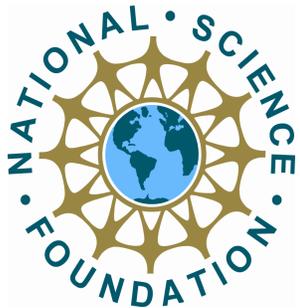


National Sponsors



***The Second  
Annual***

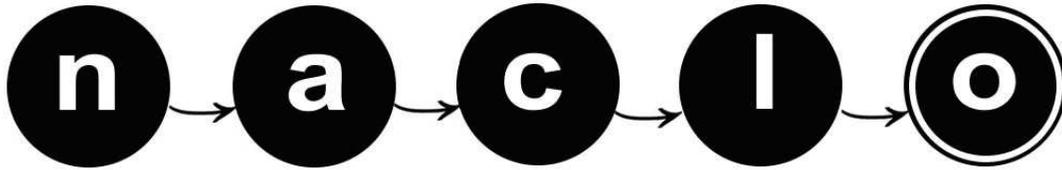
**North American  
Computational  
Linguistics  
Olympiad**

**2008**

**[www.naclo.cs.cmu.edu](http://www.naclo.cs.cmu.edu)**

**Open Round**

**February 5, 2008**



**The North American Computational Linguistics Olympiad**  
**[www.naclo.cs.cmu.edu](http://www.naclo.cs.cmu.edu)**

## Contest Booklet

Your Name: \_\_\_\_\_

Registration Number: \_\_\_\_\_

Your School: \_\_\_\_\_

City, State, Zip: \_\_\_\_\_

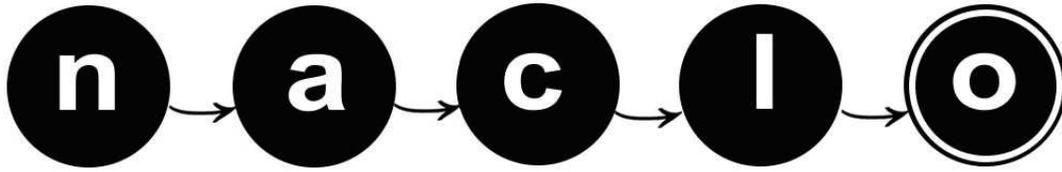
Your Grade: \_\_\_\_\_

Start Time: \_\_\_\_\_

End Time: \_\_\_\_\_

Your Teacher's Name: \_\_\_\_\_

Please also make sure to write your registration number 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.



## Rules

Welcome to the second annual North American Computational Linguistics Olympiad! You are among the few, the brave, the brilliant, to participate in this unique event! Because this is an international competition, in order to be completely fair to all participants across North America, we need you to read, understand and follow these rules completely.

1. The contest is three hours long.
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 3.
5. Each problem is worth a specified number of points, with a total of 100 points in the contest. About 40% of the points are given for "practice," that is, for getting the right answers. The rest of the points are given for "theory," that is, for the quality of your written description of how you solved the problem. It is therefore essential that you 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.
7. Write your name and registration number on each page:  
Here is an example:  
Jessica Sawyer #850
8. Approximately the top 100 participants across the continent in the open round will be invited to the second (invitational) round on March 11. The invitational round will be held under the same conditions and at the same place as the open round, though it will consist of seven problems and will last for five hours.
9. Each problem has been thoroughly checked by linguists and computer scientists as well as students like you for clarity, accuracy, and solvability. Some 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 in order to solve them. 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.

Oh, and have fun!

( 10 points )

**(A) Of monkeys and children**

Apinayé belongs to the Ge language family of Brazil. Currently it is spoken by less than 800 people, and therefore is seriously endangered. The following are some sentences in Apinayé, along with their English translations. You will see some letters here that do not occur in the English or Portuguese writing systems. You do not need to know exactly how these letters are pronounced in order to solve this problem:

Kukrẽ kokoi.

'The monkey eats.'

Ape kra.

'The child works.'

Ape kokoi ratš.

'The big monkey works.'

Ape mĩ metš.

'The good man works.'

Ape metš kra.

'The child works well.'

Ape punui mĩ piŋetš.

'The old man works badly.'

**A1 (practical).** Translate the following into English:

Ape ratš mĩ metš.

Kukrẽ ratš kokoi punui.

Ape piŋetš mĩ.

**A2 (practical).** Translate the following into Apinayé:

'The big child works a long time.'

'The old monkey eats a lot.'

**A3 (theoretical).** Explain the meanings of the following words:

ratš:

metš:

piŋetš:



( 20 points )

**(B) Springing up baby (page 1/3)**

The following sentence pairs are translation equivalents in English and Hindi. There is not enough information in this data set to fully decode the Hindi from the English, but there is enough information to pinpoint the translations of certain words.

In particular, we are interested in the words **spring** and **कल**, each of which is ambiguous, and translate differently in different cases.

|   |
|---|
| The flowers bloom in the spring.<br>कलियाँ वसन्त में खिलती हैं ।  |
| Sita came yesterday.<br>सीता कल आयी थी ।  |
| The gymnast makes springing up to the bar look easy.<br>कसरतबाज डंडे के ऊपर से कूदने के कार्य को आसान बना देता है ।                             |
| It rained yesterday.<br>कल बारिश हुई थी ।   |
| School will commence tomorrow.<br>विद्यालय कल से आरम्भ होगा ।   |
| With a spring the cat reached the branch.<br>वह बिल्ली एक टहनी पर कूद गयी ।   |
| I will come tomorrow.<br>मैं कल आऊँगा ।   |
| The train stopped, and the child sprang for the door and in a twinkling was gone.<br>रेलगाडी के रुकते ही बच्चा दरवाजे से कूदकर रफूचककर हो गया । |
| Sita loves the spring season.<br>सीता को वसन्त ऋतु अच्छी लगती है ।  |
| He will e-mail us tomorrow.<br>वह हमें कल ईमेल करेगा ।  |



## (B) Springing up baby (page 2/3)

**B1 (practical).** In the data on the previous page, **spring** translates to two different Hindi words (in different sentences). Indicate which words from the following list are translations of **spring** (*Just circle the words, rather than trying to copy them over*):

|             |           |          |          |           |
|-------------|-----------|----------|----------|-----------|
| a. ऋतु      | b. कलियाँ | c. कार्य | d. कूद   | e. के     |
| f. गया      | g. टहनी   | h. पर    | i. बच्चा | j. बिल्ली |
| k. रफूचक्कर | l. वसन्त  | m. सीता  | n. से    | o. है     |

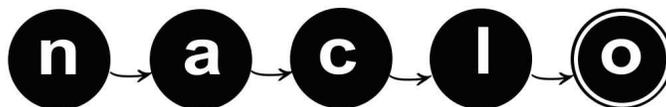
**B2 (practical).** Which of these words is the most likely translation of **spring** in the following sentence?

We always look forward to the **spring** holidays.

Indicate your answer by writing the letter associated with the word here: \_\_\_\_\_.

**B3 (theoretical).** What is it about this sentence that suggests that the word you chose is the proper translation?

**B4 (practical).** In the Hindi sentences on the previous page, **कल** translates to two different English words (in different sentences). What are those words?



## (B) Springing up baby (page 3/3)

**B5 (practical).** What is the most likely translation of कल in the following sentence?

अनामिका यहाँ कल आयी थी ।

**B6 (theoretical).** What is it about this sentence that suggests that the word you chose is the proper translation?



( 20 points )

# (C) Reach for the top (page 1/2)

The Ilocano language is one of the major languages of the Philippines, spoken by more than 8 million people. Today it is written in the Roman alphabet, which was introduced by the Spanish, but before that Ilocano was written in the *Baybayin* script. *Baybayin* (which literally means “spelling”) was used to write many Philippine languages and was in use from the 14<sup>th</sup> to the 19<sup>th</sup> centuries.

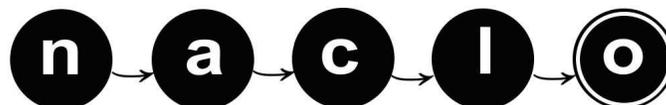
**C1 (practical).** Below are twelve Ilocano words written in Baybayin. Match them to their English translations, listed in scrambled order below.

|              |       |
|--------------|-------|
| ᜆᜄᜇ          | _____ |
| ᜆᜄᜇᜆᜄᜇ       | _____ |
| ᜆᜄᜇᜆ         | _____ |
| ᜆᜄᜇᜆᜆᜆᜆᜆ     | _____ |
| ᜆᜆᜆᜆᜆᜆ       | _____ |
| ᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆ | _____ |
| ᜆᜆᜆᜆ         | _____ |
| ᜆᜆᜆᜆᜆᜆᜆᜆ     | _____ |
| ᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆ   | _____ |
| ᜆᜆᜆᜆᜆᜆᜆ      | _____ |
| ᜆᜆᜆᜆᜆᜆ       | _____ |
| ᜆᜆᜆᜆᜆᜆᜆ      | _____ |

{ to look, is skipping for joy, is becoming a skeleton, to buy, various skeletons, various appearances, to reach the top, is looking, appearance, summit, happiness, skeleton }

**C2 (practical).** Fill in the missing forms.

|              |                  |
|--------------|------------------|
| ᜆᜆᜆᜆᜆᜆ       | _____            |
| ᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆ   | _____            |
| ᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆᜆ | _____            |
| _____        | (the/a) purchase |
| _____        | is buying        |

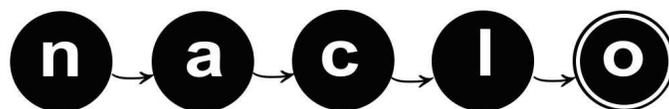


YOUR NAME:

REGISTRATION #:

# (C) Reach for the top (page 2/2)

C3 (theoretical). Explain the reasoning behind your solutions to C1 and C2.



( 25 points )

**(D) Spare the Rod (1/3)**

An excerpt from a well known text is shown below. It is in two languages (X and Y) that are closely linguistically related to each other and also to English. However the two versions are not perfect translations of one another.

Text in language X

- X1. Rödluvan: Men mormor, varför har du så stora ögon?  
 X2. "Mormor": Det är bara för att jag skall se dig bättre, mitt barn.  
 X3. Rödluvan: Men mormor, varför har du så stora öron?  
 X4. "Mormor": Det är bara för att jag skall höra dig bättre, mitt barn.  
 X5. Rödluvan: Men mormor, varför har du så stora tänder?  
 X6. "Mormor": Det är bara för att jag skall kunna äta upp dig!

(almost) the same text in language Y

- Y1. - Så store ører du har, bestemor, sa Rødhette.  
 Y2. - Det er fordi jeg skal kunne høre deg bedre, svarte ulven.  
 Y3. - Så store øyne du har, bestemor, sa Rødhette.  
 Y4. - Det er fordi jeg skal kunne se deg bedre, svarte ulven.  
 Y5. - Så store hender du har, bestemor, sa Rødhette.  
 Y6. - Det er fordi jeg skal kunne klemme deg bedre, svarte ulven.  
 Y7. - Så stor munn du har, bestemor, sa Rødhette.  
 Y8. - Det er fordi jeg skal kunne ete deg bedre, svarte ulven.

**D1 (practical).** Translate sentences X1 and X2 into grammatical English using your own words and word order.

**D2 (practical).** Align the eight sentences in text Y with the six sentences in text X by content. Which two sentences in Y remain unaligned?

|       |    |       |    |
|-------|----|-------|----|
| _____ | Y1 | _____ | Y5 |
| _____ | Y2 | _____ | Y6 |
| _____ | Y3 | _____ | Y7 |
| _____ | Y4 | _____ | Y8 |



# (D) Spare the Rod (2/3)

**D3 (practical).** Fill the leftmost column of the table below. If you believe that a given word is not translated at all, use an X to indicate that.

**D4 (practical).** Now fill out the rightmost column of the table.

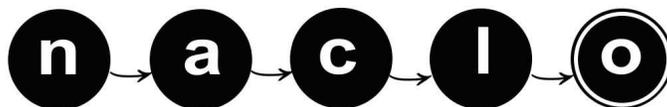
| Language X | Language Y | English |
|------------|------------|---------|
|            | så         |         |
|            | store      |         |
|            | ører       |         |
|            | du         |         |
|            | har        |         |
|            | bestemor   |         |
|            | sa         |         |
|            | Rødhette   |         |
|            | det        |         |
|            | er         |         |
|            | fordi      |         |
|            | jeg        |         |
|            | skal       |         |
|            | kunne      |         |
|            | høre       |         |
|            | deg        |         |
|            | bedre      |         |
|            | svarte     |         |
|            | ulven      |         |
|            | øyne       |         |
|            | se         |         |
|            | hender     |         |
|            | klemme     |         |
|            | stor       |         |
|            | mun        |         |
|            | der        |         |
|            | ete        |         |



# (D) Spare the Rod (3/3)

**D5 (theoretical).** Explain the reasoning that you used to answer parts D1-D4 in detail.

**D6 (theoretical).** Say as much as you can about languages X and Y.

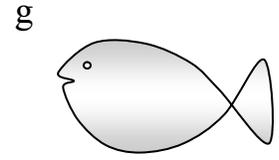
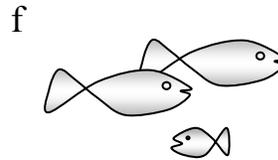
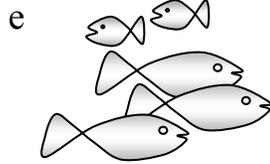
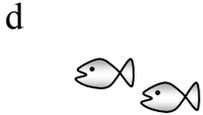
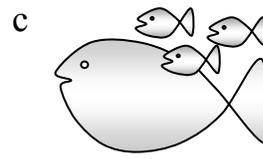
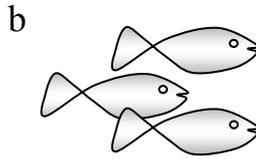
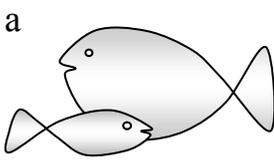


( 25 points )

**(E) A Fish Story (1/2)**

Aymara is a South American language spoken by more than 2 million people in the area around Lake Titicaca, which, at 12,507 feet above sea level, is the highest navigable lake in the world. Among the speakers of Aymara are the *Uros*, a fishing people who live on artificial islands, woven from reeds, that float on the surface of Lake Titicaca.

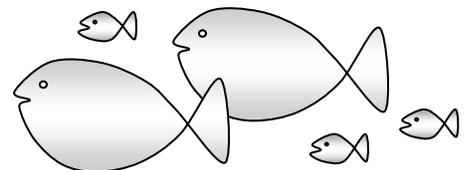
**E1 (practical).** Below, seven fishermen describe their catch. Who caught what?



- \_\_\_ 1. "Mä hach'a challwawa challwataxa."  
 \_\_\_ 2. "Kimsa hach'a challwawa challwataxa."  
 \_\_\_ 3. "Mä challwa mä hach'a challwampiwa challwataxa."  
 \_\_\_ 4. "Mä hach'a challwa kimsa challwallampiwa challwataxa."  
 \_\_\_ 5. "Paya challwallawa challwataxa."  
 \_\_\_ 6. "Mä challwalla paya challwampiwa challwataxa."  
 \_\_\_ 7. "Kimsa challwa paya challwallampiwa challwataxa."

Also, watch out! *One of the fishermen is lying.*

**E2 (practical).** Your daily catch is pictured to the right. Describe it in Aymara, and don't lie!

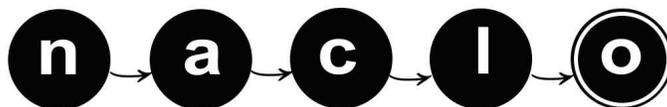


YOUR NAME:

REGISTRATION #:

# (E) A Fish Story (2/2)

E3 (theoretical). Describe your reasoning.



# NACLO 2008 organizers

## General co-chairs:

Lori Levin, Carnegie Mellon University  
Thomas Payne, University of Oregon

## Administrative assistant:

Mary Jo Bensasi, Carnegie Mellon University

## Program committee chair and booklet editor:

Dragomir Radev, University of Michigan

## School liaison:

Amy Troyani, Taylor Allderdice High School

## Program committee:

Emily Bender, U. Washington  
Eric Breck, Cornell U.  
Lauren Collister, U. Pittsburgh  
Eugene Fink, Carnegie Mellon U.  
Adam Hesterberg, Princeton U.  
Joshua Katz, Princeton U.  
Stacy Kurnikova, Carnegie Mellon U.  
Lori Levin, Carnegie Mellon U.  
Will Lewis, U. Washington  
Patrick Littell, U. Pittsburgh  
David Mortensen, U. Pittsburgh  
Barbara Partee, U. Massachusetts  
Thomas Payne, U. Oregon  
James Pustejovsky, Brandeis U.  
Richard Sproat, U. Illinois  
Todor Tchervenkov, U. Lyon 2

## Web site and registration:

Justin Brown, Carnegie Mellon University

## Problem credits:

Problem A: Ronnie Sim and Thomas Payne  
Copyright © 2007, University of Oregon Department of Linguistics  
Problem B: Ankit Srivastava and Emily Bender  
Problem C: Patrick Littell  
Problem D: Dragomir Radev  
Problem E: Patrick Littell

## Additional thanks to:

Tanya Korelsky, NSF  
Paula Chesley, U. Minnesota  
Alina Johnson, U. Michigan  
Patti Kardia, U. Michigan  
Jennifer Wofford, Cornell U.  
Cindy Robinson, Cornell U.  
More than 65 high school teachers from 28 states,  
provinces, and the District of Columbia

## Sponsorship chair:

James Pustejovsky, Brandeis U.

## Site coordinators:

Brandeis University: James Pustejovsky  
Carnegie Mellon University: Lori Levin  
Columbia University: Kathy McKeown and Barry Schiffman  
Cornell University: Claire Cardie and Lillian Lee  
Middle Tennessee State University: Aleka Blackwell  
San José State University: Roula Svorou  
University of Illinois: Roxana Girju and Richard Sproat  
University of Michigan: Steve Abney and Sally Thomason  
University of Oregon: Thomas Payne  
University of Ottawa: Diana Inkpen and Stan Szpakowicz  
University of Pennsylvania: Mitch Marcus  
University of Toronto: Elaine Gold  
University of Wisconsin: Catherine Arnott Smith  
"Online" (high school sites): Dragomir Radev

## NACLO 2008 sites

---



Cornell University

Carnegie Mellon



San José State  
UNIVERSITY



uOttawa



THE UNIVERSITY  
of  
**WISCONSIN**  
MADISON



COLUMBIA UNIVERSITY  
IN THE CITY OF NEW YORK



UNIVERSITY of TORONTO



ILLINOIS  
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN



University of Pittsburgh



UNIVERSITY OF MICHIGAN



Brandeis University

MIDDLE  
TENNESSEE  
STATE UNIVERSITY



Penn  
UNIVERSITY of PENNSYLVANIA



UNIVERSITY OF OREGON

as well as more than 65 high schools throughout the USA and Canada

---