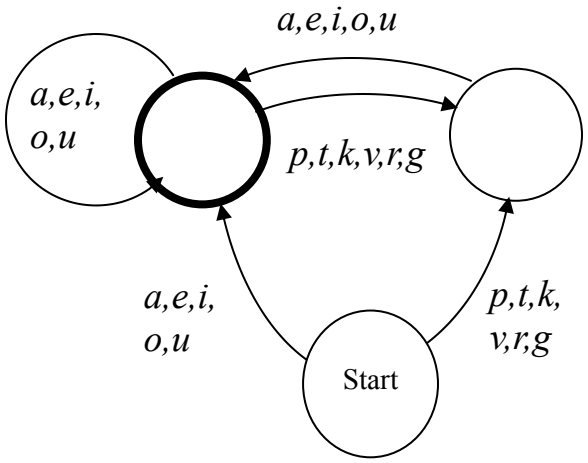
YOUR NAME: REGISTRATION #:

(15 points)

(I) aw-TOM-uh-tuh (1/2)

Finite-state automata (FSA) are a type of abstract "machine" with many possible uses. One possible use is to guess what language a document (such as a webpage) is in. If we make an automaton that can distinguish between possible English words and impossible ones, and then give it a webpage with a bunch of words that are impossible in English (like "*aioaepa*" or "*ragaiiare*"), we can be pretty sure that the webpage isn't written in English. (Or, at least, isn't *entirely* written in English.)

Here is a finite state automaton that can distinguish between possible and impossible words in Rotokas, a language spoken on the island of Bougainville off the coast of New Guinea. Rotokas has a very simple system of sounds and allows us to create a very small FSA.



An FSA works like a board game. Choose a word, and place your pencil on the space marked "Start". Going through the letters of the word one at a time, move your pencil along the path marked with that letter. If the word ends and you're at a space marked with a thicker circle, the word succeeds: it's a possible Rotokas word! If the word ends and you're not at a thicker circle, or you're midway through the word and there's no path corresponding to the next letter, the word fails: it's *not* a possible Rotokas word!

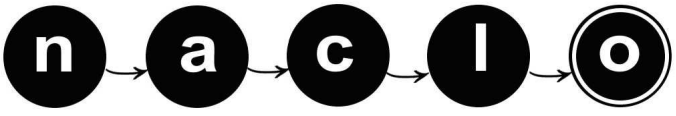
Try it out with these possible and impossible words; the automaton should accept all the possible words and reject the impossible ones.

Possible Rotokas words	
tauo	*kareveiepa*
puraveva	*ovokirovuia*
avaopa	*ouragaveva*

Impossible Rotokas words	
grio	*ouag*
ovgi	*vonoka*
gataap	*oappa*

I1. Now, using the automaton above, put a check mark next to each possible Rotokas word:

___ *iu*	___ *uente*	___ *voav*
___ *idau*	___ *urioo*	___ *uaia*
___ *oire*	___ *raorao*	___ *oratreopaveiepa*

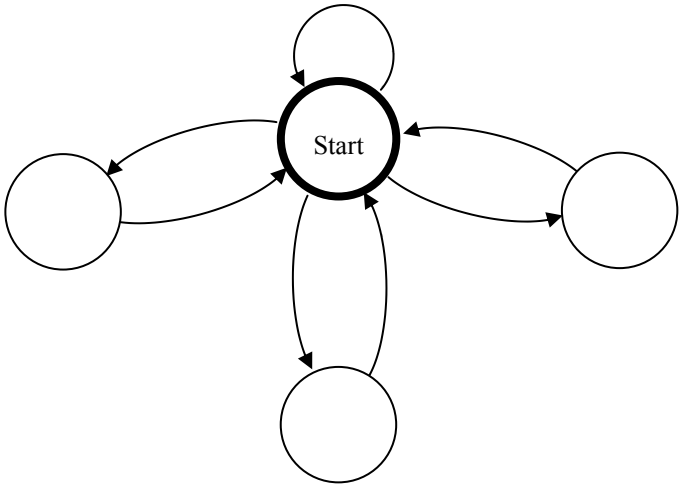


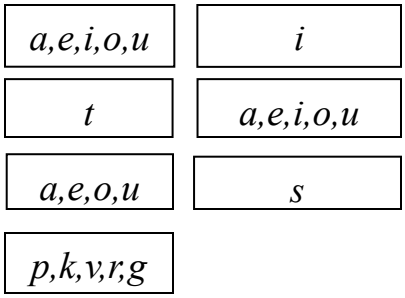
YOUR NAME: REGISTRATION #:

(I) aw-TOM-uh-tuh (2/2)

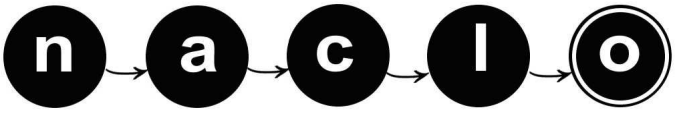
I2. Actually, the system that people on Bougainville use to write Rotokas is a little more complicated than the one we've presented here. In addition to the eleven letters above, the real Rotokas alphabet has a twelfth letter, S. This letter represents the sound "s" or "ts", which in Rotokas only occurs in very specific situations.

Below is a skeleton of an FSA for Rotokas with all the path labels removed and set to one side.





Replace the path labels so that exactly half of the words below succeed and the other half fail.

oisio	*tiravau*	*saiuu*	*kotoe*
uasau	*utsa*	*sioparoia*	*parauos*
puapuata	*sisigarue*	*porouativeve*	*aasiia*

I3. Why do T and S get their own paths? What is special about these letters?