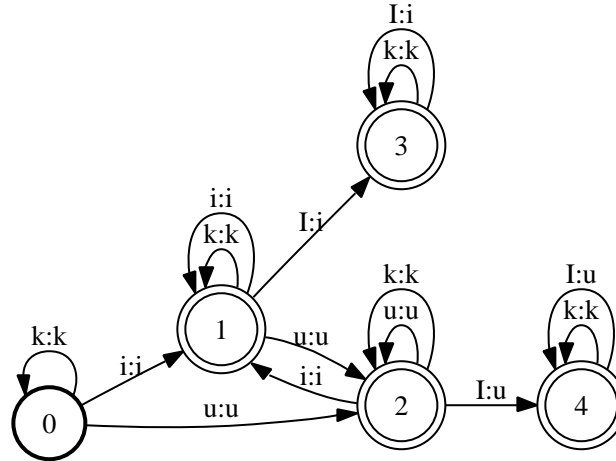


CMPT 413 - Spring 2011 - Midterm #1

Please write down "Midterm #1" on the top of the answer booklet.

When you have finished, return your answer booklet along with this question booklet.

- (1) (20pts) Consider the following finite-state transducer (FST) γ , where we assume that u and i are the only vowels in this particular natural language and that k is used as a placeholder for any consonant. I is an abstract vowel that appears in the input language and depending on the context is mapped to either u or i in the output language.



- a. The *input* language of the above FST can be written as the regular expression:

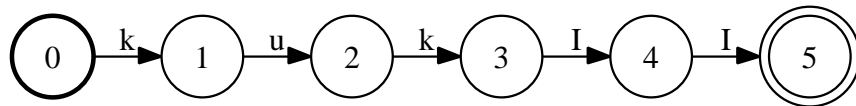
$(k^* (u|i))^+ k^* (k^* I)^* k^*$

Does the string *kiiukI* belong to the input language?

Answer: 1pt: Yes.

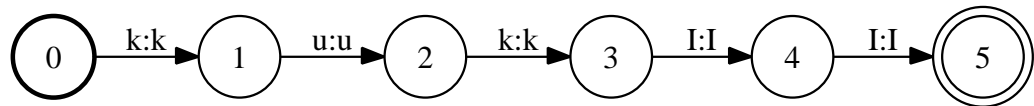
- b. Provide a finite-state machine (FSM) f that accepts the string *kukII* and nothing else.

Answer: 1pt:



- c. Id is a function that takes a FSM as input and produces a FST that accepts pairs of strings, where each string accepted by the FSM is paired with itself. Provide $Id(f)$.

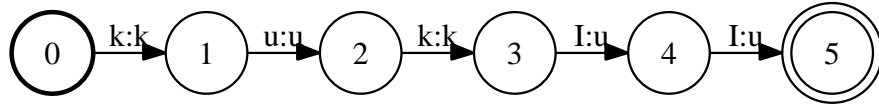
Answer: 2pt:



- d. Provide $Id(f) \circ \gamma$, the composition of $Id(f)$ with the FST γ given above.

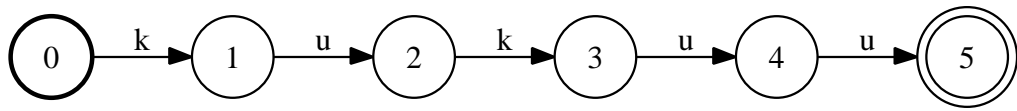
Answer: 6pt:

(0,0) (1,0) k:k
 (1,0) (2,2) u:u
 (2,2) (3,2) k:k
 (3,2) (4,4) I:u
 (4,4) (5,4) I:u



- e. Provide the FSM $\pi_2(Id(f) \circ \gamma)$, where π_2 projects the output language of a FST to provide an FSM.

Answer: 2pt:



- f. Provide an alternate name used to denote steps (1b) through (1e) that produced $\pi_2(Id(f) \circ \gamma)$.

Answer: 2pt: $\pi_2(Id(f) \circ \gamma) = \text{transduce}(\gamma, \text{kukII})$

- g. Provide two rewrite rules of the type $\alpha \rightarrow \beta / \lambda _ \rho$ that are together equivalent to the FST γ . Mention whether they are *obligatory* or *optional* and if they are to be applied *left to right* or *right to left*, and whether they are used *iteratively* or *simultaneously*.

Answer: 6pt: These two obligatory rewrite rules are left to right, iterative:

I \rightarrow i / i k* ____
 I \rightarrow u / u k* ____

- (2) (10pts) **Edit distance:** Assume insertion of a character has cost 1, deletion has cost 1, and substitution of one character for another has cost 2.

- a. What is the minimum edit distance value between target word *goal* and source word *hole*?

Answer:

4pts:

levenshtein distance = 4

- b. The following is a visual display of one possible alignment between target word *goal* and source word *hole*. Using the same visual display notation, provide all other alignments that have the same edit distance.

```

g o a l _
|   |
h o _ l e

```

Answer:

6pts:

levenshtein distance = 4

alignment number 1 for [4,4]:

```
_ g o a l _  
      |   |  
h _ o _ l e
```

alignment number 2 for [4,4]:

```
g _ o a l _  
      |   |  
_ h o _ l e
```

alignment number 3 for [4,4]:

```
g o a l _  
      |   |  
h o _ l e
```

total of 3 alignments