Formal Languages and Parsing

CS 462

Jeffrey Shallit

LaTeXed by Sibelius Peng

Preface

Disclaimer Much of the information on this set of notes is transcribed directly/indirectly from the lectures of CS 462 during Winter 2022 as well as other related resources. I do not make any warranties about the completeness, reliability and accuracy of this set of notes. Use at your own risk.

For any questions, send me an email via https://notes.sibeliusp.com/contact.

You can find my notes for other courses on https://notes.sibeliusp.com/.

Sibelius Peng

Contents

Preface **1**

1 CS 462 notation **3**

1

CS 462 notation

- Natural numbers $\mathbb{N} = \{0, 1, 2, \ldots\}$ and we use letters $i, j, k, \ell, m, n \in \mathbb{N}$.
- Finite string/word: a map from $[0, n-1]$ (an interval) to Σ (a finite alphabet of symbols)

  $w[i]$ is ith symbol of w
- infinite strings/words: a map from $\mathbb{N}$ to Σ. We denote infinite strings by bold-face:

$$\mathbf{w} = \mathbf{w}[0]\mathbf{w}[1]\mathbf{w}[2]\cdots$$

- Σ^* is the set of all finite words over Σ.
- Σ^ω is the set of all infinite words over Σ. Also written $\Sigma^{\mathbb{N}}$.
- $\Sigma^\infty = \Sigma^* \cup \Sigma^{\mathbb{N}}$.

Finite words typically denote by s, t, u, v, w, x, y, z

Some refreshers from CS 360/365:

- x is a **prefix** of z if there exists y such that $z = xy$
- x is a **suffix** of z if there exists y such that $z = yx$
-