naclo

The Third Annual

North American Computational Linguistics Olympiad

2009

www.naclo.cs.cmu.edu

JustSystems Evans Research

The Association for Computational Linguistics North American Chapter

Invitational Round

March 11, 2009

The North American Computational Linguistics Olympiad
www.naclo.cs.cmu.edu

Contest Booklet

Your Name: ______________________________
Registration Number: ______________________________
Your School: ______________________________
City, State, Zip: ______________________________
Your Grade: ________

Start Time (part I): ______________________________
End Time (part I): ______________________________

Start Time (part II): ______________________________
End Time (part II): ______________________________

Your Teacher's Name: ______________________________

Please also make sure to write your registration number and your name on each page that you turn in. Each problem will be graded by a different judge and pages with no registration numbers will not be graded.

n a c l o

Welcome to the invitational round for NACLO 2009. You have shown to be among the top 10% of all participants this year.

Rules

1. The contest is five hours long and includes seven problems, labeled G to M.
2. Follow the facilitators' instructions carefully.
3. If you want clarification on any of the problems, talk to a facilitator. The facilitator will consult with the jury before answering.
4. You may not discuss the problems with anyone except as described in item 3.
5. Each problem is worth a specified number of points, with a total of 100 points. On all problems, points are given for "practice," that is, for getting the right answers. Some problems also assign points for "theory," that is for written descriptions of how you solved the problem. You should therefore show all your work.
6. We will grade only work in this booklet. All your answers should be in the spaces provided in this booklet. DO NOT WRITE ON THE BACK OF THE PAGES.
7. Write your name and registration number on each page:
   Here is an example: Jessica Sawyer #850
8. The top participants in this round across the continent will be eligible to participate in the ILO which is scheduled for July 2009 in Poland.
9. Each problem has been thoroughly checked by linguists and computer scientists as well as students like you for clarity, accuracy, and solvability. Some problems are more difficult than others, but all can be solved using ordinary reasoning and analytic skills. You don't need to know anything about linguistics or about these languages to solve them.
10. If we have done our job well, almost no one will solve all these problems completely in the time allotted. So don't be discouraged if you don't finish everything.
11. If you have any comments, suggestions or complaints about the competition, we ask you to remember these for the web based evaluation. We will send you an e-mail shortly after the competition is finished with instructions on how to fill it out.
12. **DO NOT DISCUSS THE PROBLEMS UNTIL THEY HAVE BEEN POSTED ONLINE!**

Oh, and have fun!

PART I
(problems GHIJ)
(before the break)

Do not work on this part after the break. You have three hours for this part.

YOUR NAME: REGISTRATION #:

(15 points)

(G) Sk8 Parsr (1/4)

Languages are everywhere… even in places where you don't expect them.

Consider the "combo rules" of *P-Little's Triple-I XTreem Hyp0th3tica7 Sk8boarding Game*. In it, players press a series of buttons (left, right, down, up, circle, triangle, square, and X) to make their on-screen avatar perform skateboard tricks that illustrate pro boarder P-Little's "Triple-I" philosophy of Insane, Ill-Advised, and Impossible According to the Laws of Physics. Underneath, the game is using the methods of computational linguistics to turn this "little language" of button presses into tricks and combos. The game uses a simple ***shift-reduce parser*** to parse button "words" into combo "sentences". As each button-press comes in, the corresponding symbols are placed, in order, in a *buffer* (that is, temporary storage space).

1.	↑
2.	↑ ←
3.	↑ ← ▢
4.	↑ ← ▢ ⊗

If, at any point, the *rightmost* symbols in this buffer match any of the patterns on the next page, they are removed and replaced with a new symbol indicating a combo. So, since ▢⊗ corresponds to an "ollie", we replace it with the new symbol **Ollie**.

5.	↑ ← **Ollie**
6.	↑ ← **Ollie** ▢
7.	↑ ← **Ollie** ▢ ▢
8.	↑ ← **Ollie** ▢ ▢ ⊗
9.	↑ ← **Ollie** ▢ **Ollie**

More complex combos can then be built out of simpler combos. You see in the fifth rule on the next page that **Ollie** and **Nollie** can be joined by ▢ to make a new combo. There are also ***rule schemas*** that can create new combos out of *any* kind of combo. The tenth rule on the next page says that *any* combo (represented by α), whether it's an Ollie or an Inverted-360-Kickflip, can be joined with itself by a ▢ to make a Double combo:

10.	↑ ← **Double-Ollie**

YOUR NAME: REGISTRATION #:

(G) Sk8 Parsr (2/4)

The chart of shift-replace rules is given below… but with some holes in it.

If the right side of the input matches…		*… replace it with …*
← ↑ △	⇒	**Backside-180**
________________	⇒	**Frontside-180**
□ ⊗	⇒	**Ollie**
________________	⇒	**Nollie**
Nollie □ Ollie	⇒	________
↓ ↓	⇒	**Crouch**
________________	⇒	**Backside-360**
________________	⇒	**360-Kickflip**
________________	⇒	________
α □ α	⇒	**Double-α**
Double-α □ α	⇒	**Triple-α**
Double-α □ **Double-α**	⇒	**Quadruple-α**
________________	⇒	**Atomic-α**

YOUR NAME: REGISTRATION #:

(G) Sk8 Parsr (3/4)

Complex combos can get pretty involved. Here are a few combos from the manual to give you an idea:

Inverted-Nollie: ↓⊗▢↑
Double-Inverted-Woolie: ↓⊗▢▢▢⊗↑▢↓⊗▢▢▢⊗↑
Inverted-Triple-Backside-180: ↓←↑△▢←↑△▢←↑△↑
Atomic-Double-Frontside-180: →↓◎▢→↓◎▢↓→↓◎▢→↓◎↑
Inverted-Backside-360: ↓←↑△→↓◎↑
Triple-360-Kickflip: ↓↓←↑△→↓◎▢↓↓←↑△→↓◎▢↓↓←↑△→↓◎

G1. How would you perform an "Inverted-Atomic-Backside-360"?

G2. How about an "Atomic-Atomic-Ollie"?

YOUR NAME: REGISTRATION #:

(G) Sk8 Parsr (4/4)

G3. The shift-reduce rules given above are incomplete. Using the descriptions of advanced combos in the manual, can you fill in the missing pieces? State them as concisely as possible. Use the space on page 2/4.

G4. During playtesting, the testers discover that even though combos like "Quadruple -Ollie" and "Quadruple-Inverted-Woolie" are listed in the manual, the game can never actually recognize any Quadruple combo that the player performs. Why not? How could you fix the game so that it can?

G5. What other types of combinations of the listed combos can never actually be pulled off by the player, and why not?

YOUR NAME: REGISTRATION #:

(15 points)

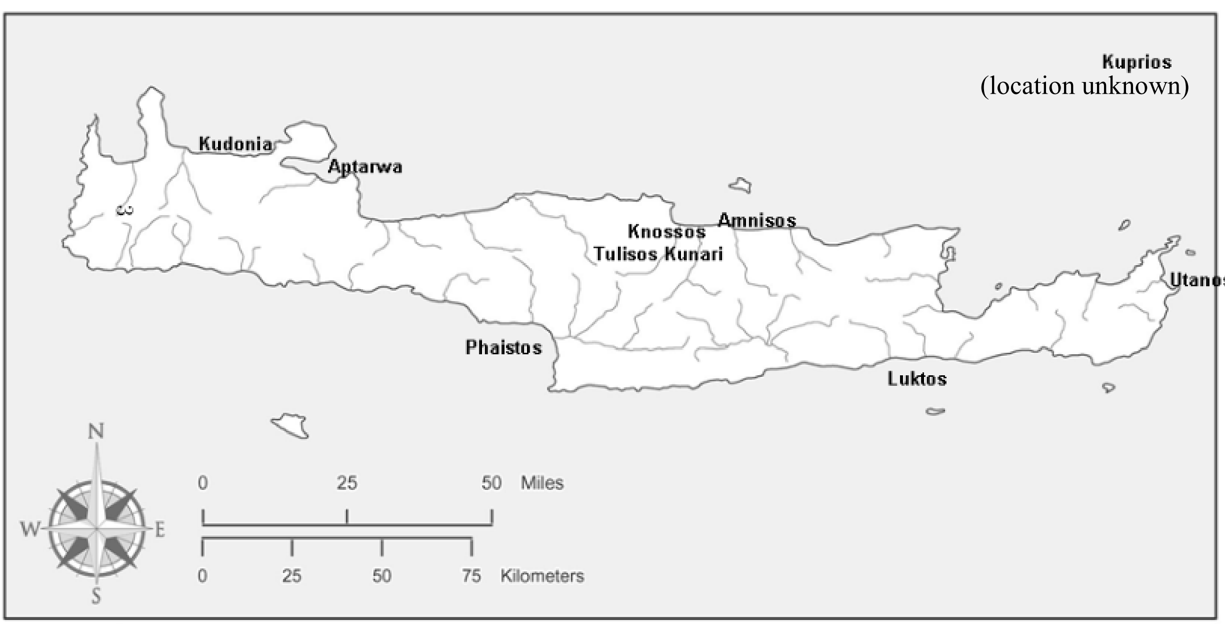
(H) Linear Combinations (1/3)

The script Linear B, deciphered by the architect and amateur epigrapher Michael Ventris in 1953, was used to write Mycenean Greek around the 15th century BCE. Linear B tablets, all of which were accounting records, have been found both on Crete and at various Mycenaean sites on the Greek mainland.

Linear B isn't perfectly suited for Greek; it is an adaptation of another script (Linear A) that was used to write a language about which very little is known. Linear A/B didn't distinguish /l/ and /r/, nor did it have a way to distinguish similar triples of sounds like /ba/, /pa/, and /p^ha/ (which were distinct sounds in Greek), and apparently could only write sequences of V or CV syllables so that a syllable in a Greek word like *k^hru-sos* 'gold' had to be broken up as something like *ku-ru-so.* Here V stands for vowel and C stands for consonant. The superscript ^{h} is used to indicate aspiration.

You will now be asked to decipher a portion of the Linear B symbol set.

The map below shows the approximate locations of a number of ancient Cretan towns: the spellings reflect their probable pronunciation in Mycenaean Greek (not their pronunciation in Modern Greek). Note that we do not know the location of *Kuprios*; also, *Tulisos* and *Kunari* are two different places. Most of these names have stayed more or less the same up until the present day. However one of the names on the map (a "distractor") is not the name that was used in Mycenaean times.

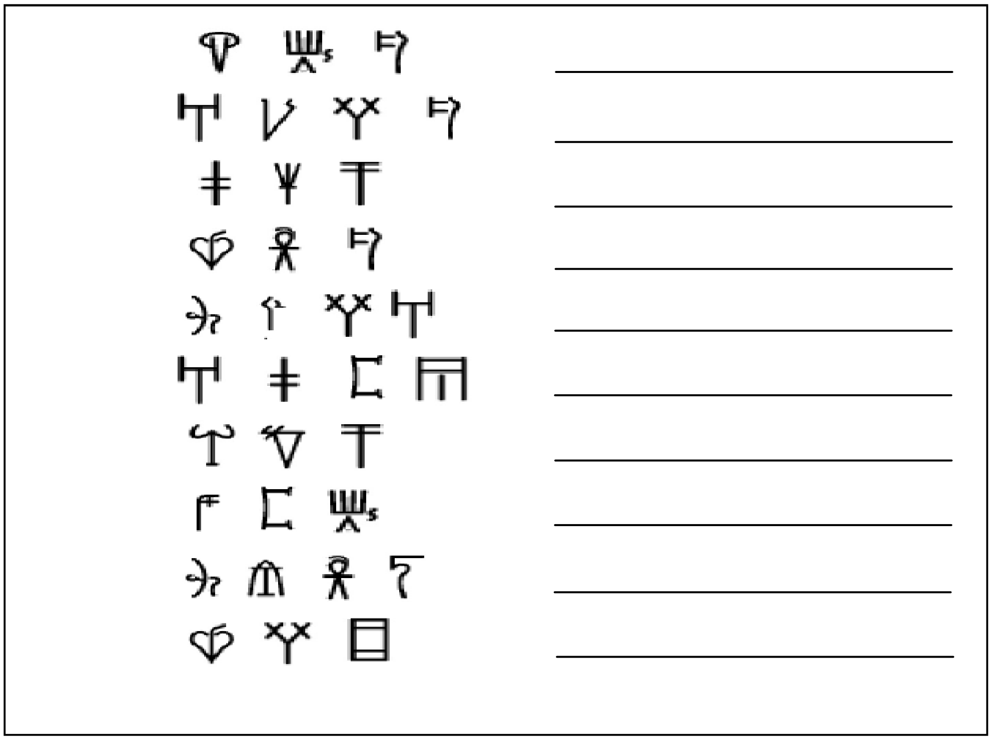


Map by Tom Elliott. Copyright 2003, Ancient World Mapping Center

YOUR NAME: REGISTRATION #:

(H) Linear Combinations (2/3)

H1. Given the spellings of these names in Linear B, can you figure out which is which?

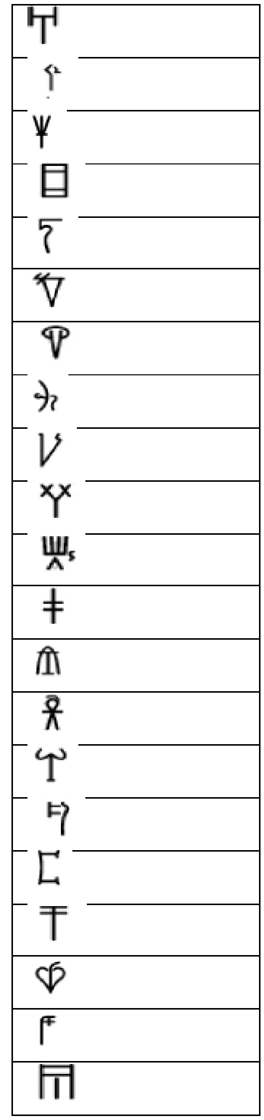


H2. Identify the distractor mentioned on the previous page. What was the approximate Mycenaean pronunciation, given the Linear B spelling? Since one of the symbols used there only occurs in that name in this set, we will even give you a hint: that symbol represents *ja* (pronounced *ya*). Note that if you solve all the others, you will be able to read this name.

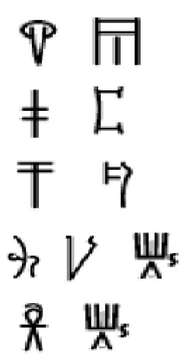
YOUR NAME: REGISTRATION #:

(H) Linear Combinations (3/3)

H3. Write the most likely pronunciation of each Linear B symbol at right.



H4. What are the probable pronunciations of the following words:



	'girl'	____________
	'all'	____________
	'this'	____________
	'cumin'	____________
	'linen'	____________

YOUR NAME: REGISTRATION #:

(15 points)

(I) Easy Pieces (1/2)

Below are phrases in Bulgarian and their translations into English:

	Bulgarian phrase	**English translation**
1	červẹni yạbəlki	red apples
2	kọsteni iglị	bone needles
3	studẹni napịtki	cold drinks
4	dosạdni decạ	annoying children
5	obiknovẹn čovẹk	ordinary person
6	gnẹvni dụmi	angry words
7	červẹn plod	red fruit
8	lẹnen plat	linen fabric
9	sọčni plodovẹ	juicy fruits
10	kọžni zabolyạvaniya	skin diseases
11	gnẹven sədiyạ	angry judge
12	rịbeni kyuftẹta	fish croquettes
13	kirpịčeni kẹšti	adobe houses
14	kọženi rəkavịci	leather gloves
15	lẹsen ịzpit	easy exam
16	cẹnni knịgi	precious books
17	sọčen grẹypfrut	juicy grapefruit
18	cẹnen predmẹt	precious object

Note. Bulgarian is written in the Cyrillic alphabet. Here it is given in transcription, where *č*, *š*, *ž*, and *c* stand for specific consonants of the Bulgarian language and *ə* is a vowel. Stressed vowels are marked with a dot underneath them. For words with only one vowel stress is not marked. Note: a croquette is a small ball of food.

YOUR NAME: REGISTRATION #:

(I) Easy Pieces (2/2)

I1. Three rules govern the formation of the plural of the adjectives. What are they? Write them down below:

I2. Apply the rules from part I1 and fill in the gaps in the following table:

	Bulgarian phrase	English translation
19	... procedụri	ordinary procedures
20	... urọci	easy lessons
21	... restorạnti	fish restaurants
22	... zabolyạvaniya	bone diseases
23	... čaršạfi	linen sheets

YOUR NAME: REGISTRATION #:

(15 points)

(J) Hypo-Hmong-driac (1/2)

The meanings of words may be related in various ways. One of these relations is called *hyponymy.* A word is a *hyponym* of another word if the things or events to which the first word can refer are a subset of the things or events to which the second word can refer. Thus, *spaniel* is a hyponym of *dog* (every spaniel is a dog), *crimson* is a hyponym of *red* (anything that is crimson is red), and *devour* is a hyponym of *eat* (since you cannot devour something without eating it).

Below, you are given a number of words in Mong Leng, also known as Green Hmong (a language of Southern China, Laos, Thailand, and Vietnam) that are related by hyponymy. The subset symbol ⊂ is used here to mean "are hyponyms of." A, B ⊂ C would mean that both A and B are hyponyms of C. Some of the items are simple words, consisting of a single root; others are compound words, made by combining two or more roots. You are also provided with a list of English translations of these words (on the next page). Write the number for the Hmong word next to its English translation.

```
                          (1) sab, (2) ntswg ⊂ (3) sab-ntswg
                          (4) dlej, (5) cawv ⊂ (6) dlej-cawv
                          (7) nyaj, (8) txaj ⊂ (9) nyaj-txaj
                        (10) dlev, (11) npua ⊂ (12) dlev-npua
                        (13) qab,  (11) npua ⊂ (14) qab-npua
                (15) nyuj-twm, (14) qab-npua ⊂ (16) qab-npua-nyuj-twm
                              (17) nqaj-nyuj ⊂ (18) nqaj
                              (19) maum-npua ⊂ (20) maum
                                 (21) sab-twm ⊂ (1) sab
            (22) lug-txaj, (23) lug-dlev-npua ⊂ (24) lug
(25) poob-sab, (26) poob-nyaj, (27) poob-dlej ⊂ (28) poob
  (29) mob-sab, (30) mob-hlwb, (31) mob-ntswg ⊂ (32) mob
```

Note that some Hmong words occur more than once but are always assigned the same number.

YOUR NAME: REGISTRATION #:

(J) Hypo-Hmong-driac (2/2)

```
___ be lost
___ beef
___ beverage
___ bovine* livestock
___ chicken (the animal)
___ dog (the animal)
___ filthy animals; filth
___ filthy language
___ flesh; meat
___ hurt
___ internal organs; soul
___ language
___ liver (the organ)
___ livestock
___ lose heart ("liver"); lose one's wits; panic
___ lose life to water; drown
___ lose money ("silver")
___ lungs
___ money
___ small, non-bovine livestock
___ pig (the animal)
___ poetic genre ("money-language")
___ silver
___ suffer from a headache ("brain-ache")
___ suffer from grief ("liver-ache")
___ suffer from lung disease ("lung-ache")
___ water
___ water-buffalo liver
___ wealth
___ whisky
___ young female
___ young sow**
```

* Bovines are a group of large hooved mammals including cattle, water buffalo, bison, and yaks.
**A sow is a female pig.

PART II

(problems KLM)

(after the break)

Do not work on this part before the break. You have two hours for this part.

(10 points)

(K) The gerbil arrived (1/2)

Below are given Dyirbal sentences and their English translations.

N	Dyirbal	English
1	ŋinda bayi ñalŋga walmbin.	You woke the boy.
2	bayi ŋuma baŋgul ñalŋgaŋgu buṛan.	The boy saw the father.
3	ŋađa banagañu.	I returned.
4	bayi yuṛi baniñu.	The kangaroo came.
5	ŋađa bayi yaṛa buṛan.	I saw the man.
6	bayi ñalŋga baŋgul yaṛaŋgu ñiman.	The man caught the boy.
7	bayi ŋuma ñinañu.	The father sat.

Note: **đ**, **ñ**, **ŋ** and **ṛ** are specific consonants. Dyirbal (pronounced "jirble") is from the Pama-Nyungan language family, and was spoken in Queensland, Australia. It is practically extinct.

K1. How are Dyirbal words and sentences formed? This is an important part of the problem's solution.

YOUR NAME: REGISTRATION #:

(K) The gerbil arrived (2/2)

K2. Give the English translations for these sentences.

bayi ñalŋga banagañu.

bayi yaṛa baŋgul yuṛiŋgu walmbin.

ŋinda bayi yuṛi buṛan .

K3. Give Dyirbal translations for these sentences.

You sat.

I caught the kangaroo.

The father woke the man

YOUR NAME: REGISTRATION #:

(15 points)

(L) Yak, Du, Dray (1/3)

Consider the following arithmetic expression in Kuvi (a language from southeastern India):

(PA:SA x **SA:RI) + (NO:** x **A:TA) = (PA:SA** x **DOS) + (SO:** x **SA:TA)**

The "x" symbol above is the multiplication symbol. The ":" symbol denotes a long vowel. All seven words in the expression above are distinct integers from 1 to 10.

Your task is to order the following expressions by value (in increasing order). No two expressions have the same value.

(A)	**A:TA – RINDI**	(in Kuvi)
(B)	**DHJETË – GJASHTË**	(in Albanian)
(C)	**HASHT – SE**	(in Farsi)
(D)	**SÉ – CÚIG**	(in Irish)
(E)	**CHA – CHA:R**	(in Nepali)
(F)	**NAYN – EYNS**	(in Yiddish)
(G)	**DAS – TIN**	(in Pengo)
(H)	**AŠTUONI – PENKI**	(in Lithuanian)

The "–" symbol above is a minus. The eight expressions correspond to eight distinct positive integers.

As you can easily guess, solving this problem with only the information given above is impossible.

However, we have some additional information that we can use. On the next page you can see the numbers from 1 to 10 in a few languages. Each line lists all these numbers in the given language.

YOUR NAME: REGISTRATION #:

(L) Yak, Du, Dray (2/3)

Nepali: **A:T, CHA, CHA:R, DAS, DUI, EK, NAU, PA:NCH, SA:T, TIN**
Pengo: **AT, CAR, CO, DAS, NOV, PÃC, RI, RO, SAT, TIN**
Farsi : **CHAHA:R, DAH, DO, HAFT, HASHT, NOH, PANJ, SE, SHESH, YAK**
Lithuanian: **AŠTUONI, DEŠIMT, DEVYNI, DU, KETURI, PENKI, SEPTYNI, ŠEŠI, TRYS, VIENAS**
Albanian: **DHJETË, DY, GJASHTË, KATËR, NËNTË, NJË, PESË, SHTATË, TETË, TRE**
Yiddish: **AKHT, DRAY, EYNS, FINF, FIR, NAYN, TSEN, TSVEY, ZEKS, ZIBN**
Irish: **AON, CEATHAIR, CÚIG, DEICH, DÓ, NAOI, OCHT, SÉ, SEACHT, TRÍ**

Note that on each row **above**, the numbers are sorted **alphabetically** (using their Latin transcriptions) and **not numerically**. The languages themselves are sorted geographically from East to West. Pengo and Kuvi are from the Dravidian family of languages. The other languages used in this problem belong to the Indo-European language family. The Dravidian languages use several number words of Indo-European origin.

Next you also have access to the following lists of numbers (**this time** sorted **numerically** from 1 to 10 on each line):

German: eins, zwei, drei, vier, fünf, sechs, sieben, acht, neun, zehn
Latin: unus, duo, tres, quattuor, quinque, sex, septem, octo, novem, decem
Ancient Greek: en, duo, tria, tettara, pente, hex, hepta, octo, ennea, deca

YOUR NAME: REGISTRATION #:

(L) Yak, Du, Dray (3/3)

L1. Fill in the blanks in the table below with the letters A — H as appropriate. One cell should remain blank. (It should be obvious why there are no 0 or 10 columns).

1	2	3	4	5	6	7	8	9

L2. Use the space below to explain (concisely, yet precisely) the key insights that you used in solving this problem.

YOUR NAME: REGISTRATION #:

(15 points)

(M) Orwellspeak (1/5)

Part 1. Opposites Attract

Here is a fragment of an English grammar. If you speak according to this grammar, you will utter sentences like `happy people love charming bad people.`

1.	**Sentence**	→	**NounPhrase + Verb + NounPhrase**
2.	**NounPhrase**	→	**Noun**
3.	**NounPhrase**	→	**Adjective + NounPhrase**

4.	**Noun**	→	`people`
5.	**Verb**	→	`love`
6.	**Adjective**	→	`good`
7.	**Adjective**	→	`charming`
8.	**Adjective**	→	`happy`
9.	**Adjective**	→	`bad`
10.	**Adjective**	→	`obnoxious`
11.	**Adjective**	→	`unhappy`

What do the above grammar rules mean? For example,

- Rule 1 says that to utter a **Sentence**, one must utter a **NounPhrase**, then a **Verb**, then a **NounPhrase**
- Rules 2-3 offer two choices for uttering a **NounPhrase**: one may either utter a **Noun**, or utter an **Adjective** followed by another **NounPhrase**.
- Rules 6-11 offer several choices of **Adjective**.

YOUR NAME: REGISTRATION #:

(M) Orwellspeak (2/5)

Now, keep in mind that **opposites attract**. So it is true that

- `good people love bad people`
- `good happy people love obnoxious people`
- `happy charming people love unhappy obnoxious bad unhappy unhappy people`

and also vice-versa,

- `bad people love good people`
- `obnoxious people love good happy people`
- `unhappy obnoxious bad unhappy unhappy people love happy charming people`

But it is false that

- `good people love good people`
- `obnoxious people love bad unhappy people`
- `people love good people`

M1. Following the example of the totalitarian government in George Orwell's famous book *1984*, we would like you to revise the grammar of English so that it does not permit false **Sentences**.[1] The above grammar permits many **Sentences**. Your revised grammar should permit only a subset of these, using the same notation. It should systematically enforce the principle that opposites (and only opposites) attract. For example, it should be possible to utter the true example **Sentences** above but not the false ones. It should also be impossible to discuss `charming bad people` or `unhappy good people`. (Such people pose intolerable problems for our moral philosophy, and their situations will be corrected forthwith.)

Please show your revisions directly on the grammar on the previous page, using the same notation, by adding new rules and by crossing out or otherwise modifying some of the old rules.

---

[1] "It was intended that when Newspeak had been adopted once and for all and Oldspeak forgotten, a heretical thought—that is, a thought diverging from the principles of [the Party] - should be literally unthinkable, at least so far as thought is dependent on words. Its vocabulary was so constructed as to give exact and often very subtle expression to every meaning that a Party member could properly wish to express, while excluding all other meaning . . . " - from "The Principles of Newspeak", an appendix to *1984* by G. Orwell, 1948.

(M) Orwellspeak (3/5)

Part 2. Censorship

Consider again the setting of the previous problem ("Opposites Attract"). You can do this problem even if you did not solve the previous problem.

In an orderly society, only true, well-formed sentences should be uttered. Thus, our censors should detect illegal utterances like

- `good people love happy people` (which is false since only opposites attract)
- `bad bad bad` (which is nonsense and possibly a subversive code)
- `good charming people love` (which is not a complete sentence)

To be precise, if an utterance is a possible **Sentence** under the revised grammar that you were asked to write in the previous problem, then it is legal. Otherwise it is illegal and must be censored.

A vendor of censorship software has proposed a faster solution that does not use a grammar. Their device censors an utterance if and only if it contains at least one **bad phrase**. Each bad phrase in the device's memory is a sequence of **up to 4** adjacent words.

An input utterance would be presented in the form

```
START good people love obnoxious happy END
```

Input utterances may be of any length. You may assume that they begin with `START` and end with `END`, and that in between, they use only words from the 8-word vocabulary

`{people, love, good, charming, happy, bad, obnoxious, unhappy}`.

YOUR NAME: REGISTRATION #:

(M) Orwellspeak (4/5)

M2. The vendor's device has been carefully constructed to censor as many illegal utterances as possible while not censoring any legal ones. What is the shortest possible list of bad phrases that will do this? Write out a summary of the phrases on the list, and be sure to give the total number of phrases. For example, the list might include the 2-word bad phrase `people charming`, since this can never occur in a legal sentence. To indicate that there are 6 bad phrases of this general form, your summary list might include a line

`people` **Adjective** (6)

or if you prefer,

Noun Adjective (6)

(Hint: The vocabulary consists of 1 **Noun**, 1 **Verb**, 3 positive **Adjective**s, and 3 negative **Adjective**s. A bad phrase may contain any of these words, and may also contain `START` and/or `END`. Remember that a bad phrase may consist of UP TO 4 adjacent words)

(M) Orwellspeak (5/5)

EXTRA SPACE FOR PART **M2.**

M3. Does the resulting device ever fail to censor an illegal utterance? If so, give an example.

M4. Suppose the government tightens its grip, and requires that the vendor modify its machine to censor ALL illegal utterances (even if this means censoring some legal ones as well). What is the shortest possible list of bad phrases that meets this new requirement?

NACLO 2009 organizers

General chair:
Lori Levin, Carnegie Mellon University

Program committee chair:
Dragomir Radev, University of Michigan

Program committee:
Emily Bender, University of Washington
Aleka Blackwell, Middle Tennessee State University
Bozhidar Bozhanov
Eric Breck, Rhodes College
Ivan Derzhanski, Bulgarian Academy of Sciences
Jason Eisner, Johns Hopkins University
Eugene Fink, Carnegie Mellon University
Adam Hesterberg, Princeton University
Anatole Gershman, Carnegie Mellon University
Boris Iomdin, Russian Academy of Sciences
Joshua Katz, Princeton University
Lori Levin, Carnegie Mellon University
Patrick Littell, University of British Columbia
Scott Mackie, University of British Columbia
Ruslan Mitkov, University of Wolverhampton
David Mortensen, University of Pittsburgh
Barbara Partee, University of Massachusetts
Thomas Payne, University of Oregon
James Pustejovsky, Brandeis University
Harold Somers, Dublin City University
Ekaterina Spriggs, Carnegie Mellon University
Richard Sproat, Oregon Health and Science University
Todor Tchervenkov, Université Lumière Lyon 2
Xiaojin "Jerry" Zhu, University of Wisconsin
Richard Wicentowski, Swarthmore College

School liaison:
Amy Troyani, Taylor Allderdice High School

Administrative assistant:
Mary Jo Bensasi, Carnegie Mellon University

NACLO 2009 organizers (cont'd)

Problem credits:
Problem G: Patrick Littell
Problem H: Richard Sproat
Problem I: Todor Tchervenkov
Problem J: David Mortensen
Problem K: Bozhidar Bozhanov
Problem L: Dragomir Radev
Problem M: Jason Eisner

Other members of the organizing committee:
Eugene Fink, Carnegie Mellon University
Ketty Gann, Boeing
Patrick Littell, University of British Columbia
Thomas Payne, University of Oregon
Susanne Vejdemo, the Linguist List
Richard Wicentowski, Swarthmore College

Web site and registration:
Ida Mayer, Carnegie Mellon University
Eugene Fink, Carnegie Mellon University

Jury members
The program committee plus:
Jae-Kyu Lee, Princeton University
Helen Mukomel, DeVry University

US Team coaches:
Dragomir Radev, University of Michigan, head coach
Lori Levin, Carnegie Mellon University, coach
Adam Hesterberg, Princeton University, assistant coach

Student assistants:
Adam Emerson, University of Michigan
Blumie Gurarie, University of Michigan
Mercedes Harvey, University of Michigan
Nate LaFave, University of Michigan
Victor Pudeyev, University of Michigan
Vahed Qazvinian, University of Michigan
Meredith Rogan, University of Michigan
David Ross, University of Michigan
Laine Stranahan, University of Michigan

NACLO 2009 organizers (cont'd)

Contest site coordinators:

Brandeis University: James Pustejovsky
Carnegie Mellon University and University of Pittsburgh: Lori Levin and David Mortensen
Cornell University: Claire Cardie, Randy Hess, Lillian Lee, and Jennifer Wofford
Dalhousie University: Connie Adsett
Georgetown University: Graham Katz
Middle Tennessee State University, Aleka Blackwell
Minnesota State University, Mankato: Rebecca Bates
New York University: Adam Meyers and Satoshi Sekine
Ohio State University: Chris Brew
Princeton University: Christiane Fellbaum, Adam Hesterberg, and Jae-Kyu Lee
San Diego State University: Mark Gawron
San José State University: Roula Svorou
Simon Fraser University: John Alderete, Cliff Burgess, and Maite Taboada
Stanford University: Josh Falk, Jenny Finkel, Dan Jurafsky, and Kyle Noe
University of Great Falls: Porter Coggins
University of Illinois: Roxana Girju and Julia Hockenmaier
University of Illinois, Chicago: Robert Sloan
University of Lethbridge: Kevin Grant
University of Maryland, Baltimore County: Marjorie McShane and Sergei Nirenburg
University of Memphis: Vasile Rus
University of Michigan: Steve Abney and San Duanmu
University of North Texas: Rada Mihalcea
University of Oregon: Thomas Payne
University of Ottawa: Diana Inkpen and Stan Szpakowicz
University of Pennsylvania: Mitch Marcus
University of Southern California: David Chiang
University of Texas at Dallas: Vincent Ng
University of Texas at El Paso: Gabby Gandara
University of Utah: Naomi Fox
University of Washington: Jim Hoard
University of Wisconsin: Nathanael Fillmore and Xiaojin Zhu
High school sites: Dragomir Radev

NACLO 2009 sponsors

Sponsorship chair:
James Pustejovsky, Brandeis University

Corporate, academic, and government sponsors
National Science Foundation
Google
Microsoft
Everyzing
M*Modal
JUST. Systems
The North American Chapter of the Association for Computational Linguistics (NAACL)
Cambridge University Press
Oxford University Press
Carnegie Mellon University's Language Technologies Institute
University of Michigan
Brandeis University
University of Pittsburgh Linguistics Department

Individual contributions
Helga and Charles Marqusee
Linnae and David Evans

Special thanks to:
Tanya Korelsky, NSF
Ben Fortson, University of Michigan
More than 80 high school teachers from 23 states and provinces

And many other individuals and organizations

NACLO 2009 sites

as well as more than 80 high schools throughout the USA and Canada