The Tenth Annual North American Computational Linguistics Olympiad 2016
www.nacloweb.org

Open Round
January 28, 2016

Serious language puzzles that are surprisingly fun!
-Will Shortz, Crossword editor of The New York Times and Puzzlemaster for NPR
Welcome to the tenth 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 eight problems, labeled A to H.
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. All your answers should be in the Answer Sheets at the end of this booklet. ONLY THE ANSWER SHEETS WILL BE GRADED.
7. Write your name and registration number on each page of the Answer Sheets’ Here is an example: Jessica Sawyer #850
8. The top 10% of participants (approximately) across the continent in the open round will be invited to the second round.
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 some basic 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. DO NOT DISCUSS THE PROBLEMS UNTIL THEY HAVE BEEN POSTED ONLINE! THIS MAY BE A COUPLE OF MONTHS AFTER THE END OF THE CONTEST.

Oh, and have fun!
NACLO 2016 Organizers

Program Committee:
Adam Hesterberg, Massachusetts Institute of Technology (co-chair)
Alan Chang, University of Chicago
Aleka Blackwell, Middle Tennessee State University
Alex Iriza, Princeton University
Alex Wade, Stanford University
Andrew Lamont, Indiana University
Babette Newsome, Aquinas College
Ben Sklaroff, University of California, Berkeley
Caroline Ellison, Stanford University
Daniel Lovsted, McGill University
David Mortensen, Carnegie Mellon University
David Palfreyman, Zayed University
David McClosky, IBM
Dick Hudson, University College London
Dorottya Demszky, Princeton University
Dragomir Radev, University of Michigan (co-chair)
Elysea Warner, University of Cambridge
Greg Kondrak, University of Alberta
Harold Somers, AILO
Harry Go, WUSTL
James Pustejovsky, Brandeis University
Jason Eisner, Johns Hopkins University
Jonathan Kummerfeld, University of California, Berkeley
Jonathan May, ISI
Jonathan Graehl, SDL International
Jordan Boyd-Graber, University of Colorado
Jordan Ho, University of Toronto
Josh Falk, University of Chicago
Julia Buffinton, University of Maryland
Lars Hellan, Norwegian University of Science and Technology
Lori Levin, Carnegie Mellon University
Lynn Clark, University of Canterbury
Mary Laughren, University of Queensland
Michael Erlewine, National University of Singapore
Oliver Sayeed, University of Cambridge
Patrick Littell, Carnegie Mellon University
Rachel McEnroe, University of Chicago
Richard Littauer
Susan Barry, Manchester Metropolitan University
Tom McCoy, Yale University
Tom Roberts, University of California, Santa Cruz
Verna Rieschild, Macquarie University
Wesley Jones, University of Chicago
Yejin Choi, University of Washington
NACLO 2016 Organizers (cont’d)

Problem Credits:
Problem A: Bozhidar Bozhanov
Problem B: Patrick Littell
Problem C: Praveen Venkataramana
Problem D: Dorottya Demszky, Yiwei Luo, and Tom McCoy
Problem E: Bill Huang
Problem F: Dorottya Demszky
Problem G: James Hyett
Problem H: Ben King

Organizing Committee:
Adam Hesterberg, Massachusetts Institute of Technology
Aleka Blackwell, Middle Tennessee State University
Alex Iriza, Princeton University
Alex Wade, Stanford University
Andrew Lamont, Indiana University
Bill Huang, Princeton University
Caroline Ellison, Stanford University
Daniel Lovsted, McGill University
David Mortensen, Carnegie Mellon University
Deven Lahoti, Massachusetts Institute of Technology
Dorottya Demszky, Princeton University
Dragomir Radev, University of Michigan
Harry Go, Washington University in Saint Louis
James Pustejovsky, Brandeis University
James Bloxham, Massachusetts Institute of Technology
Janis Chang, University of Western Ontario
Jordan Ho, University of Toronto
Josh Falk, University of Chicago
Julia Buffinton, University of Maryland
Laura Radev, Harvard University
Lori Levin, Carnegie Mellon University
Mary Jo Bensasi, Carnegie Mellon University
Matthew Gardner, Carnegie Mellon University
Patrick Littell, Carnegie Mellon University
Rachel McEnroe, University of Chicago
Simon Huang, University of Waterloo
Stella Lau, University of Cambridge
Tom McCoy, Yale University
Tom Roberts, University of California, Santa Cruz
Wesley Jones, University of Chicago
Yilu Zhou, Fordham University
NACLO 2016 Organizers (cont’d)

US Team Coaches:
Dragomir Radev, University of Michigan
Lori Levin, Carnegie Mellon University

Canadian Coordinator and Team Coach:
Patrick Littell, Carnegie Mellon University

USA Contest Site Coordinators:
Brandeis University: James Pustejovsky
Brigham Young University: Deryle Lonsdale
Carnegie Mellon University: Lori Levin, Pat Littell, David Mortensen
College of William and Mary: Dan Parker
Columbia University: Brianne Cortese, Kathy McKeown
Cornell University: Abby Cohn, Sam Tilsen
Dartmouth College: Michael Lefkowitz
Emory University: Jinho Choi, Phillip Wolff
Georgetown University: Daniel Simonson
Goshen College: Peter Miller
Indiana University: Markus Dickinson, Andrew Lamont
Johns Hopkins University: Rebecca Knowles
Massachusetts Institute of Technology: Adam Hesterberg, Sophie Mori
Middle Tennessee State University: Aleka Blackwell
Minnesota State University Mankato: Rebecca Bates, Dean Kelley, Richard Roiger
Northeastern Illinois University: J. P. Boyle, R. Hallett, Judith Kaplan-Weinger, K. Konopka
Ohio State University: Micha Elsner, Michael White
Princeton University: Dorottya Demszky, Christiane Fellbaum, Misha Khodak, Caleb South, Yuyan Zhao
San Diego State University: Rob Malouf
Sega Math Academy: Lucian Sega
Southern Illinois University: Vicki Carstens, Jeffrey Punske
SpringLight Education Institute: Sherry Wang
Stanford University: Sarah Yamamoto
Stony Brook University: Kristen La Magna, Lori Repetti
Union College: Kristina Striegnitz, Nick Webb
University of Alabama, Birmingham: Steven Benthard
University of Colorado at Boulder: Silva Chang
University of Houston: Thamar Solorio
University of Illinois at Urbana-Champaign: Julia Hockenmaier, Ryan Musa
University of Maryland: Julia Buffinton, Kasia Hitczenko
University of Memphis: Vasile Rus
University of Michigan: Steven Abney, Marcus Berger, Sally Thomason
University of Nebraska, Omaha: Ashwathy Ashokan
University of North Carolina, Charlotte: Hossein Hematicalam, Wlodek Zadrozny
University of North Texas: Genene Murphy, Rodney Nielsen
University of Pennsylvania: Chris Callison-Burch, Cheryl Hickey, Mitch Marcus
University of Southern California: Aliya Deri, Ashish Vaswani
University of Texas at Dallas: Luis Gerardo Mojica de la Vega, Jing Lu, Vincent Ng
University of Texas, Austin: Rainer Mueller
University of Utah: Aniko Czirmaz, Mengqi Wang, Andrew Zupon
University of Washington: Jim Hoard, Joyce Parvi
University of Wisconsin, Eau Claire: Lynsey Wolter
University of Wisconsin, Madison: Steve Lacy
University of Wisconsin, Milwaukee: Joyce Boyland, Hanyon Park, Gabriella Pinter
Washington University in Saint Louis: Harry Go, Brett Hyde, Jackie Nelligan
Western Michigan University: Shannon Houtrow, John Kapenga
Western Washington University: Kristin Denham
Yale University: Bob Frank, Raffaella Zanuttini, and the Yale Undergraduate Linguistics Society
NACLO 2016 Organizers (cont’d)

Canada Contest Site Coordinators:
Dalhousie University: Magdalena Jankowska, Vlado Keselj, Dijana Kosmajac, Armin Sajadi
McGill University: Junko Shimoyama, Michael Wagner
Simon Fraser University: John Alderete, Marion Caldecott, Maite Taboada
University of Alberta: Herbert Colston, Sally Rice
University of British Columbia: David Penco, Jozina Vander Klok
University of Calgary: Dennis Storoshenko
University of Lethbridge: Yliias Chali
University of Ottawa: Diana Inkpen
University of Toronto: Jordan Ho, James Hyett, Pen Long
University of Western Ontario: Janis Chang

High school sites: Dragomir Radev

Booklet Editors:
Andrew Lamont, Indiana University
Dragomir Radev, University of Michigan

Sponsorship Chair:
James Pustejovsky, Brandeis University

Sustaining Donors
Linguistic Society of America
NAACL
NSF
Yahoo!

Major Donors
ACM SIGIR
ARL
Choosito!
DARPA
Linguistic Data Consortium (LDC)

University Donors
Brandeis University
Carnegie Mellon University
University of Maryland
University of Michigan
University of Washington

Many generous individual donors

Special thanks to:
Tatiana Korelsky, Joan Maling, and D. Terrence Langendoen, US National Science Foundation
James Pustejovsky for his personal sponsorship
And the hosts of the 100+ High School Sites

All material in this booklet © 2016, North American Computational Linguistics Olympiad and the authors of the individual problems. Please do not copy or distribute without permission.
As well as more than 120 high schools throughout the USA and Canada
Malay is an Austronesian language spoken by the Malay people and people of other ethnic groups who reside in Malaysia (indicated on the map below), southern Thailand, the Philippines, and Singapore. Below are eight sentences in Malay, along with their English translations:

<table>
<thead>
<tr>
<th>Malay</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gadis cantik itu tidak kaya.</td>
<td>The beautiful girl is not rich.</td>
</tr>
<tr>
<td>Penyanyi itu tidak bahagia.</td>
<td>The singer is not happy.</td>
</tr>
<tr>
<td>Kekayaan itu bukan dari bapanya.</td>
<td>The wealth is not from his father.</td>
</tr>
<tr>
<td>Wang bukan kebahagiaan.</td>
<td>Money is not happiness.</td>
</tr>
<tr>
<td>Kereta itu tidak datang dari medan itu.</td>
<td>The car is not coming from the field.</td>
</tr>
<tr>
<td>Manusia itu depan rumah itu bukan penyanyi.</td>
<td>The man in front of the house is not a singer.</td>
</tr>
<tr>
<td>Hadiah itu bukan untuk bapa itu.</td>
<td>The gift is not for the father.</td>
</tr>
<tr>
<td>Gadis bahagia itu tidak menangis.</td>
<td>The happy girl is not crying.</td>
</tr>
</tbody>
</table>

**A1.** Translate the following sentences into Malay. Write your answers in the Answer Sheets.

a. Beauty is not a gift.
b. The rich girl is not a singer.
c. His wealth is not for the girl.
d. The man is not coming.
e. The gift from the singer is not beautiful.
“DAWG” (directed acyclic word graph) describes a diagram that stores a set of words in a graph (in the sense of a web of paths) that is directed (each path can only be traveled in one direction) and acyclic (there is no possibility of traveling in a circle).

When storing a set of words (say, HYDROGEN, OXYGEN, and NITROGEN), there is often some redundancy. All three of these words, for example, end in -GEN, and two of them end in -ROGEN. The DAWG below stores all three words without storing the redundant parts multiple times.

![Diagam of DAWG storing HYDROGEN, OXYGEN, and NITROGEN]

This DAWG can “recognize” all three words, in that each word constitutes a valid path from the start symbol to the end symbol, and no other sequence of letters forms such a path.

However, it is not correct to just “merge” any redundant letters like this, because inappropriately merged letters will lead to incorrect words being recognized.

![Diagam of DAWG recognizing NITROGEN, HYDROGEN, and OXYGEN] but incorrectly recognizing NITROXYGEN and HYDROXYGEN]

This DAWG correctly recognizes the letter sequences NITROGEN, HYDROGEN, and OXYGEN, but it also incorrectly recognizes the letter sequences NITROXYGEN and HYDROXYGEN, which were not intended.

Answer these questions in the Answer Sheets.

**B1.** On the next page are three DAWGs that recognize a list of words in a category, the way that the DAWGs above recognizes a three-word list of chemical elements. Each DAWG recognizes a different category of words.

These DAWGs are poorly-constructed, however, in that each one recognizes several incorrect letter sequences as well. We will give you the shape of the DAWG (but without letter labels) and the incorrect letter sequences; from this, deduce what words the DAWG was supposed to recognize and write these intended words on your answer sheet. (For each part of this question, the number of intended words is the same as the number of answer spaces you are given on your Answer Sheet.)

![Diagram of DAWG recognizing NaClO]
B2. For each of the DAWGs above, what would be the fewest number of letter squares needed to recognize every intended word, and only the intended words? (For example, to recognize HYDROGEN, OXYGEN, and NITROGEN, you need at least 14 squares. Any fewer than 14 squares and you would recognize an unintended word like NITROXYGEN or OXYDROGEN. Do not include the start or end spaces in your counts.)
Estonian is a Uralic language closely related to Finnish; it is spoken by over a million people in Estonia. In Estonian, as in many languages around the world, nouns (and adjectives) take different forms (called "cases") according to their role in the sentence. For example, in the sentence below, there are three forms of the word *horse*. This is similar to the difference between *he*, *him*, and *his* in English.

Hobune nägi hobust hobusega.
horse.nominative saw horse.partitive horse.comitative
'A horse saw a horse with a horse.'

Estonian nouns have a lot of different forms depending on their case and their number (singular or plural). The table below illustrates various forms for different nouns and adjectives; the exact meaning of these different cases is not relevant to this problem. Be careful, however: exactly four of the forms in the table below are mistakes! Note that õ, ä, ö, and ü are vowels.

<table>
<thead>
<tr>
<th>English translation</th>
<th>Genitive Singular</th>
<th>Partitive Singular</th>
<th>Adessive Singular</th>
<th>Nominative Plural</th>
<th>Genitive Plural</th>
<th>Adessive Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘house’</td>
<td>maja</td>
<td>maja</td>
<td>majal</td>
<td>majad</td>
<td>majade</td>
<td>majad</td>
</tr>
<tr>
<td>‘nest’</td>
<td>pesa</td>
<td>pesa</td>
<td>pesal</td>
<td>pesad</td>
<td>pesade</td>
<td>pesad</td>
</tr>
<tr>
<td>‘singer’</td>
<td>laulja</td>
<td>laulja</td>
<td>lauljal</td>
<td>lauljad</td>
<td>lauljate</td>
<td>lauljate</td>
</tr>
<tr>
<td>‘restaurant’</td>
<td>söökla</td>
<td>sööklat</td>
<td>sööklal</td>
<td>sööklad</td>
<td>sööklate</td>
<td>sööklate</td>
</tr>
<tr>
<td>‘name’</td>
<td>nime</td>
<td>nime</td>
<td>nimel</td>
<td>nimed</td>
<td>nime</td>
<td>nime</td>
</tr>
<tr>
<td>‘ice’</td>
<td>jää</td>
<td>jääd</td>
<td>jäääl</td>
<td>jääd</td>
<td>jääde</td>
<td>jäädel</td>
</tr>
<tr>
<td>‘summer’</td>
<td>suve</td>
<td>suve</td>
<td>suvel</td>
<td>suved</td>
<td>suvede</td>
<td>suvedel</td>
</tr>
<tr>
<td>‘white’</td>
<td>valge</td>
<td>valget</td>
<td>valgel</td>
<td>valged</td>
<td>valgete</td>
<td>valgetel</td>
</tr>
<tr>
<td>‘sister’</td>
<td>õe</td>
<td>õde</td>
<td>õdel</td>
<td>õed</td>
<td>õdede</td>
<td>õdedel</td>
</tr>
<tr>
<td>‘road’</td>
<td>tee</td>
<td>teed</td>
<td>teel</td>
<td>teed</td>
<td>teede</td>
<td>teedel</td>
</tr>
<tr>
<td>‘big’</td>
<td>suur</td>
<td>suurt</td>
<td>suurel</td>
<td>suured</td>
<td>suurte</td>
<td>suurtel</td>
</tr>
<tr>
<td>‘yellow’</td>
<td>kollase</td>
<td>kollaset</td>
<td>kollasel</td>
<td>kollased</td>
<td>kollaste</td>
<td>kollastel</td>
</tr>
<tr>
<td>‘man’</td>
<td>mehe</td>
<td>meest</td>
<td>mehel</td>
<td>mehed</td>
<td>meeste</td>
<td>meestel</td>
</tr>
<tr>
<td>‘bean’</td>
<td>oa</td>
<td>uba</td>
<td>oal</td>
<td>oad</td>
<td>ubade</td>
<td>ubadel</td>
</tr>
<tr>
<td>‘reason’</td>
<td>põhjuse</td>
<td>põhjust</td>
<td>põhjusel</td>
<td>põhjused</td>
<td>põhjuste</td>
<td>põhjustel</td>
</tr>
<tr>
<td>‘story’</td>
<td>lOO</td>
<td>lugu</td>
<td>lool</td>
<td>lood</td>
<td>lugude</td>
<td>lugudel</td>
</tr>
<tr>
<td>‘island’</td>
<td>saare</td>
<td>saart</td>
<td>saarel</td>
<td>saared</td>
<td>saarte</td>
<td>saartel</td>
</tr>
</tbody>
</table>

(C) The Curious Case of Estonian (1/2) [5 points]
(C) The Curious Case of Estonian (2/2)

Answer these questions in the Answer Sheets.

C1. On your answer sheet, write down the correct forms of the four incorrect words.

C2. Fill in the blanks in the table below.

<table>
<thead>
<tr>
<th>English translation</th>
<th>Genitive Singular</th>
<th>Partitive Singular</th>
<th>Adessive Singular</th>
<th>Nominative Plural</th>
<th>Genitive Plural</th>
<th>Adessive Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘moon’</td>
<td>(a)</td>
<td>kuud</td>
<td>kuul</td>
<td>(b)</td>
<td>(c)</td>
<td>(d)</td>
</tr>
<tr>
<td>‘human’</td>
<td>inimese</td>
<td>(e)</td>
<td>(f)</td>
<td>(g)</td>
<td>inimeste</td>
<td>(h)</td>
</tr>
<tr>
<td>‘fish’</td>
<td>(i)</td>
<td>kala</td>
<td>(j)</td>
<td>kalad</td>
<td>(k)</td>
<td>(l)</td>
</tr>
</tbody>
</table>
Cebuano is an Austronesian language spoken in the Philippines. The language was heavily influenced by Spanish during a period of colonialism from 1521 to 1898.

Your task is to figure out what four women, Althea, Inday, Janelle, and Maria, had for lunch. Each person had a main dish as well as a dessert and a drink. No two people ordered the same main dish, no two people ordered the same dessert, and no two people ordered the same drink.

Each of them chose from the following Philippine menu:

<table>
<thead>
<tr>
<th>Main Dishes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>adobong baboy</strong></td>
</tr>
<tr>
<td><strong>adobong sitaw</strong></td>
</tr>
<tr>
<td><strong>asado</strong></td>
</tr>
<tr>
<td><strong>bulanglang</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Desserts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>buko pie</strong></td>
</tr>
<tr>
<td><strong>kutsinta</strong></td>
</tr>
<tr>
<td><strong>maruya</strong></td>
</tr>
<tr>
<td><strong>turon saba</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Drinks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bino</strong></td>
</tr>
<tr>
<td><strong>gatas</strong></td>
</tr>
<tr>
<td><strong>serbesa</strong></td>
</tr>
<tr>
<td><strong>tubig</strong></td>
</tr>
</tbody>
</table>

D1. Based on the following clues try to determine who ordered which food and drink. Write your answers in the Answer Sheets in Cebuano.

Janelle nikaon og bulanglang apan dili og turon saba.
Ang tao nga nipalit og asado ug nikaon og kutsinta wala niinom og serbesa.
Ang tao nga nikaon og adobong baboy ug niinom og tubig apan si dili Althea.
Ang duha ka tao nga nanginon alkohol mga Maria ug ang tao nga nipalit og adobong sitaw.
Inday nipalit og maruya.
Ang tolo ka tao nga wala nipalit og maruya mga Maria, Althea, ug Janelle.
Ang tao nga nikaon og bulanglang si Janelle.
Janelle nipalit og bulanglang apan dili og tubig.
(E) Be There, or Be Squared (1/1) [20 points]

The Huli language is a Trans-New Guinea language spoken in Papua New Guinea. Answer these questions in the Answer Sheets.

E1. The perfect squares from 1 to 100 are written in Huli below, in scrambled order. Perfect squares are the numbers 1, 4, 9, 16, 25, … that are the products of multiplying a whole number by itself (e.g. 4 = 2 * 2). On your answer sheet, write the corresponding numbers in Arabic numerals (e.g., the number twelve is written as 12 in Arabic numerals).

   a. ngui ki, ngui tebone-gonaga waragaria
   b. mbira
   c. ngui dau, ngui waragane-gonaga waragaria
   d. nguira-ni pira
   e. nguira-ni mbira
   f. dira
   g. maria
   h. ngui tebo, ngui mane-gonaga maria
   i. ngui ma, ngui dauni-gonaga maria
   j. ngui waraga, ngui kane-gonaga pira

E2. Four consecutive numbers are written in Huli below, in increasing order. Write the corresponding numbers in Arabic numerals.

   ngui ka, ngui haline-gonaga bearia
   ngui ka, ngui haline-gonaga hombearia
   ngui ka, ngui haline-gonaga haleria
   ngui ka, ngui haline-gonaga deria

E3. Write the following numbers in Huli.

   a. 2
   b. 4
   c. 6
   d. 7
   e. 22
   f. 44
   g. 66
   h. 77
   i. 88
   j. 173
Hittite is an extinct language that belongs to the Anatolian branch of the Indo-European language family. It was spoken in the ancient Hittite Empire in second millennium BCE. Hittite was written using a script, called cuneiform, composed of many wedge shapes; an example of cuneiform unrelated to this problem is below.

The excerpt below is a (simplified) phonetic rendering of a cuneiform passage found on a tablet. You do not need to know how the text is pronounced to solve this problem.

```
našta illuyankan
haṭtešnaz šarā kallišta
kāšawa ezenan iyami
nuwa adanna akuwanna eḫu

našta illuyankaš qadu dumumeššu
šarā úēr nuza eter ekuer
našta palḥan ḫūmandan ekuer
neza ninkēr

ne namma ḫattešnaš kattanta
nūmān pānzi ḫupašiyašša úit
nu illuyankan išhimanta kalēliēt

imaš úit nukān illuyankan
kuenta dingirmesša kattīšši ešer
```

Its translation into English:

And he called up the snake from the hole: “Behold the feast I’m making! Come to eat and to drink!” And the snake came up with his sons. And they ate and drank. And they drank all the kettles. And they could no longer go down into the hole again. And Hupasiyas came and tied the snake with a rope. The Stormgod came and killed the snake; and the gods were with him.”

Answer the questions on the next page in the Answer Sheets provided.
(F) Take One Tablet and Call Me in the Morning (2/2)

F1. Match the following Hittite word forms with their English translations.

a. eter  
   (1) snake

b. ḥatteššar  
   (2) hole

c. ūit  
   (3) came

d. illuyankaš  
   (4) and

e. našta  
   (5) up

f. šarā  
   (6) drank

g. ekuer  
   (7) ate

F2. Match the following suffixes with their grammatical roles.

a. –aš  
   (1) marker of the infinitive (infinitives are verbs translated with to, such as “to sleep” or “to walk”)

b. –ša  
   (2) plural marker

c. –meš  
   (3) marker of 3rd person plural past tense verbs (these are past tense verbs with a subject of they or a plural noun, such as walked in “they walked” or drank in “the cats drank milk”)

d. –er/ēr  
   (4) marker of the direct object (the direct object is the recipient of the action, such as him in “she hit him” or a pizza in “he made a pizza”)

e. –a  
   (5) marker of the subject (the subject is the entity performing the action, such as the dog in “the dog chased the cat”)

f. –an  
   (6) marker for 3rd person singular past tense verbs (these are past tense verbs with a subject of he, she, it, or a singular noun, such as walked in “he walked” or drank in “the cat drank milk”)

g. –it  
   (7) marker denoting “and”
The Cistercians are an order of Christian monks still active today, who, for reasons including vows of silence, have developed rudimentary sign languages. Below are several words in the Cistercian Sign Language (each group of signs below is one word - each individual sign is labeled with a number) in arbitrary order, and their translations in English.¹

<table>
<thead>
<tr>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(D)</th>
<th>(E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>31</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(F)</th>
<th>(G)</th>
<th>(H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>20</td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(I)</th>
<th>(J)</th>
<th>(K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹The particular variety of Cistercian sign language represented here is that of a monastery in the U.S.; signs in other communities may vary from those presented here.
(G) Signs from Above (3/3)

Answer these questions in the Answer Sheets.

**G1.** Determine the correct correspondences. For each part, write the capital letter corresponding to a Cistercian Sign Language word.

<table>
<thead>
<tr>
<th>a. apple</th>
<th>i. a Cistercian monk</th>
<th>q. milk</th>
</tr>
</thead>
<tbody>
<tr>
<td>b. barn</td>
<td>j. Cistercians</td>
<td>r. a nun</td>
</tr>
<tr>
<td>c. bathroom</td>
<td>k. dormitory</td>
<td>s. poetry</td>
</tr>
<tr>
<td>d. Benedictines</td>
<td>l. (to) drink</td>
<td>t. queen bee</td>
</tr>
<tr>
<td>e. the Blessed Sacrament</td>
<td>m. England</td>
<td>u. snow</td>
</tr>
<tr>
<td>f. cake</td>
<td>n. ice</td>
<td>v. tree</td>
</tr>
<tr>
<td>g. chocolate milk</td>
<td>o. Iceland</td>
<td>w. wooden table</td>
</tr>
<tr>
<td>h. Christmas</td>
<td>p. Italy</td>
<td></td>
</tr>
</tbody>
</table>

**G2.** Translate the following into Cistercian Sign Language (for each word, the answer will be a single sign. Write the number of that sign in your Answer Sheet).

a. baby
b. (to) pour
c. rain
d. tea

---

2The Benedictines are another order of Christian monks (see pictures above on this page).
3The Blessed Sacrament is a term used to refer to the bread used in a particular ritual.
4Christmas is a holiday celebrating the birth of the Christian figure Jesus.
MARY SU.0 is a fan-fiction writing robot. Fan fiction is a fiction written by people using another author's characters. Unfortunately, she’s not very good at what she does. MARY writes fan-fiction by reading the text of a book (or series of books) and randomly generating new sentences based on the text. Her latest effort is fan-fiction based on the Harry Potter book series.

MARY SU.0 has a few different methods that she’s able to use for generating sentences. The first class of methods are called n-gram methods. The simplest of these methods is the unigram method. In the unigram method, MARY chooses each token of the sentence completely randomly from the entire vocabulary of the book she read. (A token can also be a punctuation mark.) An example of a sentence generated using this method might look like this:

gave spiral the truly poisoned, Neville the shoulder Invisibility

A second method is the bigram method. In this method, MARY first finds all the tokens that were used to start a sentence in the text and randomly chooses one of these to start the sentence. Then she builds the rest of the sentence by looking at the most recent token generated, finding all tokens that occur immediately after that token in the text, and randomly choosing one of these. For example, if the most recently generated token was “red”, MARY would find all the tokens in the text that immediately follow “red”, {“hair”, “curtains”, “as”, …} and randomly choose one of these to be the next word. A sentence generated using the bigram method might look like this:

Face your nose noisily after you saying stuff.

A third method is called the trigram method. This method is very similar to the bigram method, but uses the previous two tokens (instead of the previous one) to decide what the next token will be. A sentence generated using the trigram method might look like this:

But Harry hardly noticed that six extra chairs.”

The last method that MARY can use to generate sentences is called the Context Free method. This method starts by taking each sentence in the text and generating a grammar tree, like the one below, for it.
The symbols that aren’t words refer to labels of words or larger sequences. Some symbols refer to parts of speech, such as NNP for proper noun, PRP for personal pronoun, VBD for a verb in past tense, and TO for preposition. Other labels refer to sequences of words that form units, such as S for sentence, NP for a noun phrase, a sequence of one or more words that behaves like a noun (e.g. dogs or the big dogs), and VP for a verb phrase, which is a sequence of one or more words that behaves like a verb (e.g. goes or went to the store).

To generate a new sentence, she first generates an “S” which represents a sentence. Then she looks through her collection of grammar trees for all the sets of symbols ([NP VP .] for example) that occur immediately under an “S”. She then repeats this process recursively for each of the new items generated until the tree has no more nodes that can be expanded (once a token is generated, it cannot be expanded). A sentence generated by this method might look like this:

The next question will cast by Ron.

H1. Below is a collection of sentences. Two of them are real sentences from the Harry Potter series. The rest were generated using one of the methods above; each method generated at least two sentences. In the answer sheets, write either “u” for unigram, “b” for bigram, “t” for trigram, or “c” for context-free to indicate the method that most likely generated that sentence, or if you think the sentence was not automatically generated, write “r” for real.

a. Headmaster uninjured could that was Malfoy that badges
b. He bent over top of the water blushing furiously.
c. There were crouching in your bedroom.
d. He lived about a hundred wizards were closing.
e. Ron spooned iron bolts, keyholes, and a heavy wooden breadboard on to her back and picked up a fistful.
f. "What?" said Harry.
g. 'Sorry!' he said," said Mr. Malfoy's eyes.
h. Harry wasn't," said Dumbledore went slightly surprised.
i. years beginning at to annoyance spider!" just months Harry
j. You might have been an impostor.
k. They'll be the first to rise up in the Invisibility Cloak on," said Professor Flitwick pressed a box into his bag.
l. The broom gave them an enormous wink.
Contest Booklet

Name: ___________________________________________

Contest Site: ________________________________________

Site ID:  ____________________________________________

City, State: _________________________________________

Grade:  ______

Start Time: _________________________________________
End Time: __________________________________________

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 WEBSITE IN APRIL.

Signature:    __________________________________________________
(A) Tangled Up In Nots
1. a. 
   b. 
   c. 
   d. 
   e. 

(B) DAWG Breeds
1. a. 
   b. 
   c. 
   d. 
   e. 
2. a.  
   b.  
   c.  

(C) The Curious Case of Estonian
1. 
   
2. a.  
   b.  
   c.  
   d.  
   e.  
   f.  

n a c l o
(C) The Curious Case of Estonian (cont.)

g. 

h. 

i. 

j. 

k. 

l. 

(D) That’s an Order!

1. Althea  Inday  Janelle  Maria

   Main Dish

   Dessert

   Drink

(E) Be There, or Be Squared

1. a. 

   b. 

   c. 

   d. 

   e. 

   f. 

   g. 

   h. 

   i. 

   j. 

2. 

3. a. 

   b. 

   c. 

   d. 

   e. 

   f. 

   g. 

   h. 

n a c l o
(E) Be There, or Be Squared (cont.)

i. 

j. 

(F) Take One Tablet and Call Me in the Morning

1. 

2. 

(G) Signs from Above

1. 

2. 

(H) Fan Fiction

1. 

2. 

n a c l o