(I) Non Sequitur (1/4) [10 points]

At this very moment, scientists at NACLO Labs are hard at work on *Def21*, a definition-generating machine for English words. If they can just fix a few small problems, *Def21* is sure to be a huge popular success.

Def21 has access to a word list, L1, that includes some common English words and abbreviations, along with their definitions:

<u>L1</u>

al: American League cent: a penny grate: a grid of metal bars in: contained by ion: a charged particle rate: give a score real: truly existing

rest: stay still st: a street stat: a quantitative fact sting: sharply injure sure: certain union: a worker's organization

The NACLO Labs team also compiled a second list, L2, for *Def21* to use, by scanning many English words with their compressor machine.¹ The compressor identified patterns of letters that showed up frequently in the words it scanned, and it turned out that many of these word pieces had meanings of their own! For instance, when the compressor scanned the words "redo" (meaning "do again"), "regrow" (meaning "grow again"), and "replay" (meaning "play again"), it picked out "re" as a frequent repetition. Even better, with a little clever programming, the scientists were able to find out the meaning of "re" automatically (that is, based only on the meanings of the scanned words).

Here are the items in L2, and for each one, a sample of three words that it was found in:

<u>L2:</u>

ation (adaptation, consideration, installation) ing (jumping, knowing, wandering) ize (equalize, publicize, randomize) pre (predawn, prehistoric, premodern) re (redo, regrow, replay) un (uncertainty, uncommon, untie)

Finally, *Def21* has a sophisticated way of combining definitions from L1 with the meanings of its L2 items to guess the definition of a word it does not recognize. So, when *Def21* is given the word "unsure", which it does not recognize, *Def21* should recognize the pieces "un" from L2 and "sure" from L1 and produce the definition "the opposite of certain". In this case, *Def21* is confident about its result; when *Def21* is unsure about its guess, it will flag the result with one or more question marks.

¹ If you want to know more about how NACLO Labs' compressor machine works, try problem (H), Sequitur, also from NACLO 2021, Round 1. The NACLO Labs machine operates on the same basic principles as the algorithms shown there. Be aware, however, that solving problem (H) will not give you any advantage in solving this problem.



(I) Non Sequitur (2/4)

What *Def21* still struggles with is dividing unrecognized words into pieces. The great minds at NACLO Labs are trying out four different strategies -1, 2, 3, and 4 - for this task. Here are some words that *Def21* did not recognize, and the results of each strategy:

ingrate

- 1: contained by give a score (?)
- 2: contained by a grid of metal bars
- 3: continually give a score
- 4: contained by a grid of metal bars

resting

- 1: contained by a street again (?)
- 2: sharply injure again
- 3: continually stay still
- 4: sharply injure again

unionize

- 1: make into the opposite of a charged particle
- 2: make into a workers' organization
- 3: make into a workers' organization
- 4: make into a workers' organization

predation

- 1: a charged particle again (????)
- 2: not yet a charged particle (???)
- 3: not yet the result of (?)
- 4: not yet the result of (?)

realize

- 1: make into American League again
- 2: make into truly existing
- 3: make into truly existing
- 4: make into truly existing

station

- 1: a street a charged particle (??)
- 2: a quantitative fact a charged particle
- 3: the result of a street
- 4: the result of a street



(I) Non Sequitur (3/4)

Here are brief descriptions the designers wrote up o	of the four strategies,	with a few key words removed:
--	-------------------------	-------------------------------

1: Choose the (a)	_piece from (b) _	at each step	
2: At the first step, choose the	e (c)	piece from (d)	At each step after this, choose the
(e) piece from (f)	·		
<i>3:</i> At the first step, choose the	e (g)	_ piece from (h)	At each step after this, choose the
(i) piece from (j)	·		
4: Choose the (k)	_piece from (I) _	each step.	

I1. From the options *longest, shortest, L1, L2,* and *L1 or L2,* fill in the blanks (a) to (I).

12. Give the results of each strategy 1-4 on each of the inputs below. In a case where there is not enough information to choose between multiple possible results, you may enter any result that is consistent with the examples given:

(a) reunion	1:	
	2:	
	3:	
	4:	
(b) unrest	1:	
	2:	
	3:	
	4:	

(I) Non Sequitur (4/4)



I3. In a recent test, a Canadian scientist asked *Def21* for the definition of "centre" (which Americans typically spell "center"). This time, all four strategies *1-4* agreed on a result. What's more, *Def21* noticed that this result was the same as that given by all four strategies for a *different* unrecognized English word. What is the other word, and what definition was given to both words by all strategies?

Other word:	
Definition:	

© 2021, Daniel Lovsted, North American Computational Linguistics Open Competition, Round 1

