.2 and Proposition 6.1.
Remark. It was shown by Rabin ([3], Theorem 2; see also [4], Theorem 12) that the set of p. ac. is not a countable set (even if the transition matrices have only rational entries). But the set of Turing machines is countable.

7. CHARACTERIZATION AND THE $R_\epsilon$-RELATION

DEFINITION 7.1. Given an f.s. ac. $(\psi, \lambda)$ and $\epsilon > 0$, an $\epsilon$-cover induced by $(\psi, \lambda)$ is a finite set $\{C_i\}_{i=0}^k$ where the $C_i$ are sets of points in the interval $[0, 1]$ satisfying the following requirements:

1. $\bigcup_{i=0}^k C_i = \{\xi \mid \psi(x) = \xi, x \in \Sigma^*\}$;
2. either $C_i \subseteq \{\xi \mid \xi \geq \lambda - \epsilon\}$
   or $C_i \subseteq \{\xi \mid \xi \leq \lambda + \epsilon\}, \quad i = 0, 1 \cdots k$;
3. for any $i$ and $x$ there is $j$ such that $C_i x \subseteq C_j$.

($C_i x$ is defined as in Definition 3.1.)

THEOREM 7.1. Given an f.s. ac. $(\psi, \lambda)$ and $\epsilon > 0$ $(\psi, \lambda)$ is $\epsilon$-approximable by an f. ac. iff there is a $\epsilon$-cover induced by the f.s. ac. $(\psi, \lambda)$. If there exists an $\epsilon$-cover induced by $\psi$ then there exists an $\epsilon$-cover induced by $(\psi, \lambda)$.

The proof of the first assertion of the theorem which is similar to the proof of Theorem 3.1, is omitted. The machine $B$ here will be defined as in Theorem 3.1 and the final states of $B$ will be those $C_i$ which satisfy the relation $C_i \subseteq \{\xi \mid \xi \geq \lambda - \epsilon\}$.

It is immediate that any $\epsilon$-cover satisfying the conditions of Definition 3.1 also satisfies the conditions of Definition 7.1.

DEFINITION 7.2. The relation $R_\epsilon$, over $\Sigma^*$ induced by an f.s. ac. $(\psi, \lambda)$ is defined as follows.

$x R_\epsilon y \triangleq \text{def} (z) [\psi(xz) \cdot \lambda \cdot \epsilon \text{ and } \psi(yz) \cdot \lambda \cdot \epsilon]
= [\psi(xz) > \lambda \Rightarrow \psi(yz) > \lambda].$

It is easily seen that the relation $P_\epsilon$ defined in Lemma 5.1 is a refinement of the relation $R_\epsilon$ here and therefore $P_\epsilon$, of finite index implies that $R_\epsilon$ is of finite index. We have thus the following:
Corollary 7.2. The relation $R_e$ induced by a p. ac. is of finite index.

Proof. By Theorem 5.5 and the above remark.

Corollary 7.3. If the relation $R_e$ is induced by an f.s. ac. such that the corresponding f.s.f. is $\epsilon$-approximable, then $R_{ae}$ is of finite index.

Proof. By Proposition 5.3 and the above remark.

Using the same kind of reasoning as that used in the proof of Proposition 5.3 we can also prove the following.

Proposition 7.4. Given an f.s. ac. $(\psi, \lambda)$ and $\epsilon > 0$, if $(\psi, \lambda)$ is $\epsilon$-approximable by an f. ac., then the induced relation $R_e$ is of finite index.

Finally consider the following.

Definition 7.3. An f.s. ac. $(\psi, \lambda)$ is quasi-definite if for any $\epsilon$ there is $k(\epsilon)$ such that for any $x$ with $l(x) \geq k(\epsilon)$ the following is true for any $y \in \Sigma^*$:

\[
\psi(x) > \lambda \quad (\psi(x) \in T(\psi, \lambda)) = \quad > \psi(yx) > \lambda - \epsilon;
\]

\[
\psi(x) \leq \lambda \quad (\psi(x) \notin T(\psi, \lambda)) = \quad > \psi(yx) \leq \lambda + \epsilon.
\]

It is clear that, if $\psi$ is a quasi-definite f.s.f., then $(\psi, \lambda)$ is a quasi-definite f.s. ac. (The converse is, however, not true; e.g., suppose the matrices of a p.a. are

\[
A(a) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad A(b) = \begin{pmatrix} 3 & 1 \\ 4 & 3 \end{pmatrix}
\]

and $\psi(x)$ is the $(1, 1)$ entry in $A(x)$. Then with $\lambda = \frac{1}{3}$, $T((\psi, \lambda)) = \Sigma^*$, and therefore $(\psi, \lambda)$ is a quasi-definite acceptor. But, for any $n$,

\[
| \psi(ba^n) - \psi(a^n) | = \frac{1}{4},
\]

so that $\psi$ is not quasi-definite.)

We have therefore the following.

Corollary 7.5. If $\psi$ is a quasi-definite f.s.f. then for any $\lambda$ and $\epsilon, (\psi, \lambda)$ is an $\epsilon$-approximable f.s. ac. and the approximating acceptor may be choosen to be a definite acceptor.

Proof. By the remark after Theorem 4.1 and the above remark.

The above corollary follows also from the following theorem which can be proved in the same manner as Theorem 4.1.
THEOREM 7.7. Given a quasi-definite acceptor \((\psi, \lambda)\) and \(\epsilon, (\psi, \lambda)\) is \(\epsilon\)-approximable by a definite finite acceptor.

The proof is omitted.

We remark here that in the previously-mentioned theorem of Rabin ([3], Theorem 2; see also [4], Theorem 12) the f.s. acceptors are in fact quasi-definite so that the set of quasi-definite acceptors is not a countable set.

8. A COUNTEREXAMPLE

The following counterexample is based on an example of H. Kesten and on an idea of R. E. Stearns (private communications). The author is indebted to both of them for allowing him to use their examples.

Consider the following p.a. \(A = (S, \pi, \{A(\sigma)\}, F)\) over \(\Sigma = \{0, 1\}\) with \(S = \{s_0, s_1, s_2, s_3\}\), \(\pi = (1000)\)

\[
\eta^F = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\]

and

\[
A(0) = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A(1) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.
\]

By straightforward computations one can prove the following relations:

\[
p(x) = \begin{cases} \left(\frac{1}{2}\right)^n & \text{if } x = 0^n \quad n = 0, 1, 2, ..., \quad (0^0 = A) \\ \frac{1}{2} & \text{if } x = 0^n10^n1, 0^n11, \quad n_j \geq 0, \quad j = 1, 2, ..., k \\ > \frac{1}{2} & \text{if } x = 0^n10^n1, 0^n11, \quad n_j > 0, \quad j = 1, 2, ..., k \\ < \frac{1}{2} & \text{if } x = 0^n10^n1, 0^n10^n11, \quad n_j \geq 0, \quad j = 1, 2, ..., k \\ n_{k+1} > 0, \end{cases}
\]

where \(p(x)\) is the \((1, 1)\) entry in \(A(x)\).

Consider now the f.s.f. defined by \(A, P_A\), and let \((P_A, \lambda)\) be the p. ac. with \(\lambda = \frac{1}{2}\).

We have that

\[
T((P_A, \lambda)) = \{x \mid P_A(x) > \frac{1}{2}\}.
\]

It follows from the above inequalities that \(T((P_A, \lambda))\) for \(\lambda = \frac{1}{2}\) is the set of tapes \(x\) such that \(x = A\) or \(x\) begins with a zero, ends with a one, and contains no subtape of two or more consecutive ones. It is easily verified that this set of tapes is a regular set (there exists an f. ac. accepting it) and therefore it is \(\epsilon\)-approximable (even for \(\epsilon = 0\)) by an f. ac.
We shall show that there is \( \lambda \) such that \((P_A, \lambda)\) is not \(\varepsilon\)-approximable by an f.a. for that \( \lambda \) with the result that \( P_A \) is a f.s.f. which is not approximable by an np. a. (This will follow from Proposition 6.1.)

Let \( x_n^m \) be the word \( x_n^m = (0^n1)^m \). One can prove again, using straightforward computation, that:

\[
p(x_n^m) = \frac{1 + \left[1 - \left(\frac{1}{2}\right)^n\right]^m}{2}.
\]

Thus \( \lim_{n \to \infty} p(x_n^m) = 1 \) for fixed \( m > 0 \), while \( \lim_{m \to \infty} p(x_n^m) = \frac{1}{2} \) for fixed \( n > 0 \). Now let \( \lambda \) be a real number \( \frac{1}{2} < \lambda < 1 \), say \( \lambda = \frac{3}{4} \), and let \( \epsilon \) be a real number \( 0 < \epsilon \) and suppose that \((P_A, \lambda)\) is \(\epsilon\)-approximable for the given \( \lambda \) and \( \epsilon \). Let the approximating machine have \( k \) states. Choose \( n_0 \) so great that

\[
p(x_{n_0}^m) > \lambda + \epsilon \quad \text{for} \quad m = 1, 2, \ldots, k + 1.
\]

The first \( k + 1 \) applications of the input sequence \( x \) must send the approximating machine \( B \) through a sequence of states \( s_0, s_1, \ldots, s_{k+1} \), which are all final states of \( B \). But \( B \) has only \( k \) states so that \( s_{k+1} = s_i \) for some \( i < k + 1 \) so that all the tapes of the form \( x_{n_0}^m \), \( m = 1, 2, \ldots \) will be in \( T(B) \). Thus \( B \) cannot \(\epsilon\)-approximate \( P_A \), for there is \( m_0 \) with \( p(x_{n_0}^{m_0}) < \lambda - \epsilon \), i.e.,

\[
| p(x_{n_0}^{m_0}) - \lambda | > \epsilon,
\]

while \( x_{n_0}^{m_0} \in T(B) \) and \( x_{n_0}^{m_0} \notin T(A) \). The following are direct consequences of the above example.

1. There is an p. ac. which is not approximable by an f. ac.
2. There is an f.s.f. which is not approximable by an f.a. (This follows from our example and Proposition 6.1).
3. There is an p.a. which is not approximable by an f.a., but the acceptor defined by the p.a. with some \( \lambda \) (\( \lambda = \frac{1}{2} \) in our example) is approximable by an f.a. The two concepts of approximation are therefore not equivalent.
4. The class of p.a.'s is stronger than the class of f.a.'s, this being a consequence of the intrinsic nature of the probabilistic automata and not of the actual properties of the cut point \( \lambda \).
5. There is an f.s.f. and \( \epsilon \) such that there is no \(\epsilon\)-cover induced by this f.s.f. (See Definition 3.1.)
6. There is a f.s. ac. and \( \epsilon \) such that there is no \(\epsilon\)-cover induced by this f.s. ac. (See Definition 7.1.)
We have shown in this paper that f.s.f.'s which are computable (e.g., p.a.'s) are approximable by Turing automata but there are f.s.f.'s defined by p.a.'s (previous example) which are not approximable by f.a.'s.

We are thus faced with the following open problems:

(1) Characterize the f.s.f.'s (the p.a.'s) which are approximable by
   (a) linear bounded automata,
   (b) push-down automata,
   (c) sequential automata;

(2) What is the most powerful class of np. devices which suffices for approximating the p.a.'s? Our paper shows that Turing automata are enough but f.a.'s are not. We believe that push-down automata are not enough either, but we have no proof for this.

We have shown (Section 5) that the condition that the relation $\mathcal{P}_r$ (Lemma 5.1) induced by a given f.s.f. $\psi$ is of finite index, is a necessary condition for the function $\psi$ to be approximable by an f.a. This condition is however not sufficient for, by Theorem 5.5, that relation is of finite index for any p.a., but the previous example is a p.a. which is not approximable by an f.a.

The previous example shows also that there is a p.a., which is not approximable by f.a.'s, but the acceptor defined by that p.a. with some $\lambda$ is approximable by an f. acceptor. The two concepts are therefore not identical and the above remarks and open problems may be stated and posed also for the case of approximating fuzzy star acceptors separately.

Finally, it would be interesting to try to extend the above result to input-output fuzzy star functions whether of machine type (as introduced by Carlyle [11]) or general as studied by Arbib [12] or Ott [13]. In connection with this problem it seems more natural to compare input-output fuzzy star functions with two-tape finite automata when considering the Problem (3) and (4) in the introduction but we shall not undertake this further study here.

Acknowledgments

The author wishes to express his gratitude to the Electrical Engineering Department of the University of California, Berkeley, for their kind invitation which gives him the opportunity to be associated with this department for the academic year 1966-67.

Thanks are also due to Professors A. Gill, L. Haines, M. Harrison, and L. Zadeh, with whom he had several fruitful discussions, and to Professors H. Kesten and R. E. Stearns for allowing him to use their examples in Section 8 of the paper.
REFERENCES