IMPORTANT RULES

To ensure the integrity of the contest:

1. Do not discuss the contents of this booklet with anyone during and after the contest (until it has been posted on the NACLO web site in early March). If you have any questions during the contest, talk quietly to the local facilitators, who will relay your questions to the jury and then give you the official jury answer.

2. Students are not allowed to keep any pages of the booklet after the contest is over.

THE ACTUAL CONTEST BOOKLET STARTS ON PAGE 3

Open Round
February 4, 2010
The Fourth Annual
North American Computational Linguistics Olympiad
2010
www.naclo.cs.cmu.edu

Open Round
February 4, 2010
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: ________________________________
End Time: ________________________________

Your Teacher’s Name: ________________________________

Please also make sure to write your registration number and your name on each page that you turn in.

SIGN YOUR NAME BELOW TO CONFIRM THAT YOU WILL NOT DISCUSS THESE PROBLEMS WITH ANYONE UNTIL THEY HAVE BEEN OFFICIALLY POSTED ON THE NACLO WEB SITE IN EARLY MARCH.

Signature: ________________________________
Welcome to the fourth annual North American Computational Linguistics Olympiad! You are among the few, the brave, and the brilliant, to participate in this unique event. In order to be completely fair to all participants across North America, we need you to read, understand and follow these rules completely.

**Rules**

1. The contest is three hours long and includes seven problems, labeled A to G.
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 items 3 & 12.
5. Each problem is worth a specified number of points, with a total of 100 points.
   In this year’s open round, no points will be given for explanations. Instead, make sure to fill out all the answer boxes properly.
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 100 participants (approximately) across the continent in the open round will be invited to the second round on March 10, 2010.
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 in order to solve them.
10. If we have done our job well, very few people 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! THIS MAY BE SEVERAL WEEKS AFTER THE END OF THE CONTEST.**

Oh, and have fun!
A frequent problem in computational linguistics is that passages often use words that the computer simply doesn’t have in its dictionary. Online slang evolves very fast, people use foreign words in English passages, people make typos and invent new abbreviations, etc. You could add new words to the dictionary as fast as you can find them and the next day the program could still be stumped by a new one!

But the program doesn’t have to give up – instead, it can try to work out as much as it can. Various clues can tell a program whether something is a noun or a verb, a person or an inanimate object, etc., and you can even work out more! The following is a webpage where customers have rated their most recent experience at Gelda’s House of Gelbelgarg. Even if you’ve never heard of any of these dishes, you can still figure out some things about them…

A1. Based on the following reviews, attempt to categorize the following items into:

- I: Individual, discrete food items
- L: Liquids, undifferentiated masses, or masses of uncountably small things
- C: Containers or measurements

You won’t be able to categorize them with 100% certainty, but use the category that you think is most probable for each. Choose a single category for each word below.

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>L</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>färssel-försel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gelbelgarg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gorse-weebel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rolse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flebba</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>göngerplose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>meembel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sweet-bolger</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(A) Gelda’s House of Gelbelgarg (2/3)

Gelda’s House of Gelbelgarg

1138 Euclid Ave.
Neighborhood: Lower Uptown
Category: Ethnic, Specialty
Price Range: $$$
Hours: Mon-Fri. 10:00 a.m. - 9:00 p.m.
Sat. 10:30 a.m. - 11:00 p.m.

mosfe12
Reviews: 2

A hidden gem in Lower Uptown! Get the färsel-förssel with gorse-weebeel and you’ll have a happy stomach for a week. And top it off with a flebba of sweet-bolger while you’re at it!

SanDeE*
Reviews: 2

The portions at this place are just too big! I’d rather have half the portions at a lower price – they just bring out too many göngerplose and too much meembel for me.

wndlHghs40
Reviews: 5

i took my nana here and she said it was just like she remembered from the old country. but the service was a bit lacking – nana ordered four gelbelgarg and the waitress only brought two!
(A) Gelda’s House of Gelbelgarg (3/3)

**xMandiee7x**
**Reviews:** 4

I found the food confusing and disorienting. Where is this from? I randomly ordered the färsel-försel and had to send them back! Three words: weird, weird, and weird.

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Service</th>
<th>Atmosphere</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>☀</td>
<td>★★★★</td>
<td>☀</td>
<td>☀</td>
</tr>
</tbody>
</table>

**wrldTrvl1977**
**Reviews:** 11

I went to Wolserl last year for a holiday, and this is the real thing. If you order the gelbelgarg, though, make sure you also get at least one rolse of sweet-bolger— it’s how the locals like it!

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Service</th>
<th>Atmosphere</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>★★★</td>
<td>★★★★</td>
<td>★★★★★</td>
<td>★★★★</td>
</tr>
</tbody>
</table>

**money@home**
**Reviews:** 103

the prices are steep, but i can afford them – i make up to $75/hr working at home! find out how i do it at http://bit.ly/grhCm

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Service</th>
<th>Atmosphere</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>★★★</td>
<td>★★★★</td>
<td>★★★★★</td>
<td>★★★★</td>
</tr>
</tbody>
</table>

User is on probation

**bu_zhidao**
**Reviews:** 8

not a great date spot! i got a gelbelgarg and a rolse of meembel, but my date was so disoriented that she just ended up with some gorse-weebel. :/

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Service</th>
<th>Atmosphere</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>★★</td>
<td>★★★★</td>
<td>☀</td>
<td>☀</td>
</tr>
</tbody>
</table>

**wembley2000**
**Reviews:** 2

The food was pretty good… But I would have liked more gorse-weebel and fewer göngerplose. You really feel like the chef is skimping on the good stuff..

<table>
<thead>
<tr>
<th></th>
<th>Food</th>
<th>Service</th>
<th>Atmosphere</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>★★★</td>
<td>★★★</td>
<td>★★★★★</td>
<td>★★★</td>
</tr>
</tbody>
</table>
(B) Say it in Abma (1/2)

Abma is an Austronesian language spoken in parts of the South Pacific island nation of Vanuatu by around 8,000 people. Carefully study these Abma sentences, then answer the following questions. Note that there is no separate word for ‘the’ or ‘he’ in these Abma sentences.

Mwamni sileng. He drinks water.
Nutsu mwatbo mwamni sileng. The child keeps drinking water.
Nutsu mwegau. The child grows.
Nutsu mwatbo mwegalgal. The child keeps crawling.
Mworob mwabma. He runs here.
Mwerava Mabontare mwisib. He pulls Mabontare down.
Mabontare mwisib. Mabontare goes down.
Mweselkani tela mwesak. He carries the axe up.
Mwelebte sileng mwabma. He bring water.
Mabontare mworob mwesak. Mabontare runs up.
Sileng mworob. The water runs.

Now, here are some new words in Abma:

sesesrakan teacher
mwegani eat
bwet taro (a kind of sweet potato)
muhural walk
butsukul palm-tree

B1. Translate the following sentences into Abma.

a. The teacher carries the water down.

b. The child keeps eating.

c. Mabontare eats taro.
(B) Say it in Abma (2/2)

B1 (continued). Translate the following sentences into Abma.

d. The child crawls here.


e. The teacher walks downhill.


f. The palm-tree keeps growing upwards.


g. He goes up.


On her visit to Armenia, Millie has gotten lost in Yerevan, the nation’s capital. She is now at the Metropolitan (subway) station named Shengavit but her friends are waiting for her at the station named Barekamutyun. Can you help Millie meet up with her friends?
(C) Lost in Yerevan (2/2)

C1. Assuming Millie takes a train in the right direction, which will be the first stop after Shengavit? Put the correct letter in the box on the right. Note that all names of stations listed below appear on the map.
   (4 points)
   a. Gortsaranayin
   b. Zoravar Andranik
   c. Charbakh
   d. Garegin Njdehi Hraparak
   e. none of the above

C2. After boarding at Shengavit, how many stops will it take Millie to get to Barekamutyun (don’t include Shengavit itself in the number of stops)? (4 points)

C3. What is the name (transcribed into English) of the end station on the short, five-station line that is currently in construction, shown in a different shade on the map? Start writing from the leftmost box.
   (7 points)
(D) Huesvos y Pimientos (1/2)

Paula went shopping while her mother was sick in bed. The only items that Paula had to buy were (according to her mother’s instructions) "red peppers and cucumbers". On the way to the corner store, Paula thought more about her shopping order. It was clear that the peppers had to be red while the cucumbers didn’t have to be red (after all, Paula didn’t think that red cucumbers existed). Paula imagined what she would have had to do if her mom had sent her to buy "red peppers and grapefruit". In that case, she thought, maybe she would have to make sure that the grapefruit were red as well. Or maybe not... Paula was confused.

In linguistics, the problem pondered by Paula is called "attachment ambiguity". Does the adjective ("red") attach to (describe) the nearest noun ("peppers") only or does it attach to the entire noun phrase "peppers and cucumbers")? In some cases, world knowledge can help. We agree with Paula that cucumbers cannot be red so one of the possible interpretations of "red peppers and cucumbers" is actually unlikely. In other cases, e.g., "old boys and girls", both interpretations ("old boys and old girls" and "old boys and girls of any age") are reasonable.

Paula's best friend, Cecilia, speaks Spanish at home. Cecilia and Paula often help each other with homework or with household chores. On her way to the store, Paula ran into Cecilia and wanted to tell her about the linguistic problem that was on her mind. She remembered the Spanish words for "red" ("rojos" in plural), "peppers" ("pimientos"), "and" ("y"), "cucumbers" ("pepinos"), and "grapefruit" ("pomelos" in plural) and also remembered that in Spanish the adjective comes after the noun that it describes (e.g., "pomelos rojos", literally meaning "grapefruit (plural) red" or "niñas pequeñas" which literally translates as "girls small"). When she told Cecilia that she was on her way to buy some "pimientos y pepinos rojos" ("peppers and cucumbers red"), Cecilia started to laugh. Paula realized that in her Spanish translation, not only did the cucumbers now appear to be red but it was now also unclear whether the peppers themselves had to be red.

D1. How could Paula translate each of the following phrases into Spanish and preserve all ambiguities ("uncertainties") as well as all certainties present in their English versions? (4 points)

a. red peppers and cucumbers (give two distinct answers that work—one per line)

b. red peppers and grapefruit (give one answer)

D2. A very popular children's book by Dr. Seuss is called "Green Eggs and Ham". Ignoring the actual contents of the book, can you determine based solely on the book's title whether these statements are true. (2 points each)

a. the eggs are unambiguously green (T/F)

b. the ham is unambiguously green (T/F)
(10 points)

(D) Huevos y Pimientos (2/2)

D3. Consider the following translations from English into Spanish. (2 points)

ham = jamón
eggs = huevos
green (plural) = verdes

How would you translate the title of the book into Spanish (again disregard the actual translation, if you happen to know it) in order to preserve any ambiguities and certainties present in the English title?

This problem is based on the following paper: Kevin Knight and Irene Langkilde. Preserving Ambiguities in Generation via Automata Intersection. In Proceedings of AAAI 2000.
(E) Texting, Texting, One Two Three  (1/3)

The respected espionage-supply company Z Enterprises is about to release a new version of their Z1200 model wristwatch, popular among spies (and also among high-school students) for its ability to discreetly send text messages. Although the Z1200 had only four buttons in total, the user could input characters (letters, numbers, spaces, etc.) by pressing three-button sequences. For example, if we call the buttons 1, 2, 3, and 4, a was 112, A was 113, b was 114, SPACE was 111, the END sequence that finished the message was 444, etc.

The Z1300 has the same button layout, and it was planned that it use the same text-input method. In the design stage, however, a new engineer proposes that he can significantly reduce the number of button presses needed for each message. Unfortunately, the manual had already been printed and the new Z1300 shipped without any information regarding how to use this new input method.

Being a good spy and/or high school student, though, you can figure out how it works just from a few examples, right?

Testing testing
332221432241423411222143224142341331

Does anyone copy
3323332214313142343324221124232343331

be vewy vewy qwiet im hunting wabbits
2341211234221344341234221344341234423444412122141243123124
14222414234113443123412341412243331

Mission failed Tango not eliminated
33243414343413242124431412322123313322314234132142322212123241243414231222123331

my boss Z is a pain in the
24334312341324341332344414143131134231414214212223121331

uh oh no backspace on this thing
2412311322311423212341312422343433423124221132412223141431222314142341331

just kiddin boss
2344324143221234341233233414212341324343331
(E) Texting, Texting, One Two Three  (2/3)

E0. What are the input codes for each of the lowercase letters? Not every letter is used in the messages above, but you can still deduce how they are encoded. This table is just for your own use and it will not be graded.

| a | n |
| b | o |
| c | p |
| d | q |
| e | r |
| f | s |
| g | t |
| h | u |
| i | v |
| j | w |
| k | x |
| l | y |
| m | z |

E1. What message does the following sequence of button presses encode? Start filling the boxes from the left end, one English letter (or space) in each box. (4 points)

231212322321414313142343234132233343123241432221424142341331

n a c l o
(E) Texting, Texting, One Two Three  (3/3)

E2. With what sequences of button presses would you input the following messages? (4 points each)

<table>
<thead>
<tr>
<th>Message</th>
<th>Buttons</th>
</tr>
</thead>
<tbody>
<tr>
<td>help</td>
<td></td>
</tr>
<tr>
<td>xray</td>
<td></td>
</tr>
<tr>
<td>affirmative</td>
<td></td>
</tr>
<tr>
<td>Mayday mayday SOS</td>
<td></td>
</tr>
</tbody>
</table>

E3. This scheme only shortens the number of button presses needed on average – most messages are shorter, but there are some that will take more presses than they did on the Z1200*. Can you find a message (using only characters whose codes you know) that will be longer using the above method than it would have been if it used exactly three button presses per character (including the END sequence)? Enter your message as letters (like 'abc...') rather than as the numerical code (like '12341234...'). (5 points)

*This is true for every compression scheme, actually – for any method of compressing data into less space, there will always be some example that when “compressed” is larger than it was originally!
(15 points)

(F) Türkçe Deliti (1/2)

Given are Turkish words and their English translations:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>güreşçi</td>
<td>wrestler</td>
</tr>
<tr>
<td>B</td>
<td>ikbalsız</td>
<td>unsuccessful</td>
</tr>
<tr>
<td>C</td>
<td>gözcü</td>
<td>sentry, eye doctor</td>
</tr>
<tr>
<td>D</td>
<td>isimsiz</td>
<td>nameless</td>
</tr>
<tr>
<td>E</td>
<td>ormançı</td>
<td>forester</td>
</tr>
<tr>
<td>F</td>
<td>sonsuz</td>
<td>endless</td>
</tr>
<tr>
<td>G</td>
<td>içkici</td>
<td>drunkard</td>
</tr>
<tr>
<td>H</td>
<td>takatsız</td>
<td>lacking strength</td>
</tr>
<tr>
<td>I</td>
<td>barutçu</td>
<td>gunpowder maker</td>
</tr>
<tr>
<td>J</td>
<td>sütsüz</td>
<td>without milk</td>
</tr>
<tr>
<td>K</td>
<td>balıkçı</td>
<td>fisherman</td>
</tr>
<tr>
<td>L</td>
<td>parasız</td>
<td>cashless</td>
</tr>
<tr>
<td>M</td>
<td>mumcu</td>
<td>candlemaker</td>
</tr>
</tbody>
</table>

**F1.** Two of the above words are formed in a slightly different way from the others because their stems are loans from another language. Identify those two words.

Put their letters here: [ ] [ ] (1.5 points each)
(e.g., D L)

**F2.** Translate into Turkish (write one letter in each box, starting from the left; it is ok to leave blank boxes after your answer). Use lowercase letters only. Remember that i and ì are distinct letters. (2 points each).

milkman: [ ] [ ] [ ] [ ] [ ] [ ] [ ]
blind: [ ] [ ] [ ] [ ] [ ] [ ] [ ]
(F) Türkîş Delît (2/2)

F3. Given are the following Turkish words (not loans from another language):

<table>
<thead>
<tr>
<th>dil</th>
<th>language</th>
</tr>
</thead>
<tbody>
<tr>
<td>kalip</td>
<td>form</td>
</tr>
</tbody>
</table>

Translate into Turkish: (write one letter in each box, starting from the left). Use lowercase letters only (2 points each)

linguist:

mute:

mold maker:

shapeless:

Note: ç sounds like ch in church, c like j in job, ş like sh in shoe. e, i, o, and u are pronounced approximately like in red, reed, rod, and rude, respectively. ö and ü are respectively e and i, pronounced with the lips rounded. ı (written like an “i” but without a dot on top) is like u, pronounced with the lips spread (unrounded).

Turkish is a language from the Turkic group of the Altaic language family. It is spoken by 60 million people in Turkey and roughly 10 million other people around the world.
Tangkhul is a language spoken in the northernmost district of the Indian state of Manipur. Like Manipuri (or Meithei) and many other languages of Northeast India, Tangkhul is related to Tibetan and Burmese rather than to Hindi, Bengali, Marathi, Gujarati, or other well-known languages of India.

Tangkhul words can be very long and quite complicated in their structure. Sometimes single words may have to be translated with whole sentences in English. Also, pronouns (words like he, she, it, and they) can be left out if their meanings can still be filled in from context. Following are a list of sentences from Tangkhul and their English translations (in alphabetical order). In the English translations, pronouns are enclosed in parenthesis when they are left out of the Tangkhul sentences. Tangkhul, unlike Modern English (but like Old English), distinguishes three different grammatical numbers: singular (referring to one person or thing), dual (referring to two persons or things), and plural (referring to three or more persons or things). The abbreviations sg., dl., and pl. indicate “singular,” “dual” and “plural,” respectively.

**G1.** Match the Tangkhul sentences with their English translations by writing the number of the English translation by the corresponding Tangkhul sentence (8 points)

<table>
<thead>
<tr>
<th>Tangkhul sentences</th>
<th>English translations</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) a masikserra</td>
<td>1) Do they (pl.) want to pinch one another?</td>
</tr>
<tr>
<td>b) āni masikngarokei</td>
<td>2) Do you (sg.) see it?</td>
</tr>
<tr>
<td>c) āthum masikngarokngāilā</td>
<td>3) Have you (pl.) all come?</td>
</tr>
<tr>
<td>d) ini thāngarokei</td>
<td>4) He/she will pinch all (of them).</td>
</tr>
<tr>
<td>e) na thālā</td>
<td>5) (They) all have come.</td>
</tr>
<tr>
<td>f) ithum thāngāihāirara</td>
<td>6) They (dl.) pinched one another.</td>
</tr>
<tr>
<td>g) rāserhāira</td>
<td>7) They (dl.) will come.</td>
</tr>
<tr>
<td>h) āni rāra</td>
<td>8) We (pl.) will have wanted to see (it).</td>
</tr>
<tr>
<td>i) nathum rāserhāiralā</td>
<td>9) We (dl.) saw one another.</td>
</tr>
</tbody>
</table>

| A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
(20 points)

(G) Tangkhul Tangle (2/2)

G2. Translate the following sentences into English. Always start with the leftmost box. Please follow the style of the English translations given in G1 as closely as possible. (6 points)
   a) nathum masikserngäära

   

   b) äthum thäici

   

   c) i thäiserhāiralā

G3. Translate the following sentences into Tangkhul (6 points).

1) Do you (dl.) want to come?

2) You (sg.) have seen (it) all.

3) We (pl.) will want to see one another.
NACLO 2010 organizers

General chair:
Lori Levin, Carnegie Mellon University

Program committee chair:
Dragomir Radev, University of Michigan

Program committee:
Emily Bender, University of Washington
John Berman, Massachusetts Institute of Technology
Steven Bird, University of Melbourne
Aleka Blackwell, Middle Tennessee State University
Bozhidar Bozhanov, Bulgaria
Eric Breck, Cornell University
Ivan Derzhanski, Bulgarian Academy of Sciences
Jason Eisner, Johns Hopkins University
Dominique Estival, Australia
Eugene Fink, Carnegie Mellon University
Adam Hesterberg, Princeton University
Richard Hudson, University College London
Anatole Gershman, Carnegie Mellon University
Boris Iomdin, Russian Academy of Sciences
Rebecca Jacobs, University of Chicago
Joshua Katz, Princeton University
Mary Laughren, University of Queensland
Lori Levin, Carnegie Mellon University
Patrick Littell, University of British Columbia
Scott Mackie, University of British Columbia
K P Mohanan, National University of Singapore
Ruslan Mitkov, University of Wolverhampton
David Mortensen, University of Pittsburgh
Ani Nenkova, University of Pennsylvania
Barbara Partee, University of Massachusetts
James Pustejovsky, Brandeis University
Nathan Schneider, Carnegie Mellon University
Catherine Sheard, Yale University
Harold Somers, Dublin City University
Ekaterina Spriggs, Carnegie Mellon University
Richard Sproat, Oregon Health and Science University
Amy Troyani, Taylor Allderdice High School
Susanne Vejdomo, Eastern Michigan University
Xiaojin "Jerry" Zhu, University of Wisconsin
Richard Wicentowski, Swarthmore College

Administrative assistant:
Mary Jo Bensasi, Carnegie Mellon University
NACLO 2010 organizers (cont’d)

Problem credits:
Problem A: Patrick Littell
Problem B: Cindy Schneider
Problem C: Dragomir R. Radev
Problem D: Dragomir R. Radev
Problem E: Patrick Littell
Problem F: Bozhidar Bozhanov
Problem G: David Mortensen

Other members of the organizing committee:
Mary Jo Bensasi, Carnegie-Mellon University
Aleka Blackwell, Middle Tennessee State University
Josh Falk, Stanford University
Eugene Fink, Carnegie Mellon University
Katy Gann, Boeing
Adam Hesterberg, Princeton University
Lori Levin, Carnegie-Mellon University
Patrick Littell, University of British Columbia
James Pustejovsky, Brandeis University
Dragomir Radev, University of Michigan
Amy Troyani, Taylor Allderdice High School
Susanne Vejdemo, Eastern Michigan University
Michael White, Ohio State University
Julia Workman, University of Pittsburgh
Yilu Zhou, George Washington University

Web site and registration:
Adam Emerson, University of Michigan

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

Canadian coordinator:
Patrick Littell, University of British Columbia
Contest site coordinators:
Brandeis University: James Pustejovsky
Carnegie Mellon University and University of Pittsburgh: Lori Levin and David Mortensen
Central Connecticut State University: Seunghun Lee
Columbia University: Kathy McKeown
Dalhousie University: Connie Adsett
Georgetown University: Graham Katz
Indiana University: Markus Dickinson and Sandra Kuebler
Johns Hopkins University: Mark Dredze
Middle Tennessee State University: Aleka Blackwell
Minnesota State University, Mankato: Rebecca Bates
Northeastern Illinois University: Judith Kaplan
Princeton University: Christiane Fellbaum and Adam Hesterberg
Queens College, CUNY: Heng Ji, Matt Huenerfauth, Andrew Rosenberg, Crystal Slaughter, Xiuyi Huang
San José State University: Roula Svorou
Simon Fraser University: John Alderete, Cliff Burgess, and Maite Taboada
Stanford University: Josh Falk, Spence Green, Dan Jurafsky, and Kyle Noe
University at Buffalo: Carl Alphonce
University of Great Falls: Porter Coggins
University of Illinois: Roxana Girju and Julia Hockenmaier
University of Illinois, Chicago: Barbara di Eugenio
University of Memphis: Vasile Rus
University of Michigan: Sally Thomason and Steve Abney
University of North Texas: Rada Mihalcea
University of Pennsylvania: Mitch Marcus
University of Rochester: Mary Swift
University of Southern California: David Chiang and Liang Huang
University of Texas at Dallas: Vincent Ng
University of Washington: Jim Hoard
University of Wisconsin: Nathanael Fillmore and Xiaojin Zhu
High school sites: Dragomir Radev

All material in this booklet © 2010, North American Computational Linguistics Olympiad and the authors of the individual problems. Please do not copy or distribute without permission.
NACLO 2010 sponsors

**Student assistants:**
Marcus Berger, University of Michigan  
Reed Blaylock, University of Michigan  
Adam Emerson, University of Michigan  
Amy Hemmeter, University of Michigan  
Ridley Jones, University of Michigan  
Nate LaFave, University of Michigan  
Andrew Lamont, University of Michigan  
Carrie Pichan, University of Michigan  
David Ross, University of Michigan  
Andrea Sexton, University of Michigan  
Samuel Smolkin, University of Michigan  
Laine Stranahan, University of Michigan

**Booklet editors:**
Dragomir R. Radev, University of Michigan  
Nate LaFave, University of Michigan

**Sponsorship chair:**
James Pustejovsky, Brandeis University

**Corporate, academic, and government sponsors**
National Science Foundation
The North American Chapter of the Association for Computational Linguistics (NAACL)
Carnegie Mellon University’s Language Technologies Institute  
University of Michigan  
Brandeis University

**Special thanks to:**
Tanya Korelsky, NSF  
More than 70 high school teachers from 25 states and provinces

And many other individuals and organizations

All material in this booklet © 2010, North American Computational Linguistics Olympiad and the authors of the individual problems. Please do not copy or distribute without permission.
NACLO 2010 sites

as well as more than 70 high schools throughout the USA and Canada