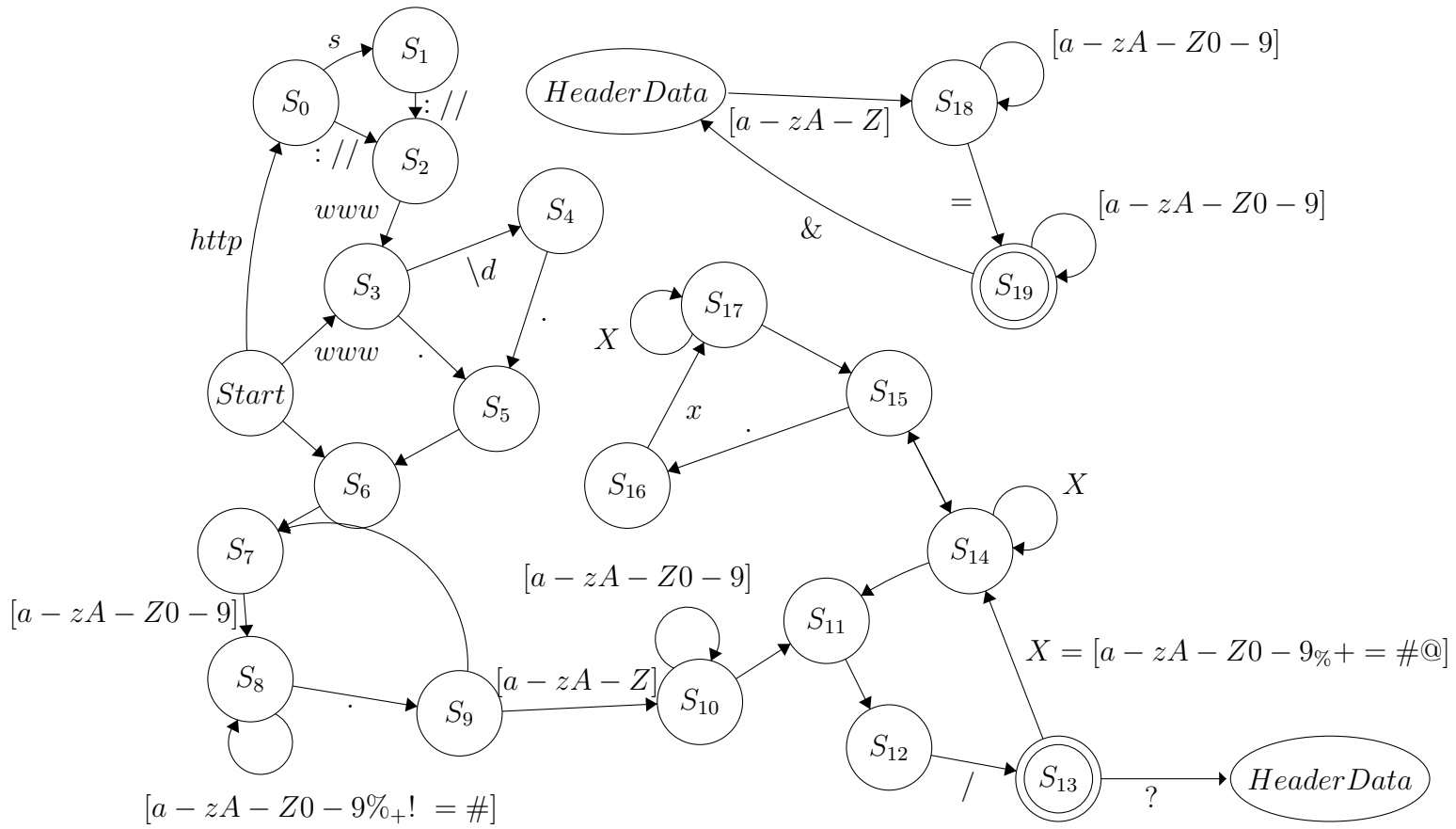


Automata of Regexes Used in the Tokenizer

Kasra Eskandari

955361005

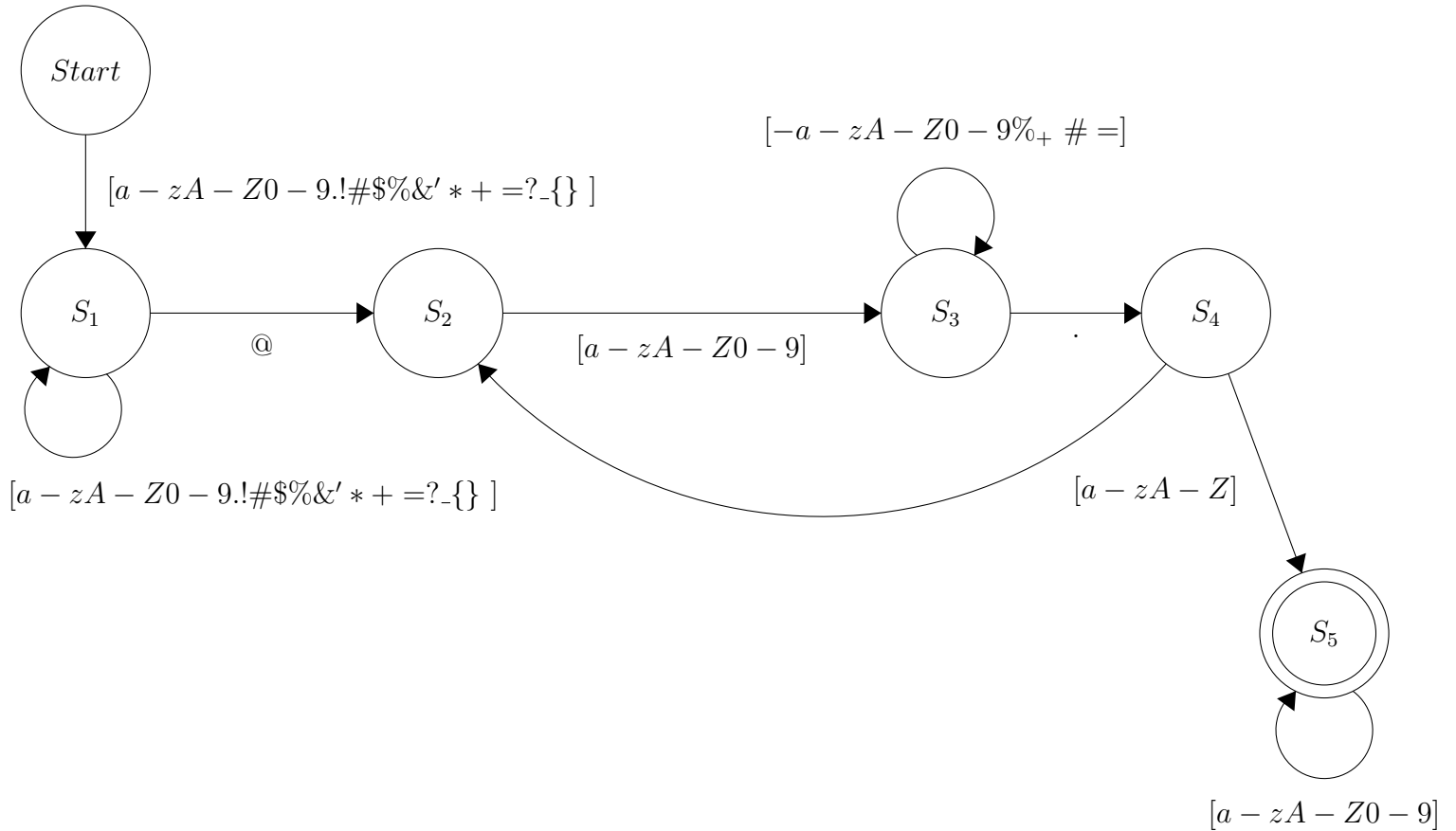
0.1 LINK



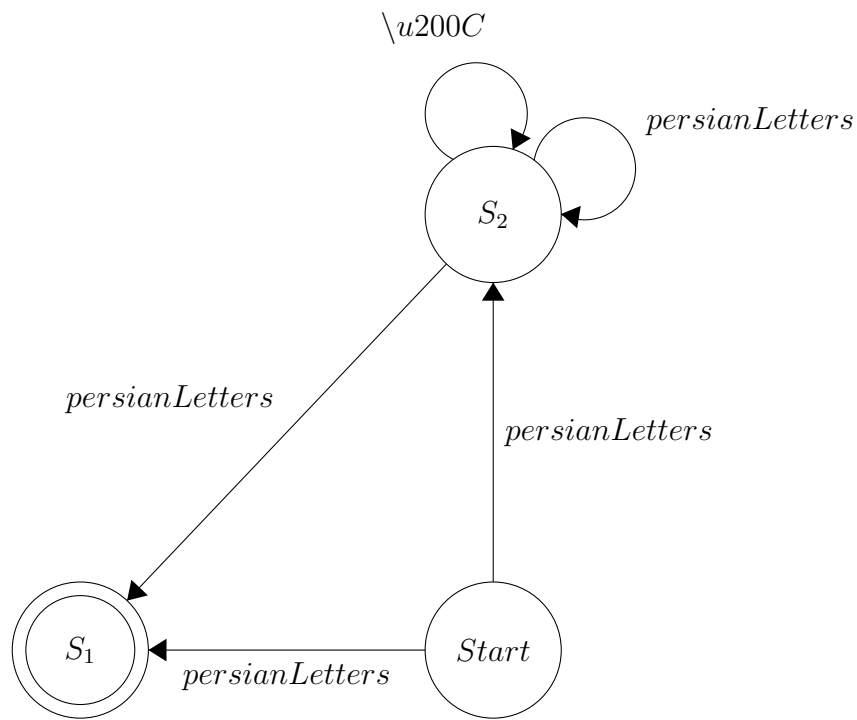
Also, mind this assignment:

$$X = [a-zA-Z0-9\%+ = \#@]$$

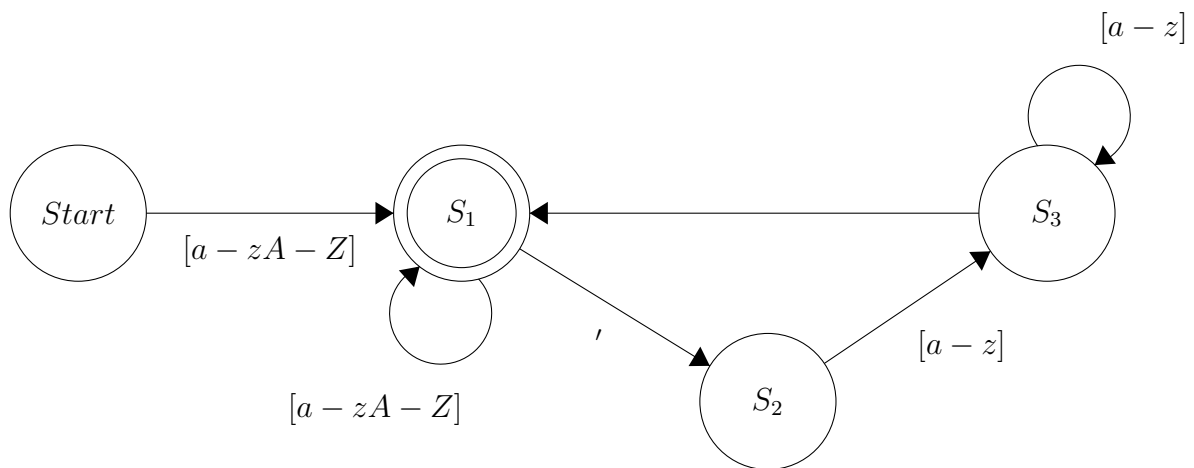
0.2 EMAIL



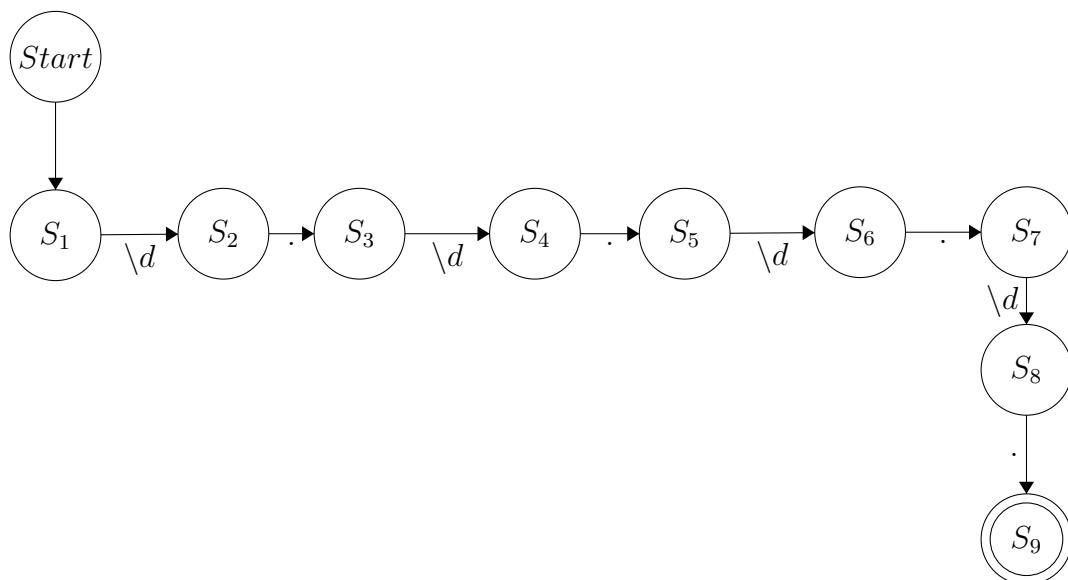
0.3 Persian Words



0.4 English Words



0.5 IP



0.6 English Human Name

