A Bibliography on Fuzzy Automata,
Grammars and Languages

Peter R.J. Asveld

Department of Computer Science, Twente University of Technology
P.O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: infprja@cs.utwente.nl

In a recent attempt to formalize a simple aspect of the notion “robustness in parsing” [7, 8, 9, 10], I needed the concept of fuzzy context-free grammar. So the question “What is known with respect to fuzzy analogues of automata, grammars and formal languages?” arose naturally. Scanning the titles of several thousands papers on fuzzy sets, fuzzy logic and their applications (in almost every area) yielded about a 160 articles on this fuzzy counterpart of formal language theory.

The result of this scanning process is listed below: the list consists of references on fuzzy formal languages, the generation of fuzzy languages by means of fuzzy grammars, the recognition of fuzzy languages by fuzzy automata and machines, as well as some applications of fuzzy set theory to syntactic pattern recognition, linguistics and natural language processing.

Fuzzy grammars should not be confused with probabilistic grammars, although both formalisms are instances of the more general notion of so-called weighted grammars [108, 77, 129], i.e., grammars in which the production rules have been provided with weights. Roughly, the differences between fuzzy and probabilistic grammars are as follows.

A fuzzy grammar generates a fuzzy language, i.e., a language for which the boundary between in and out the language is “soft” or rather gradually. So in this framework we are able to distinguish between strings that are more and less out of (the crisp part of) that language, e.g., between severe and small grammatical errors, respectively. On the other hand, a probabilistic grammar is an ordinary grammar in which the set of rules has been provided with a probability distribution in order to control the derivation process. One of the main objectives in this latter approach is to cope with ambiguity by means of the distinction between more and less probable derivations. Obviously, the difference between fuzzy and probabilistic grammars is reflected in the corresponding acceptors. A fuzzy automaton, machine or parser computes the degree of membership of the input string. Similarly, a probabilistic automaton, machine or parser determines the most probable computation or derivation, respectively, corresponding to the input.

Secondly, the underlying mathematical models differ: viz. probabilistic grammars are due to the constraint that the sum of the probabilities for all rules with the same left-hand side should be equal to 1. On the other hand there is no such normalization constraint for fuzzy grammars. Usually, weights are real numbers in between 0 and 1. Since fuzzy grammars can be defined in terms of fuzzy sets, the set of possible weights in fuzzy grammars (i.e., the interval [0, 1]) may be replaced by a much more general structure (e.g., a distributive complete lattice [34]). Such a generalization for probabilistic grammars is highly uncommon.
In defining a context-free language parsing or recognizing algorithm to be robust, the algorithm should be able to manage input strings that, on the one hand, are not in the language but, on the other hand, closely resemble a string in the language. One obvious way to describe those input strings is, of course, by means of a (generalized) fuzzy context-free grammar. This results in a framework in which correctly derived sentences as well as the effect of grammatical errors can be described in a uniform fashion; see [7, 8, 9, 10] for details.

The items in this bibliography are divided – according to their title – into the following categories:

- **A** automata, sequential machines, other accepting devices such as cellular automata and neural networks
- **B** (general) bibliography on fuzzy sets and their applications
- **F** (general aspects of) fuzzy formal languages
- **G** grammars, rewriting systems, other generating mechanisms
- **L** linguistics, natural language processing
- **P** (syntactic) pattern recognition
- **Z** miscellaneous

The last column in this table contains the number of papers put into this category; some papers are placed in more than a single category.

Earlier bibliographies included in the list below are those by de Kerf from 1975 [53], Gaines & Kohout (1977) [30], Kandel & Yager (1979) [50], Kaufmann (1980) [51], Kandel (1982) [48], Novák & Ramík (1991) [94], and by Klir & Yuan (1995) [59]. The differences with the bibliography below are obvious: the list below also contains some more recent contributions, but its scope is limited to the areas mentioned above.

My impression is that this bibliography is rather complete with respect to the categories [A], [F] and [G], less complete with respect to [L], whereas [P] just contains some papers I encountered occasionally during my search. Most references are rather old: the average age of these papers is about 17 years! In recent years the interest in fuzzy formal language theory seems to be rather limited and it slightly tends towards applications in linguistics; viz. the average ages in the categories [A + F + G] and [L] is 18 and 15½ years, respectively. Anyway, additions to this list – consisting of either old or, especially, recent contributions – are very welcome.

---


46. J. Kacprzyk & C. Iwański: Fuzzy logic with linguistic quantifiers in inductive learning, pp. 29-38 in [95]. [L]


64. S. Lakshmivarahan & K.S. Rajaseethupathy: Considerations for fuzzifying formal languages and synthesis for fuzzy games, J. Cybern. 8 (1978) 83-100. [F, Z]
91. V. Novák: On the logical basis of approximate reasoning, pp. 17-27 in [95]. [L]


