

5 Formal Models of Language (pjb48)

An intrepid linguist encounters an alien while exploring a remote planet, she documents an alien language that has four sounds, $\Sigma = \{a, e, f, m\}$. The following are examples of the alien's speech and a translation:

$e\ amaf$	nice spaceship
$e\ faem$	nice spaceboots
$e\ e\ amaf$	very nice spaceship
$e\ e\ e\ faem$	extremely nice spaceboots

The linguist hypothesises that Expression 1 matches sound sequences in the alien language.

Expression 1: $e^*(a|e|f|m)(a|e|f|m)^*$

- (a) Describe the strings that are matched by Expression 1. [2 marks]
- (b) Provide a Finite State Automaton that can generate the language defined by Expression 1. [2 marks]

As the linguist continues to explore, she encounters a grumpy alien and documents the following utterances:

$fama\ e\ amaf$	not nice spaceship (unpleasant spaceship)
$meaf\ e\ e\ faem$	very unpleasant spaceboots
$afaf\ e\ e\ e\ fafa$	extremely unpleasant gift

The linguist hypothesises that Expression 2 matches sound sequences in the language.

Expression 2: $w^{-1}e^*w$ where $w \in \Sigma^*$

- (c) Can the linguist draw a Finite State Automaton to generate the language defined by Expression 2? Provide a proof for your answer. [5 marks]
- (d) Provide a grammar that can generate the language defined by Expression 2. [5 marks]

On the far side of the planet, the linguist encounters a new dialect and documents the following utterances:

$amaf\ e\ amaf$	unpleasant spaceship
$faem\ e\ e\ faem$	very unpleasant spaceboots
$mefaem\ e\ e\ e\ mefaem$	extremely unpleasant earthling

- (e) Provide a general expression to match such sound sequences. [1 mark]
- (f) Can you use a Context Free Grammar to generate the language defined by your expression? Provide a proof for your answer. [5 marks]