

LING 106: Final

May 1-11, 2009

This is the final exam. Please read through *all* of the instructions on this page before proceeding. Failure to follow *any* instruction on this page will cost you **five points**. In part, because this is a formal course, and following precise directions is part of the material. And in part because these instructions aren't arbitrary; they're here to make the exam easier to grade, which means that, if you don't follow them, you've made my life harder. Don't annoy the guy in charge of your grade.

- The exam consists of 8 pages—including a lot of blank space—plus this page of instructions (“page 0”). Do not miss any pages.
- The exam is due at **noon on Monday, May 11**. No exceptions without a note from a dean, doctor, or clergyman. If you turn it in late, you will lose **ten points every hour**, rounded up.

That means that I must be in physical possession of your answers by that time. If send your answers via postal service, having the envelope postmarked by the deadline is not sufficient. If you're emailing answers, assuring me that you sent the answers by the deadline and you have no idea why I didn't get them is not sufficient. If you leave a page out of your email or hard copy, sending me the missing page at 3pm is not sufficient.

This is unforgiving, I know. However, it's the end of the semester, and I'm under a deadline here too; any exam that's late will add considerably to my stress levels (and see the last sentence of the first paragraph).

- Your answers should appear in the same order as the questions on the exam, and each problem (not each question!) should begin on a new page. You can achieve this by either...
 - Printing out the exam, writing your answers in the space provided, and turning in the exam pages. Use the back of the page if necessary, or attach another sheet of paper immediately after the page; or
 - Using other paper (typed, printed, handwritten on blank paper, lined paper, etc.), starting a new page where appropriate.

Either way, indicate at the top of each page which question is being answered. *Do not submit page 0 or the appendix as part of your answers*. If you do not answer the bonus question, you do not need to include that page.

- Your name should appear on each page of your answers.

You may use your notes and any notes from the class. You **may not** discuss the exam with classmates, or indeed with anyone else other than myself. Mail me if you have questions.

Good luck!

1. RECURSIVE NOTATION

10 points

Set S consists of the 38 different words that appear in the Preamble to the Constitution of the United States. For reference, those words are:

a	america	and	blessings	common	constitution
defense	do	domestic	establish	for	form
general	in	insure	justice	liberty	more
of	ordain	order	our	ourselves	people
perfect	posterity	promote	provide	secure	states
the	this	to	tranquility	union	united
we	welfare				

Set N is defined as follows:

- **to** $\in N$
- If $x \in N$ and $y \in S$, and the last letter of x is the first letter of y , then $y \in N$.
- Nothing else is in N .

Give N in list notation. (Note: $10 \leq |N| \leq 15$.)

2. THE MARS-EARTH COLONY**10 points**

Ever since the Martians made contact with Earth ten years ago, people have been discussing a joint Mars/Earth colony. One barrier to making it happen has always been the language—which Earth language would be used? Which Martian language? You’ve been corresponding in Martian for several years now with a Martian linguist named Gxlp_x, and it thinks it has an answer.

“What we need to do,” Gxlp_x explained in its last letter to you, “is use a Martian language with an Earth writing system. I thought your Latin alphabet would be good, and we should use Southern Icecap—it’s easy, because it only has five vowel sounds and four consonant sounds. Here’s what we should represent them with....”

Vowels: **a, e, i, o, u**

Consonants: **s, t, l, r**

“Here’s how Southern Icecap works. A word in the language consists of one or more syllables. And a syllable has to start with a consonant—and only one consonant!—which is followed by a vowel—and only one vowel! The syllable can have one more consonant at the end, though it doesn’t have to. It’s pretty easy.” It included a small chart...

Some syllables: **sa ru lat sor lol**

Some non-syllables: **a ur raa slo lost**

“Now, the other thing about Southern Icecap is that its speakers are very superstitious: they never mix ‘Phobos vowels’ and ‘Deimos vowels’—those are our moons, you know. Phobos vowels are the ones in their word for Phobos, **tuslo**; Deimos vowels are the ones in their word for Deimos, **telli**. The fifth vowel is a Sun vowel, because their word for the sun is **sal**, and can appear with either.” Once again, it provided a chart...

Some words: **sal tuslo telli roroto latlotlut latletlit selise**

Some non-words: **alala urtol raali rorati litlotlut latletlot solise**

“We’ll need to write a textbook, and of course, in order to truly make this a Mars/Earth collaboration, you and I should both contribute something to it. I have a computer program that generates strings from a finite state automaton (either deterministic or nondeterministic). All you need to do is give me a FSA—either the 5-tuple or a drawing—that models the Southern Icecap language.”

So that’s your task. In the interest of continued peace between Earth and Mars, you’ll want to do your best with this!

3. OLLA PODRIDA**20 points (5 points each)**

State whether each of the following statements is true or false. Provide explanations (except for problem 3.4, as noted).

3.1. $[[p \rightarrow q] \rightarrow [p \vee q]]$ is a tautology.

3.2. The function $f: C \rightarrow S$ is (a) one-to-one, (b) onto; where:
 C is the set of Visa credit card accounts in the United States,
 S is the set of sixteen-digit numbers, and
 $f(\langle x, y \rangle)$ if and only if y is the credit card number of x .

3.3. The encoding E may be monoalphabetic, given that
 $E(\text{help me}) = \text{gukf lu}$, and
 $E(\text{Mother's Day}) = \text{Lesxdh'r Tzo}$

3.4. Given the set $N = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}\}$:

- a. $\{\mathbf{c}\} \in N$
- b. $\{\mathbf{c}\} \subseteq N$
- c. $\emptyset \subseteq N$
- d. $\{\mathbf{c}\} \subseteq \wp(N)$
- e. $\emptyset \in \wp(N)$

(Explanations are not necessary for this problem. “True” or “False” will suffice.)

4. APPLIED PROPOSITIONAL LOGIC 32 points (8 pts each)

For each of the following arguments:

- (a) Express the premise or premises in propositional logic. (Don't forget any brackets!)
 (b) Using truth tables, determine the validity of the argument.

- 4.1** Either Jimmy Carter was the president in 1976, or Gerald Ford was a Democrat.
Gerald Ford was not a Democrat.
 Therefore, Jimmy Carter was the president in 1976.¹

(hint: **p** = "Carter was president in 1976"; **q** = "Gerald Ford was a Democrat")

- 4.2** Either Michele requests earmarks and she is corrupt, or she does not request earmarks.
If Michele signed a pledge against earmarks, then she does not request earmarks.
 Therefore, if Michele is not corrupt, then she signed a pledge against earmarks.²

(hint: **p** = "Michele requests earmarks", **q** = "Michele is corrupt", **r** = "Michele signed a pledge against earmarks". Don't forget to consider all *eight* possibilities!)

- 4.3** If global currency is imaginary and Michele argues against it, Michele is a kook.
 Global currency is not imaginary.
Michele argues against global currency.
 Therefore, Michele is not a kook.³

- 4.4** Either Michele is crazy or Michele is an idiot.
 If Michele is crazy, she shouldn't be a Congresswoman.
If Michele is an idiot, she shouldn't be a Congresswoman.
 Therefore, Michele shouldn't be a Congresswoman.⁴

¹ Argument courtesy of Michele Bachmann (R – MN): "I find it interesting that it was back in the 1970s that the swine flu broke out then under another Democrat president Jimmy Carter. And I'm not blaming this on President Obama, I just think it's an interesting coincidence." Note that the swine flu outbreak in the 1970s began in February 1976, several weeks before New Hampshire's Democratic primary and a month before Carter came in fourth in the Massachusetts primary.

² Argument courtesy of Michele Bachmann (R – Unuts): "I took a pledge in my own district. I have not taken earmarks in the last three years that I have been in Congress because the system is so corrupt. It's possible to make that pledge." In fact, Bachmann requested seven earmarks in 2008; and she apparently did sign Club for Growth's "No Earmark" pledge in 2008.

³ Argument courtesy of Glenn Beck (of Fox News) and Michele Bachmann (Y – Dopeoplevote4her). Beck: "Congresswoman, what happens is when you stand up and when you say those things, then you're deemed a kook." Bachmann: "But all we have to do is point to the treasury secretary on tape, on camera. This is not Michele Bachmann being a kook. This is our treasury secretary on tape and on camera."

⁴ Argument courtesy of me (D – PA).

5. HARRIS'S CONDITIONS ON MARS**20 points**

While you were working on your FSA for Gxlp, you get another message from it. Gxlp doesn't speak English, and as an experiment it's been trying to work out the syntax of English using the distribution of words, and you've been helping it. You've got a very large corpus of English texts that the two of you have been using.⁵

According to the message, Gxlp was looking over an early 20th century medical textbook—maybe it liked the pictures. At any rate, it says, “I've found a few contrasting phrases! Here...”

- (1) **This granulation tissue is gradually replaced by young cicatricial tissue**
- (2) **as the granulation tissue is slowly replaced by young cicatricial tissue**
- (3) **it must be brought into accurate contact with the new soil**
- (4) **the packing should if possible be brought into actual contact with the leaking point in the vessel**

“Based on (1) and (2), I'm pretty sure that **gradually** and **slowly** are the same category. And I think they're bimorphemic: **gradual-ly** and **slow-ly**. And then based on (3) and (4), I'm pretty sure that **accurate** and **actual** are the same category. And I think they're bimorphemic: **ac-curate** and **ac-tual**.”

You, of course, know better: only one of those pairs is bimorphemic. But you also know that, because it's trying to learn English on its own, Gxlp will get mad if you just tell it what the words mean; you need to use distributional methods.

Your task, then, is to explain to Gxlp which pair really does consist of bimorphemic words, *without referring to the meanings of any of the words*. Since the two of you have been using a very large corpus, you should feel free to “find” any words or sentences you need in there (i.e., make them up, and we'll assume you could find them).

⁵ You decided to leave *The War of the Worlds* out of this corpus. No need to give Gxlp ideas.

6. FINITE STATE MORPHOLOGY

15 points

In the lecture notes, I alluded to the fact that morphology, like phonology, could be modeled with finite state transducers. In this question, you will determine how. Consider the following four classes of English nouns.

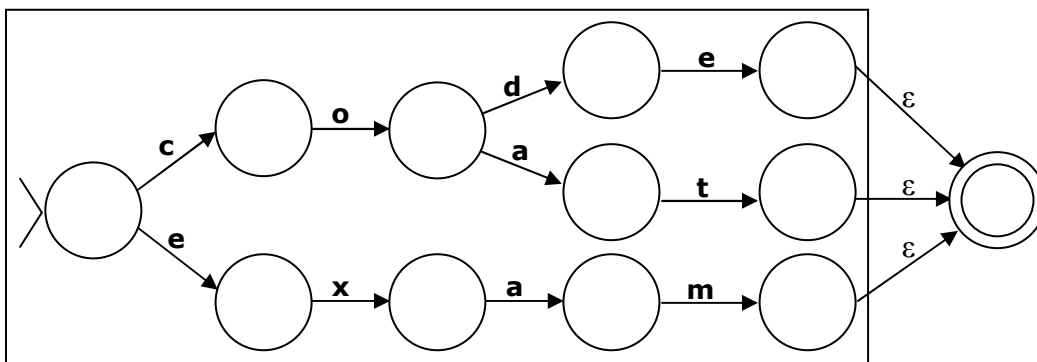
Class One	coat coats code codes exam exams	Class Three	baby babies body bodies lady ladies
Class Two	inch inches hoax hoaxes kiss kisses	Class Four	sheep sheep goose geese mouse mice

Our alphabet will consist of the 26 letters of the English writing system and one additional symbol, **2**, which indicates that a word is plural. (Since we're dealing only with single words, we don't need **#**, the indicator of word breaks.)

A few reminders: the transition labels of a transducer are input:output pairs, which can be abbreviated to a single symbol when the input and output are the same. The following are equivalent statements:

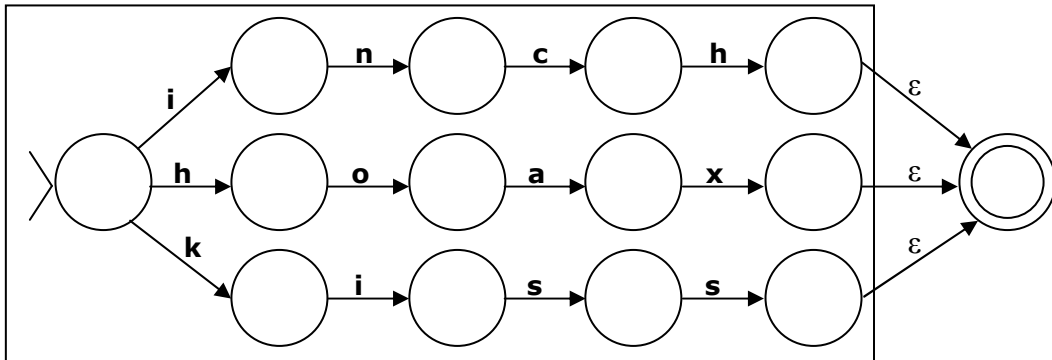
- Transducer T accepts the pair $\langle x, y \rangle$
- Transducer T maps x to y
- Transducer T takes x and returns y

Question 1: Here is machine C_1 , which accepts Class One singular nouns. That is, it takes the string **coat** and returns **coat**, it takes **code** and returns **code**, and it takes **exam** and returns **exam**.



Add states and transitions to this machine so that it also accepts Class One plural nouns: that is, so that it takes the string **coat2** and returns **coats**; so that it maps **code2** to **codes**; and so on. Don't change anything inside the rectangle! (If you answer on a separate sheet of paper, you don't need to copy the contents of the rectangle; drawing the rectangle is sufficient.)

Question 2: Here is machine C_2 , which accepts Class Two singular nouns.



Once again, add states and transitions to this machine so that it also accepts Class Two plural nouns: it should map **inch2** to **inches**, **hoax2** to **hoaxes**, **kiss2** to **kisses**.

Question 3: Draw a machine C_3 , which accepts both Class Three singular nouns and Class Three plural nouns. (Note that the “rectangle” part of C_3 may have to look somewhat different than that of C_1 and C_2 .)

Note: Throughout this question, “accepts plural nouns” specifically means “accepts the pair \langle [noun]**2**, [plural of noun] \rangle ”. For instance, in the first part, the machine should accept the pair \langle **coat2**, **coats** \rangle ; making the machine accept the pair \langle **coats**, **coats** \rangle does not count.

7. EXTRA CREDIT

As on the midterm, extra credit *cannot* be used to make up for questions you do not answer (or answer so perfunctorily that you might as well not have answered).

7.1. Machine C_4

5 points

Draw a machine C_4 for problem 7, i.e. a machine that accepts both Class Four singular nouns and Class Four plural nouns. (As in that question, it should accept **<mouse, mouse>** and **<mouse2, mice>**, and not **<mice, mice>**.)

7.2. Other Codes

5 points

Each of the lines below represents a different way to encode text. Additionally, each line has an extra letter in it. Determine those letters and perform a final (familiar) decoding to get the one-word answer.

1. **1 12 16 5 8 1 14 21 13 5 18 9 3**

2. **November Alpha Tango Oscar Alpha Lima Romeo Papa
Hotel Alpha Bravo Echo Tango**

3. 

4. 

5. **— — — • — • • • • — — • — • — — — — • • •**