Journal of Automata, Languages and Combinatorics **27** (2022) 1–3, 79 – 90

ROUGH-SET-LIKE APPROXIMATION SPACES FOR FORMAL LANGUAGES

Péter Battyányi Tamás Mihálydeák György Vaszil

Faculty of Informatics, University of Debrecen
Kassai út 26, 4028 Debrecen, Hungary
{battyanyi.peter,vaszil.gyorgy}@inf.unideb.hu
mihalydeak@unideb.hu

ABSTRACT

We define rough-set-like approximation spaces for formal languages based on similarity relations which are defined over the alphabet symbols. Our approach is motivated by situations when some uncertainty is present in our knowledge about the exact characters making up a text which need to be processed by some formal system. We define the lower and upper approximations of languages and consider the regular and context-free cases. We present characterizations of the approximations of languages accepted by deterministic finite automata or generated by context-free grammars.

Keywords: language approximation, rough language, rough finite state automaton, rough grammar

1. Introduction

Pawlak's original theory of rough sets (see in, e. g., [6, 8, 7]) is concerned with different approximating systems (such as, for example covering systems relying on tolerance relations [13], general covering systems [16, 9], decision-theoretic rough set theory [15], general partial approximation spaces [3], or similarity based approximation spaces [5]), but there is a very important common property. All these systems take into consideration that our background knowledge about the elements of the universe is limited in the sense that all we can know about a set or about its members are its lower and upper approximations. Our limited background knowledge is represented by a system of base sets: the members of a given base set have to be treated in the same way, they are indiscernible from each other.

Generally, a formal language is a set of words over a given alphabet, so it is natural to ask whether it is possible to approximate these sets in a similar way. In [10], context-free languages are approximated by regular languages, while the authors of [2, 12, 11, 1] embed different versions of roughness in the definition of automata or the definition of generative grammars by allowing uncertainty in these devices