

## Quiz 3

CS 753 2021

January 30, 2021

Submit your answers as a pdf. You can draw your FSTs on paper and submit a scan.

**(I) Spelling correction system:** A spelling correction system converts misspelled strings into valid spellings of words. For example, the input letter sequence “dac” could be mapped to the English words “dab” or “dad”. This problem deals with correcting spellings of isolated words using finite state machines.

(a) Let  $x$  be a string representing a misspelled word and  $A$  be an FSA for  $x$ . The misspellings are defined over the alphabet  $\{a, b, c, d, e\}$ . Design an unweighted edit distance transducer  $E$  such that  $B = A \circ E$ , on input  $x$ , produces  $y$  iff  $y$  is at an edit distance of at most 1 from  $x$  (where the edit operations are substitutions, insertions and deletions). **[3 points]**

(b) The composed machine  $A \circ E$  can generate spellings that do not necessarily correspond to valid English words. Design a finite-state machine  $T$  such that  $A \circ E \circ T$  only produces valid English words. You can assume the only valid English words come from the output alphabet  $V = \{\text{“bed”, “dad”, “bead”, “cede”, “dead”, “accede”}\}$ . **[3 points]**

(Extra credit for designing  $T$  with smallest number of states.)

**(II) Deterministic machine** Draw a deterministic FSA that accepts all binary strings over the alphabet  $\{0, 1\}$  that do not contain the substring 001. **[2 points]**