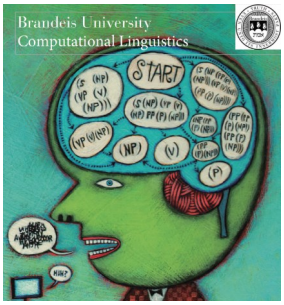
naclo

Carnegie Mellon University
Language Technologies Institute



Yale University

The Fourteenth Annual

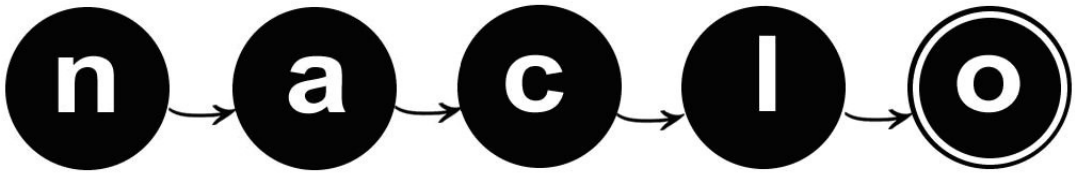
North American Computational Linguistics Open Competition 2020

www.nacloweb.org

Open Round
January 23, 2020

Serious language puzzles that are surprisingly fun!

-Will Shortz, Crossword editor of The New York Times and Puzzlemaster for NPR



Welcome to the fourteenth annual North American Computational Linguistics Open Competition! 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 & 11.
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 United States and Anglophone Canada in the open round will be invited to the second round.
9. 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 2020 Organizers

Program Committee:

Adam Hesterberg — Massachusetts Institute of Technology
Aleka Blackwell — Middle Tennessee State University
Ali Sharman — University of Michigan
Andrés Salanova — University of Ottawa
Annie Zhu — Harvard University
Babette Verhoeven — Aquinas College
Ben LaFond — Harvard University
Corinne Soucy — Université de Montréal
Daniel Lovsted — McGill University
David Mortensen — Carnegie Mellon University
Dick Hudson — University College London
Dragomir Radev — Yale University
Elisabeth Mayer — Australian National University
Ethan Chi — Stanford University
Harold Somers — All Ireland Linguistics Olympiad
Heather Newell — UQAM
James Pustejovsky — Brandeis University
James Hyett — University of Toronto
Ji Hun Wang — Stanford University
Kevin Liang — University of Pennsylvania
Lori Levin — Carnegie Mellon University
Margarita Misirpashayeva — Massachusetts Institute of Technology
Mary Laughren — University of Queensland
Matthew Gardner — Carnegie Mellon University
Mihir Singhal — Massachusetts Institute of Technology
Oliver Sayeed — University of Pennsylvania
Patrick Littell — University of British Columbia
Pranav Krishna — Massachusetts Institute of Technology
Shuli Jones — Massachusetts Institute of Technology
Sonia Reilly — Massachusetts Institute of Technology
Tom Roberts — University of California, Santa Cruz

NACLO 2020 Organizers

Problem Credits:

Problem A: Kevin Liang
Problem B: Aleka Blackwell
Problem C: Ethan Chi
Problem D: Ethan Chi
Problem E: Michael Salter
Problem F: Sam Ahmed
Problem G: Tom McCoy
Problem H: Michael Salter

Organizing Committee:

Adam Hesterberg — Massachusetts Institute of Technology
Aleka Blackwell — Middle Tennessee State University
Ali Sharman — University of Michigan
Andrés Salanova — University of Ottawa
Annie Zhu — Harvard University
Ben LaFond — Harvard University
Brian Xiao — Massachusetts Institute of Technology
Corinne Soucy — Université de Montréal
Daniel Lovsted — McGill University
David Mortensen — Carnegie Mellon University
Dragomir Radev — Yale University
Ethan Chi — Stanford University
James Pustejovsky — Brandeis University
Ji Hun Wang — Stanford University
Kevin Liang — University of Pennsylvania
Lori Levin — Carnegie Mellon University
Margarita Misirpashayeva — Massachusetts Institute of Technology
Matthew Gardner — Carnegie Mellon University
Mihir Singhal — Massachusetts Institute of Technology
Patrick Littell — University of British Columbia
Pranav Krishna — Massachusetts Institute of Technology
Ryan Guan — Stanford University
Shuli Jones — Massachusetts Institute of Technology
Sonia Reilly — Massachusetts Institute of Technology
Tom Roberts — University of California, Santa Cruz

NACLO 2020 Organizers (cont'd)

US Team Coaches:

Aleka Blackwell, Middle Tennessee State University
Dragomir Radev, Yale University
Lori Levin, Carnegie Mellon University

Canadian Coordinators and Team Coaches:

Daniel Lovsted, McGill University

USA Contest Site Coordinators:

Austin Math Circle: Joshua Dever
Brandeis University: James Pustejovsky, Sarah Irwin, Nikhil Krishnaswamy
Brigham Young University: Deryle Lonsdale
California State University: Michael Ahland
Carnegie Mellon University: Mary Jo Bensasi, Lori Levin
College of William and Mary: Dan Parker
Columbia University: Kathy McKeown, Brianne Cortese, Smaranda Muresan
Cornell University: Abby Cohn, Sam Tilsen
Emory University: Jinho Choi, Phillip Wolff
Georgetown University: Emma Manning
Indiana University: Sandra Kuebler
Johns Hopkins University: Rebecca Knowles, Pamela Shapiro, Tom McCoy
Kansas State University: Natasha Rojkovskaia
Keiser University, Jacksonville: Deborah Williams
Long Island University Post: Rachel Szekely
Massachusetts Institute of Technology: Margarita Misirpashayeva, Sonia Reilly
Middle Tennessee State University: Aleka Blackwell
Minnesota State University Mankato: Rebecca Bates, Dean Kelley, Louise Chan
Montclair State University: Anna Feldman
Northeastern Illinois University: Ariana Bancu, Lewis Gebhardt
Northwestern University: Alaina Arthurs
Ohio State University: Michael White, Micha Elsner, Marie de Marneffe
Oregon Mathletics: Shirley Ma
Oregon State University: Liang Huang
Princeton University: Christiane Fellbaum, Eve Fleisig, Rishi Mago, Abinitha Gourabathina

NACLO 2020 Organizers (cont'd)

USA Contest Site Coordinators (cont'd):

San Diego State University: Rob Malouf

Southern Illinois University: Vicki Carstens, Jeffrey Punske

SpringLight Education Institute: Sherry Wang

Stanford University: Daisy Leigh, Eli Strauss-Reis, Matthew Early

Stony Brook University: Lori Repetti, Sarena Romano, Jeffrey Heinz

UCLA: Sue Han, Lily Kawaoto

Union College: Kristina Striegnitz, Nick Webb

University of California, Irvine: Sameer Singh, Zhengli Zhao, Kristen Salsbury

University of Colorado at Boulder: Silva Chang

University of Houston: Thamar Solorio, Deepti Bhat, Giovanni Molina

University of Illinois at Urbana-Champaign: Benjamin Leff, Greg Sigalov, Julia Hockenmaier

University of Kentucky: Kevin McGowan, Andrew Byrd

University of Maryland: Kasia Hitczenko, Laurel Perkins, Aaron Doliana, Sigwan Thivierge, Nadiya Klymenko

University of Massachusetts, Lowell: Anna Rumshisky, David Donahue, Willie Boag

University of Massachusetts: Andrew Lamont, Anastasia Chobany

University of Memphis: Vasile Rus

University of Michigan: Steven Abney, Sally Thomason, Marcus Berger

University of Nebraska, Omaha: Parvathi Chundi

University of North Carolina Charlotte: Wlodek Zadrozny, Hossein Hematialam, Kodzo Wegba

University of North Texas: Rodney Nielsen

University of Notre Dame: David Chiang

University of Pennsylvania: Chris Callison-Burch, Mitch Marcus, Cheryl Hickey, Oliver Sayeed, Marianna Apidianaki, Derry Wijaya, Anne Cocos

University of Southern California, ISI campus: Jon May, Nima Pourdamghani

University of Texas at Dallas: Vincent Ng, Jing Lu, Gerardo Ocampo Diaz

University of Utah: Brendan Terry, Jessica Larsen, Justin Nistler, Joselyn Rodriguez

University of Washington: Jim Hoard, Joyce Parvi

University of Wisconsin, Milwaukee: Joyce Boyland, Hanyon Park, Gabriella Pinter, Anne Pycha

Western Washington University: Kristin Denham

Yale University: Raffaella Zanuttini

NACLO 2020 Organizers (cont'd)

Canada Contest Site Coordinators:

Dalhousie University: Vlado Keselj, Dijana Kosmajac

Simon Fraser University: Ashley Farris-Trimble

University of British Columbia: Yadong Liu, Jozina Vander Klok

University of Ottawa: Andrés Pablo Salanova

University of Toronto: Minh-Tam Nguyen, Lola Bradford

University of Western Ontario: Janis Chang

University of Calgary: Dennis Storoshenko

University of Alberta: Herbert Colston, Sally Rice

McGill University: Lisa Travis

NACLO 2020 Organizers (cont'd)

Booklet Editors:

Dragomir Radev, Yale University

Ethan Chi, Stanford University

Ji Hun Wang, Stanford University

Sponsorship Chair:

James Pustejovsky, Brandeis University

Sustaining Donors:

Linguistic Society of America

North American Chapter of the Association for Computational Linguistics

National Science Foundation

University Donors:

Brandeis University

Carnegie Mellon University

Massachusetts Institute of Technology

Middle Tennessee State University

Yale University

Many generous individual donors

Special thanks to:

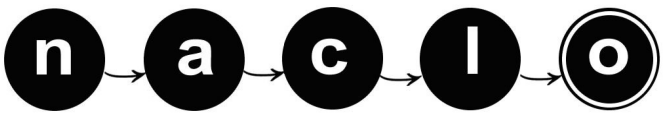
The hosts of the 200+ High School Sites

(A) Let That Mongo! (1/2) [5 Points]

Mongo (also known as Nkundo or Lomongo) is a Bantu language spoken by the Mongo Peoples of the central Democratic Republic of the Congo. Presently, there are around 400,000 native speakers spread out over a large area around the Congo River.

Below is a table showing a few verb conjugations in Lomongo. [d͡ʒ] is a consonant pronounced like the [j] in the English word *jump*. [ŋ] is a consonant pronounced like the [ng] at the end of the English word *sing*.

Imperative	***2nd singular***	***3rd singular***	***3rd plural***	***English***
bota	oota	aota	baota	'beget' (to give rise to; to bring about)
kamba	okamba	akamba	bakamba	'work'
imed͡ʒa	wimed͡ʒa	imed͡ʒa	bimed͡ʒa	'consent'
usa	wusa	usa	busa	'throw'
bata	oata	aata	baata	'get'
ɛna	wɛna	ɛna	bɛna	'see'
isa	wisa	isa	bisa	'hide'
d͡ʒila	od͡ʒila	ad͡ʒila	bad͡ʒila	'wait'
ina	wina	ina	bina	'hate'
bina	oina	aina	baina	'dance'
asa	wasa	asa	basa	'search'
saŋga	osaŋga	asaŋga	basaŋga	'say'



(A) Let That Mongo! (2/2)

A1. Explain how these Mongo verb forms work by filling in the blanks below in the Answer Sheets.

Each Mongo verb has a root form. The 4 verb forms shown here are formed by adding a prefix before the root form. The prefix for the imperative form is ___(1)___, the prefix for the 2nd singular form is ___(2)___, the prefix for the 3rd singular form is ___(3)___, and the prefix for the 3rd plural form is ___(4)___. (Note that some of these prefixes may be empty; to note this, write the symbol Ø-).

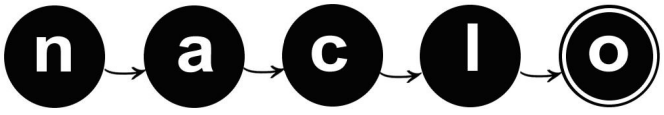
However, we are not done yet: to get the final verb form, we must apply some sound change rules. The relevant rules are:

1. If there are two ___(5)___, in a row, delete ___(6)___.
2. Delete ___(7)___ when it appears between two ___(8)___.
3. Change ___(9)___ to ___(10)___ when it appears before a ___(11)___.

There is one final wrinkle: the order that these rules are applied in matters. The rules must be applied in this order: First apply rule ___(12)___, then rule ___(13)___, then rule ___(14)___.

A2. Fill in the blanks (1-19) in the table below. Record your answers in the Answer Sheets.

Imperative	***2nd singular***	***3rd singular***	***3rd plural***	***English***
bakisa	(1)	(2)	(3)	‘add’
(4)	wanda	(5)	(6)	‘begin’
solola	(7)	(8)	basolola	‘chat’
ponama	(9)	aponama	(10)	‘elect’
(11)	oowa	(12)	(13)	‘cure’
(14)	(15)	aalusa	(16)	‘turn’
loŋga	(17)	(18)	(19)	‘win’



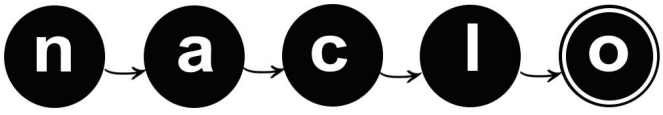
(B) A Cat in a Hat (1/2) [15 Points]

Chintang (Chintang: छिन्ताङ् Chintāṅ / *Chhintang*) is an eastern Kiranti language spoken by 5,000 to 6,000 people in Chhintang and Ahale municipalities of Dhankuta District, Koshi Zone, Nepal. Its dialects are Mulgaun and Sambhugaon. A few of the characters used to represent the words in this language may be unfamiliar. ʔ is a glottal stop, the sound heard in the middle of "uh-oh." ŋ is the sound made by the *ng* in *sing*. ʌ is the sound made by the *o* in *won*.

Below are 16 sentences in Chintang written in the International Phonetic Alphabet, and their unordered English translations:

1.	*cuwa uthurumbeʔ yuŋno*	a.	There is a hole in the towel.
2.	*appa chintaŋbeʔ yuŋno*	b.	The woman has gone away.
3.	*sencak sie*	c.	The rice has been cooked.
4.	*wapaŋa topi wadaŋse*	d.	There is a hat on the head.
5.	*kok thuktaŋse*	e.	My mother-in-law slept.
6.	*ram harinɨŋ khoŋno*	f.	You go to the market.
7.	*kʌp kedadaŋse*	g.	The water is in his mouth.
8.	*taŋbeʔ topi yuŋno*	h.	Joge sent a letter.
9.	*menwaŋa sencak sede*	i.	The mouse died.
10.	*tawelbeʔ uhoŋ yuŋno*	j.	The woman has bought a chicken.
11.	*anambaŋa cuwa thuŋno*	k.	The cup has been broken.
12.	*mechacha khadaŋse*	l.	My father is in Chintang.
13.	*jogeŋa citthi hakte*	m.	The rooster has put on a hat.
14.	*anamma imse*	n.	A cat killed a mouse.
15.	*hana bajar akhaʔno*	o.	Ram plays with Hari.
16.	*mechachaŋa wa khedoŋse*	p.	My father-in-law drinks water.

B1. Match the Chintang sentences (1-16) with their corresponding English translation (a-p). Record your answers in the Answer Sheets.



(B) A Cat in a Hat (2/2)

B2. Translate the following into Chintang. Note that some English words may be translated as part of a Chintang word. Record your answers in the Answer Sheets.

a. There is a cat in a hat.

b. Joge plays with a chicken.

c. My hen is in the market.

d. parent-in-law

B3. Translate the following into English. Your answers should be structurally similar to the English translations given in the data. Record your answers in the Answer Sheets.

a. *athurumbeʔ kok yuŋno*

b. *appa khade*

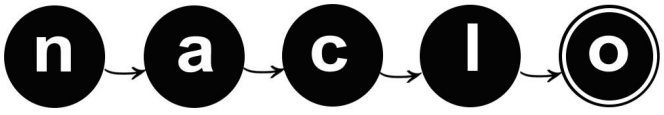
c. *anamma ammanɨŋ yuŋno*

(C) Set in Stone (1/2) [20 Points]

Old Persian, one of the two attested Old Iranian languages, was spoken from 600-300 BCE in Achaemenid Persia. Old Persian was written in *cuneiform*, a writing system produced using wedge-shaped marks in clay tablets inherited from the Sumerian Empire.

Below are some words written in Old Persian Cuneiform. On the next page are their transcriptions and English translations in no particular order. Note that one word can be written in two different ways in cuneiform!

1.
2. 
3.
4. 
5.
6. 
7.
8. 
9.
10.
11. 
12.
13. 
14.
15. 
16.
17. 
18.
19. 
20.
21. 
22.
23.
24. 
25.
26. 



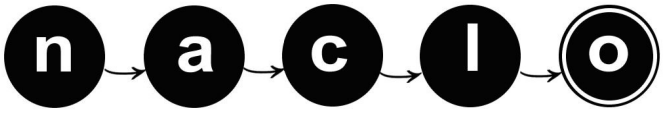
(C) Set in Stone (2/2)

Here are the transcriptions and English translations of the words on the previous page in no particular order. Note that y, š and č are consonants; ā, ī, and ū are long versions of their corresponding vowels (a, i, and u). Capitalization in transcription is not reflected in the Persian script.

A.	*paruvam*	previously	N.	*Kuruš*	Cyrus
B.	*hadugā*	record, statue	O.	*asmā*	sky
C.	*āyadana*	sanctuary	P.	*Bagadāta*	(a male name)[1]
D.	*dāruv*	wood	Q.	*Hiduš*	India
E.	*duruxta*	lie	R.	*hača*	from
F.	*Čišpaiš*	Tespes (a king of Persia)	S.	*bājim*	tribute, toll
G.	*saiymam*	silver	T.	*Arabāya*	Arabia
H.	*bar*	to bear	U.	*Haraiva*	(a female name)
I.	*Skudra*	Thrace	V.	*daiva*	false god
J.	*baga*	god	W.	*Ūvja*	Elam (a place name)
K.	*pasāva*	after	X.	*hakaramčīy*	once
L.	*Ākaufačiyā*	a tribe of southeastern Iran	Y.	*Dārayauš*	Darius (a king of Persia)
M.	*radīy*	because of			

C1. Determine the correct correspondences. Record your answers in the Answer Sheets.

[1] Equivalent to the Greek name *Theodoros* ("God-given")



(D) Pay Attention (1/2) [10 Points]

The meaning of a word depends on its context. For example, in the sentence "The farmer *seeded* the field with corn," the word *seeded* means "added seeds to." However, in the sentence "The chef seeded the tomato," the word *seeded* means "took seeds away from."

If you were building a model of language, how would you get it to to recognize the way that a word's meaning depends on context? One popular technique for achieving this goal is a mechanism called **attention**. In the way that attention is implemented in current state-of-the-art models of language, the model has a large number of **attention heads**, each of which is denoted by a pair of numbers (for examples, 8-10). When the model processes a sentence, for every pair of words in the sentence, each head calculates the "relatedness" of the two words.

The one wrinkle is that we do not know what exactly "relatedness" should mean, so instead of telling the model how to define "relatedness," we let the model learn its own definition of relatedness. Recently, computer scientists have started to analyze what these attention heads have learned, and this analysis shows that they often reflect linguistic information! For example, here's the output of one attention head (head 8-10) when we feed the following sentence into BERT, which is one of the most popular models that uses attention heads:

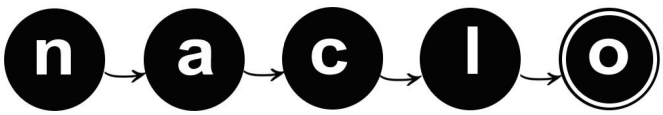
Example: I_1 see_2 my_3 $sister_4$, but_5 she_6 $can't_7$ see_8 me_9 $because_{10}$ she_{11} is_{12} $reading_{13}$ a_{14} $linguistics_{15}$ $book_{16}$.

Output: $4 \rightarrow 2$, $9 \rightarrow 8$, $16 \rightarrow 13$

This output signifies that head 8-10 connects word #4 (*sister*) to word #2 (*see*), as well as word #9 (*me*) to word #8 (*see*) and word #16 (*book*) to word #13 (*reading*). If you consider what all of those pairs of words have in common, you'll see that each one is a verb and its direct object: *sister* is the direct object of the first instance of *see*, *me* is the direct object of the second instance of *see*, and *book* is the direct object of *reading*. It appears that head 8-10 has learn to connect verbs to their objects! (Note that these activations are directional; for example, word #2 is not connected to word #4.)

Why is this information useful? If we go back to the example with the verb seed, this sort of information can help the model figure out which version of *seed* is being used: If its direct object is something like *field* or *lawn*, then it probably means "to add seeds to;" if its direct object is something like *tomato* or *watermelon*, then it probably means "to take seeds away from." Of course, one sentence isn't enough to draw strong conclusions. Instead, computer scientists tend to use a *corpus*, or a database, of example sentences to find patterns in the data. On the next page is a small corpus, the NacloWeb Corpus[1], which has 7 sentences.

[1] Some sentences derived from data in the English Web Treebank.



(D) Pay Attention (2/2)

NacloWeb Corpus

1. *My$_1$ experience$_2$ with$_3$ Gelda$_4$'s$_5$ House$_6$ of$_7$ Gelbelgarg$_8$ has$_9$ been$_{10}$ extremely$_{11}$ wonderful$_{12}$*
2. *We$_1$ use$_2$ Google$_3$'s$_4$ models$_5$ to$_6$ delve$_7$ into$_8$ the$_9$ inner$_{10}$ workings$_{11}$ of$_{12}$ language$_{13}$*
3. *At$_1$ this$_2$ corporation$_3$'s$_4$ meeting$_5$ people$_6$ are$_7$ concerned$_8$ about$_9$ the$_{10}$ company$_{11}$'s$_{12}$ plans$_{13}$*
4. *In$_1$ July$_2$ we$_3$ will$_4$ interview$_5$ the$_6$ candidate$_7$ and$_8$ review$_9$ her$_{10}$ resumé$_{11}$ again$_{12}$*
5. *The$_1$ platypus$_2$ is$_3$ a$_4$ strange$_5$ animal,$_6$ with$_7$ its$_8$ eggs$_9$ and$_{10}$ its$_{11}$ webbed$_{12}$ feet$_{13}$*
6. *I$_1$ think$_2$ that$_3$ although$_4$ my$_5$ NACLO$_6$ exam$_7$ was$_8$ difficult,$_9$ it$_{10}$ was$_{11}$ a$_{12}$ lot$_{13}$ of$_{14}$ fun$_{15}$*
7. *Linguistics$_1$ is$_2$ a$_3$ beautiful$_4$ science$_5$ that$_6$ provides$_7$ interdisciplinary$_8$ insight$_9$ into$_{10}$ the$_{11}$ human$_{12}$ experience$_{13}$*

Note that the NacloWeb corpus treats the possessive clitic *'s* as a separate word. (So in Sentence #1, word #5 is *'s* and word #6 is *House*.)

In our experiment on the NacloWeb Corpus, we ran each of the corpus' sentences through BERT and recorded the outputs of four attention heads (8-11, 7-6, 9-6, and 5-4). Unfortunately, due to some extremely sloppy experimental procedure, we don't remember which order we ran them through the model; in addition, we forgot to record some data. Your job is to fill in the blanks! Note that some blanks may have more than one connection, and some may have none at all.

Sentence	**8-11**	**7-6**	**9-6**	**5-4**
Sentence A	12 → 13	5 → 6	14 → 13	10 → 6
Sentence B	**(a)**	8 → 9, 11 → 12	None	8 → 2, 11 → 2
Sentence C	2 → 3, 10 → 11	4 → 5, 12 → 13	**(b)**	11 → 3
Sentence D	**(c)**	4 → 5	12 → 13	None
Sentence E	3 → 4, 11 → 12	**(d)**	10 → 12	5 → 1
Sentence F	**(e)**	1 → 2, 5 → 6	7 → 8	None
Sentence G	**(f)**	**(g)**	**(h)**	**(i)**

D1. Identify sentences A-G. Record your answers in the Answer Sheets.

D2. Fill in the missing data, labelled **(a)** to **(i)**, in the table above. Record your answers in the Answer Sheets.

(E) Breton Numbers (1/1) [15 Points]

Breton is a language spoken by approximately 200,000 people in Brittany, France. Part of the Celtic family of the Indo-European languages, it is distantly related to English, as well as other European languages such as French and Russian.

Below are some equations in Breton:

trizek + daouzek = pemp warn ugent

unan ha pevar-ugent ÷ nav = nav

pevar ha tri-ugent − ugent = pevar ha daou-ugent

seizh warn ugent + pevarzek = unan ha daou-ugent

daou × seizh = pevarzek

kant ÷ daou = hanter kant

nav × c'hwec'h = pevar ha hanter kant

c'hwec'h ha tri-ugent ÷ tri = daou warn ugent

c'hwezek × c'hwec'h = c'hwezek ha pevar-ugent

daouzek × pemp = tri-ugent

E1. In the Answer Sheets, fill in the gaps in the following equations.

a. **pevar-ugent ÷ pemp = _____**

b. **pemp ha hanter kant − daouzek = _____**

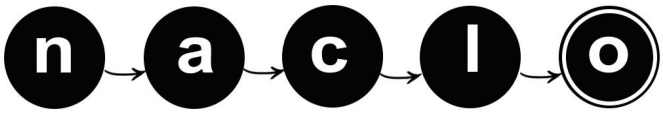
c. **nav warn ugent + _____ = tri ha daou-ugent**

d. **kant ÷ _____ = pemp**

E2. As well as an interesting numbering system, Breton has an unusual method of forming the plurals of some nouns, which is partly connected to the numbering system. Here are three plural nouns in Breton: which of them means "(one person's) eyes"? Record your answer in the Answer Sheets.

(a) **elerc'h** (b) **daoulagad** (c) **perennoù**

E3. An "irregular" number in the Breton system is *triwec'h*. It represents a number between 10 and 20. Which one? Record your answer in the Answer Sheets.



(F) Coming into Focus (1/2) [15 Points]

Paiwan is an Austronesian language spoken by around 66,000 people in southern Taiwan. One of the major components of Paiwan sentence structure is "focus," which marks a new piece of information conveyed by the sentence. In English, we might represent this through phrasing, such as in sentence "It is ***the dog*** which the man likes". In this example, ***the dog*** is focused.

Below are some sentences in Paiwan with their English translations. Sentence elements in ***italics*** are focused.

Paiwan	**English**
1. *kanan nua uqaɬay tua kuka a quma nua tsakaw*	The man eats the chicken in ***the thief's field***.
2. *tjmalaw tua vavuy a kuvatu*	***My dog*** angers the pig.
3. *djavisen nua tsemas a kukama*	The spirit snatches ***my father***.
4. *sitarang nua uqaɬay a vuluq*	The man protects (it) with ***the spear***.
5. *tjalawen nua suvavuy i tua umaq a aɬak nua vavaian*	Your pig angers ***the woman's child*** in the house.
6. *kman tua kuka a tsemas*	***The spirit*** eats the chicken.
7. *djavisan nua pulingaw tua vuluq a gadu*	The shaman snatches the spear in ***the mountain***.
8. *langedaen nua sivitay a qaya-qayam*	The soldier hears ***the bird***.

F1. How would you say these sentences in Paiwan? Record your answers in the Answer Sheets.

a. The man protects the field with ***the dog***.

b. Your shaman angers my bird in ***the spirit's mountain***.

c. ***The woman*** snatches the pig.

(F) Coming into Focus (2/2)

Now look at the following question-answer dialogues. The first answer has been translated for you.

Question	Answer
9. inu a tmarang a tsakaw?	*tmarang a tsakaw i tua quma* The thief who protects (it) is in the field.
10. anema a sitjalaw nua sukama tua vatu?	*sitjalaw nua kukama tua vatu a kuvuluq*
11. anema a sudjavisen?	*kudjavisen a qaya-qayam nua vavaian*
12. inu a kanan nua uqaɬay?	*kanan nua uqaɬay a gadu*

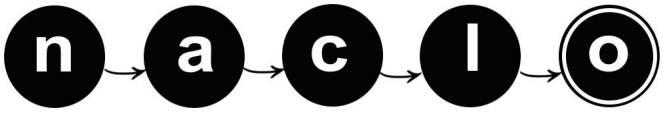
F2. Based on these examples, translate the following responses into English, underlining focused elements, and saying what questions (in Paiwan) could have prompted them. Record your answers in the Answer Sheets.

a. *susitarang tua qaya-qayam nua pulingaw a tsemas*

b. *kanen nua uqaɬay a vavuy i tua gadu*

c. *kutjalawan tua suvatu a kuquma*

*(The letter pairs **dj** and **tj** each represent a single sound. The word **langedaen** in sentence 8 is actually **langedain**, but this was edited for the sake of simplicity.)*



(G) Password Confusion (1/3) [10 Points]

Mary, Larry, and Harry are three friends sharing an apartment. Unfortunately, their nosy neighbor Perry is always trying to use their wifi, so Mary changes their wifi password frequently to thwart Perry's efforts. Whenever she changes the password, Mary texts the new password to Larry and Harry.

One day, Mary opens their wifi bill and can tell from the exorbitant charge that Perry has been using it again. Since this can only mean that Perry is somehow reading their texts, Mary hires a company called the Rearranging Expressions Organization (or REOrganization for short) to give her advice on making her messages more secure. REOrganization advises her to send messages that can be decoded by the following 2-step process:

1. Convert the sentence to a question;
2. Read the first letter of each word in the question to yield the password.

As it turns out, Larry and Harry are not very inquisitive people, so they have never encountered questions before. Therefore, Mary gives them the following example messages to show how the system works:

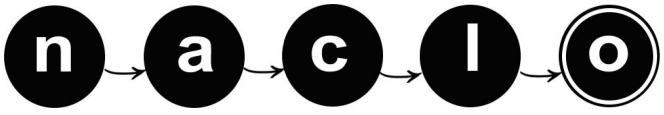
Text message	Text message converted to a question	Decoded password
Deciphering each Alaskan license plate abbreviation is lovely.	Is deciphering each Alaskan license plate abbreviation lovely?	IDEAL PAL
A new German exchange rate beneath a newspaper did assist national ambassadors.	Did a new German exchange rate beneath a newspaper assist national ambassadors?	DANGER BANANA
Every industrial geographer hired through your legal action was yelling enthusiastic random stuff.	Was every industrial geographer hired through your legal action yelling enthusiastic random stuff?	WEIGHTY LAYERS

Why have this extra step of forming a question? The idea is that the initial letters from the original text messages can also spell two-word chunks (DEAL PAIL, ANGER BANDANA, and EIGHTY LAWYERS), so perhaps Perry will be thrown off the scent by these distractor phrases.

(G) Password Confusion (2/3)

G1. Larry and Harry both seemed to understand the examples Mary showed them, so she began using this system to encode the new passwords. It went swimmingly for the first few weeks, but then Larry and Harry occasionally began to get the password wrong. Below are all of the messages Mary sent out, along with the passwords that her roommates extracted from the messages. A few cells of the table have been left blank; in the Answer Sheets, fill them in.

Text message	Distractor password	Larry's password guess	Harry's password guess	Correct password
Uplifting tales told extremely rapidly in New Guinea and Romania may reveal each storyteller's trade secrets.	UTTERING ARMRESTS	MUTTERING ARRESTS	MUTTERING ARRESTS	MUTTERING ARRESTS
Lively orangutans using discarded branches as tambourines can harmonize.	LOUD BATCH	CLOUD BATH	CLOUD BATH	CLOUD BATH
Horned owls should endure every modern orange tree iguanas can offer next summer.	HOSE EMOTICONS	CHOSE EMOTIONS	SHOE EMOTICONS	SHOE EMOTICONS
Every loud electronic creature that interesting odd numbers should provide at reunions should escape.	ELECTION SPARSE	SELECTION SPARE	SELECTION PARSE	SELECTION SPARE
Alligators that can launch airplanes may prosper.	AT CLAMP	MAT CLAP	CAT LAMP	MAT CLAP
Unions should identify novel geometric systems that one may possibly seek.	USING STOMPS	MUSING STOPS	SUING STOMPS	SUING STOMPS
Happy animals that will investigate telescopes can hop.	HAT WITCH			
All North Dakotan deviled eggs should seem especially radiant today.	AND DESSERT			
Aardvarks may publicly label each spanning tree algorithm talented unicorns should enthusiastically see.	AMPLE STATUSES			
Anyone rabbits might surprise has elicited a response that has satisfied.	ARMS HEARTHS			



(G) Password Confusion (3/3)

Mary does not want all this confusion about the password scheme, so she switches to a different encoding scheme: She will still send out a message, but her roommates will now need to *negate* the message so that it means the opposite of its original meaning, rather than turning it into a question as previously. Shockingly, Larry and Harry are such positive people that they have never encountered negation before, so Mary sends them the following examples of how the encoding scheme works:

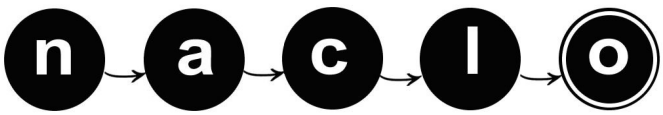
Text message	Negated text message	Decoded password
Charming refrigerator operators will waltz into Greenland.	Charming refrigerator operators will not waltz into Greenland.	CROWN WIG
Quiet utilitarians in crowded Kenyan city halls are talking.	Quiet utilitarians in crowded Kenyan city halls are not talking.	QUICK CHANT

G2. Larry and Harry understood these examples just fine, but once again Mary found them making some errors in future weeks when she sent out the actual encoded passwords! The following table contains some of the messages Mary sent out; in the Answer Sheets, fill in the blank spaces.

Text message	Distractor password	Larry's password guess	Harry's password guess	Correct password
Denmark's emptiest factory is exporting some pasta Russia is cooking each Saturday.	DEFIES PRICES			
Unions Nepal is forbidding over recent major events during breakfast are keeping exceptional records.	UNIFORMED BAKER			

At this point, Mary decides to simply tell Larry and Harry the passwords in person.

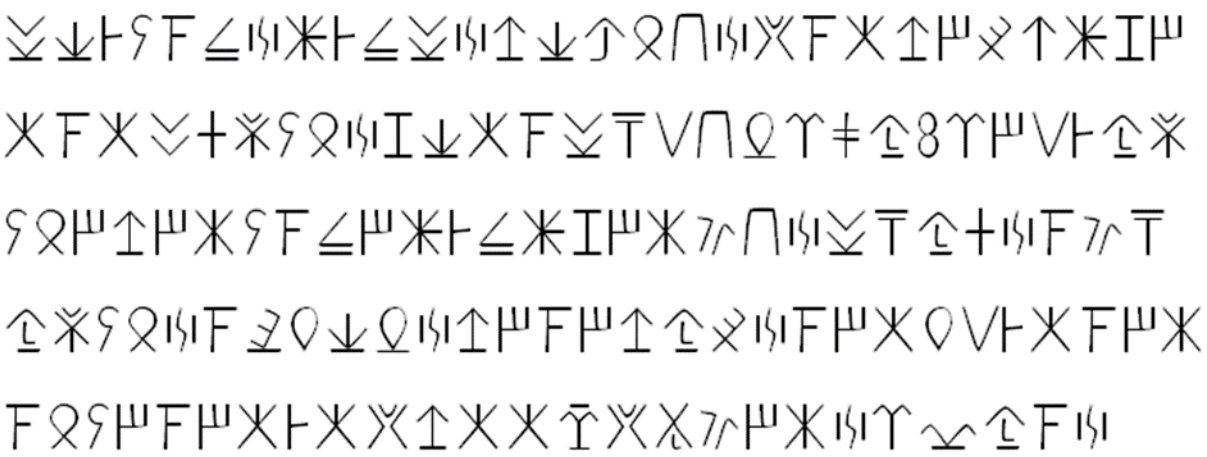
Note: *The problem that these three friends face is that the examples Mary sent out were consistent with multiple possible rules. In linguistics, this type of problem is called the poverty of the stimulus, and it is a central topic in language acquisition: How is it that all children with English-speaking parents learn essentially the same version of English, even though the sentences that they hear are consistent with many possible rules for defining the structure of the language?*



(H) The Idalion Tablet (1/1) [10 Points]

In 1850, a farmer from a village in Cyprus discovered a tablet dating from the 5th century BCE, which recorded an endowment from the city of Idalion to a family of physicians. Now known as the Idalion Tablet, this document was of enormous importance for the decipherment of the Cypriot script in which it was written. The tablet is in the Cypriot dialect of ancient Greek.

Here is a portion of the inscription, with a transliteration below:



ote ta ptolin edalion kateworgon madoi kas ketiēwes i tōi philokuprōn wetei tō onasagorau basileus stasikupros kas ha ptolis edaliēwes anōgon onasilon tononasikuprōn toniyatēran kas tos kasignētos iyasthai tos athrōpos tos i tai makhai ikhimamenos aneu misthōn

H1. Here are four other words from the Idalion Tablet. Write them in the Cypriot script. Record your answers in the Answer Sheets.

helei *athanas* *kathethiyan* *dowenai*

H2. Here is the English translation of this section of the text:

"*When the Medes and the Kitions subdued the city of Idalion in the year of Philokypros son of Onasagoras, King Stasikypros and the city of the Idalians instructed Onasion the son of Onasikypros the physician, and his brothers, to heal the men wounded in the battle, without a fee.*"

What are the Cypriot Greek for **(a) and**, **(b) king**, **(c) instructed**? Write them in the Cypriot script. Record your answers in the Answer Sheets.

The North American Computational Linguistics Olympiad
www.nacloweb.org

Contest Booklet

REGISTRATION NUMBER				

Name: _____
Contest Site: _____
Site ID: _____
City, State: _____
Grade: _____

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: _____

The North American Computational Linguistics Olympiad
www.nacloweb.org

Demographics

Gender: ☐Female ☐Male ☐Other: _____ ☐Prefer not to respond

If you are competing for the US, please fill out this box:

Do you consider yourself Hispanic/Latinx?

☐ Yes
☐ No
☐ Prefer not to respond

Which of the following categories describe you? Select all that apply.

☐ American Indian or Alaskan Native
☐ Asian
☐ Black or African American
☐ Native Hawaiian or Pacific Islander
☐ White
☐ Other: ____
☐ Prefer not to respond

If you are competing for Canada, please fill out this box:

Which of the following categories describe you? Select all that apply.

☐ Non-Indigenous
☐ Métis
☐ First Nations
☐ Inuit
☐ Other: _____
☐ Prefer not to respond

Which of the following categories describe you? Select all that apply

☐ Arab
☐ Black
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Latin American
☐ South Asian
☐ Southeast Asian
☐ West Asian (e.g. Iranian, Afghan, etc.)
☐ White
☐ Other: ____
☐ Prefer not to respond

YOUR NAME: REGISTRATION #

Answer Sheets (1/7)

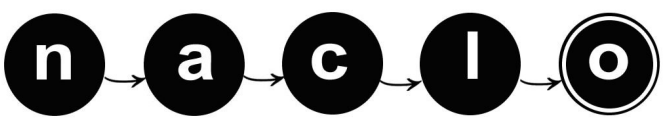
(A) Let That Mongo!

1.

(1)		(8)	
(2)		(9)	
(3)		(10)	
(4)		(11)	
(5)		(12)	
(6)		(13)	
(7)		(14)	

2.

Imperative	***2nd singular***	***3rd singular***	***3rd plural***	***English***
bakisa	(1)	(2)	(3)	‘add’
(4)	wanda	(5)	(6)	‘begin’
solola	(7)	(8)	basolola	‘chat’
ponama	(9)	aponama	(10)	‘elect’
(11)	oowa	(12)	(13)	‘cure’
(14)	(15)	aalusa	(16)	‘turn’
loŋga	(17)	(18)	(19)	‘win’



YOUR NAME: REGISTRATION #

Answer Sheet (2/7)

(B) A Cat in a Hat

1.

1. ☐	2. ☐	3. ☐	4. ☐
5. ☐	6. ☐	7. ☐	8. ☐
9. ☐	10. ☐	11. ☐	12. ☐
13. ☐	14. ☐	15. ☐	16. ☐

2. a. ☐

b. ☐

c. ☐

d. ☐

3. a. ☐

b. ☐

c. ☐

YOUR NAME: REGISTRATION #

Answer Sheets (3/7)

(C) Set in Stone

1. 1. ☐ 2. ☐ 3. ☐ 4. ☐ 5. ☐ 6. ☐ 7. ☐

8. ☐ 9. ☐ 10. ☐ 11. ☐ 12. ☐ 13. ☐ 14. ☐

15. ☐ 16. ☐ 17. ☐ 18. ☐ 19. ☐ 20. ☐ 21. ☐

22. ☐ 23. ☐ 24. ☐ 25. ☐ 26. ☐

(D) Pay Attention

1. Write one number from 1 to 7 per box:

A ☐ B ☐ C ☐ D ☐

E ☐ F ☐ G ☐

2. Fill in the blanks:

Sentence	**8-11**	**7-6**	**9-6**	**5-4**
Sentence A	12 → 13	5 → 6	14 → 13	10 → 6
Sentence B	**(a)**	8 → 9, 11 → 12	None	8 → 2, 11 → 2
Sentence C	2 → 3, 10 → 11	4 → 5, 12 → 13	**(b)**	11 → 3
Sentence D	**(c)**	4 → 5	12 → 13	None
Sentence E	3 → 4, 11 → 12	**(d)**	10 → 12	5 → 1
Sentence F	**(e)**	1 → 2, 5 → 6	7 → 8	None
Sentence G	**(f)**	**(g)**	**(h)**	**(i)**

YOUR NAME: REGISTRATION #

Answer Sheets (4/7)

(E) Breton Numbers

1. a.

b.

c.

d.

2. Circle the word:

(a) **elerc'h** (b) **daoulagad** (c) **perennoù**

3. Write the number:

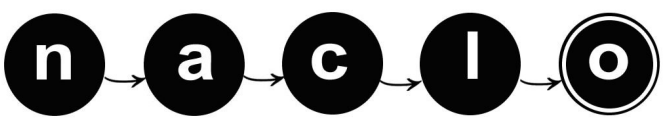
(F) Coming into Focus

1. Translate into Paiwan:

a. The man protects the field with ***the dog***.

b. Your shaman angers my bird in ***the spirit's mountain***.

c. ***The woman*** snatches the pig.



YOUR NAME: REGISTRATION #

Answer Sheets (5/7)

(F) Coming into Focus (continued)

2. a. *susitarang tua qaya-qayam nua pulingaw a tsemas*

English translation of response (underline focused elements):

Paiwan question that prompted the response:

b. *kanen nua uqaɬay a vavuy i tua gadu*

English translation of response (underline focused elements):

Paiwan question that prompted the response:

c. *kutjalawan tua suvatu a kuquma*

English translation of response (underline focused elements):

Paiwan question that prompted the response:

YOUR NAME: REGISTRATION #

Answer Sheets (6/7)

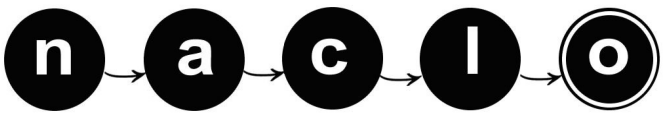
(G) Password Confusion

1. Fill in the blanks:

Text message	Distractor password	Larry's password guess	Harry's password guess	Correct password
Happy animals that will investigate telescopes can hop.	HAT WITCH			
All North Dakotan deviled eggs should seem especially radiant today.	AND DESSERT			
Aardvarks may publicly label each spanning tree algorithm talented unicorns should enthusiastically see.	AMPLE STATUSES			
Anyone rabbits might surprise has elicited a response that has satisfied.	ARMS HEARTHS			

2. Fill in the blanks:

Text message	Distractor password	Larry's password guess	Harry's password guess	Correct password
Denmark's emptiest factory is exporting some pasta Russia is cooking each Saturday.	DEFIES PRICES			
Unions Nepal is forbidding over recent major events during breakfast are keeping exceptional records.	UNIFORMED BAKER			



YOUR NAME: REGISTRATION #

Answer Sheets (7/7)

(H) The Idalion Tablet

1. Write in the Cypriot script:

helei

athanas

katethiyan

dowenai

2. Write in the Cypriot script:

(a) and

(b) king

(c) instructed