2010 Solutions

(A) Gelda's House of Gelbelgarg (1/1)

English systematically differentiates classes of nouns between whether they're Count – that is, are treated grammatically as if they can be counted, like *five cows* – or whether they're considered Mass, which can't themselves be counted. (This is a grammatical property of the *words*, not the items in question – even though rice comes in individual pieces you can't refer to five of them as "five rices" – you have to specify some measure word like "five *grains* of rice".)

Mass nouns tend to be liquids, undifferentiated masses, or masses of many, many tiny things (like rice), but as above it's a grammatical property: that's why even once you know a word is Count or Mass you can't be *sure* of the type of object it refers to. But you can still take a pretty good guess.

The properties of Count nouns are: they can co-occur with numerals, they can take "a"/"an" as an article, they co-occur with "fewer" but not "less" and "many" but not "much", and you can't leave a singular count noun "bare" – that is, *without* an article ("the","a"/"an"), quantifier (like "some", "every"), or numeral. Meanwhile, Mass nouns can occur "bare", can't occur with numerals or "a"/"n" without a "measure" or "container" word like "grain", "tablespoon", "plate", and co-occur with "less" but not "fewer" and "much" but not "many".

In addition, some words act as "measures" or "containers" – they can take an "of <something>" phrase and, whether or not it's Mass or Count, turn it into Count. Words like these are necessary to use Mass nouns with numerals, "a"/"an", etc.

How could you determine these properties in this problem if you didn't already know all this? Easy – put in words you do know in place of the unknown ones. For example, if a word like "water", "rice", "porridge", etc. fits in the same places that "meembel" does and makes good English sentences, but not in the places "gelbelgarg" does, then it's very likely that "meembel" is something like water, rice, or porridge. Meanwhile, "burger(s)" fits in the same places "gelbelgarg" does, but not "meembel", making it very likely that a "gelbelgarg" is some kind of discrete item.

2010 Solutions

(B) Say it in Abma (1/2)

In order to work out which word in the above sentences encodes the meaning associated with the individual words in the English translation we need to compare sentences with common meanings. For example 'water' appears in the translation of a., b., i., and k. Since 'runs' appears in e., j., and k. we can conclude that *sileng* = 'water' and *mworob* = 'runs'. This then allows us to conclude that *mwamni* (a. & b.) = 'drink', *mwabma* (e. & i) = 'here', *mwelebte* (i) = 'carry (as one carries water)', *mwesak* (h. & j.) = 'up'. We can deduce from these examples that the word order is virtually the same as in English.

A comparison of c. and d. shows us that *nutsu* = 'child'. A comparison of f. and g. shows us that *mwisib* = 'down' and also 'go down' or 'move downwards'. Since we know that f. is literally 'pulls + 'Mabontare' + 'down', we can deduce that *tela* (h.) = 'axe' and *mweselkani* (h.) = 'carry (as one does an axe)'. An analysis of i. shows us that the meaning encoded by English 'bring' is expressed by two words in Abma: *mwelebte* 'carry as one does water' (since it is not the same verb as in h. which involves carrying an axe) followed by the word *mwabma* translated as 'here' in e. We can analyse the meaning of English 'bring' as being made up of the meaning of 'carry' plus the idea of 'moving towards speaker's location', i.e., 'here'.

A comparison of b. and d. shows that *mwatbo* = 'keep doing something', and that (as in English) it immediately precedes the other verb in the same sentence.

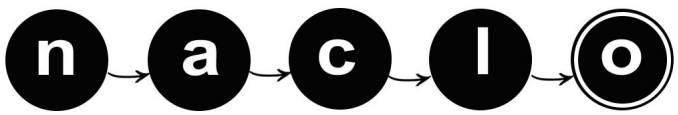
The words/verbs which express directional meanings, 'up', 'down', 'here' follow the other verb (d.) or the verb+Object noun (b.)

Having worked out the meaning of each word and the order in which words must combine in sentences, we can accurately translate the English sentences in 1 and the Abma sentences in 2.

While analysing the data it is very useful to create a dictionary as one goes along. Our analysis of sentences a. to k. above gives us the following results:

mwamni	drink	*mwisib*	(go) down
mwatbo	keep (doing)	*mwesak*	(go) up
mwerava	pull	*mweselkani*	carry (of axe)
mworob	run	*mwelebte*	carry (of water)
mwegau	grow	*sileng*	water
mwegalgal	crawl	*nutsu*	child
mwabma	(go) here, approach	*tela*	axe

We can now refer to this word list plus the words given in the table below when constructing our answers to Questions 1 and 2.



2010 Solutions

(B) Say it in Abma (2/2)

	ENGLISH	ABMA
1.	The teacher carries the water down.	***Sesesrakan mweselkani sileng mwisib.***
2.	The child keeps eating.	***Nutsu mwatbo mwegani.***
3.	Mabontare eats taro.	***Mabontare mwegani bwet.***
4.	The child crawls here.	***Nutsu mwegalgal mwabma.***
5.	The teacher walks uphill.	***Sesesrakan muhural mwesak.***
6.	The palm-tree keeps growing downwards.	***Butsukul mwatbo mwegau mwisib.***
7.	He goes up.	***Mwesak.***

	ABMA	ENGLISH
1.	*Sesesrakan mweselkani bwet mwabma.*	**The teacher carries the taro here/in this direction.**
2.	*Sileng mworob mwisib.*	**The water runs down.**
3.	*Mwelebte bwet mwesak.*	**He brings the taro up.**

(C) Lost in Yerevan (1/1)

This is a fairly simple problem to solve. One possible way to start is to look for the station on the map that has three words in its name. Then, based on its spelling as well as the spelling of the other given station names, one can easily determine that Armenian is written left to right and that both vowel and consonant sounds are represented as single characters in Armenian. Note that some sounds (e.g., "ts", map to multiple letters in English). After a few iterations, it is easy to reconstruct the entire part of the Armenian alphabet that is needed for this problem. The only "tricky" part has to do with the "T" character which doesn't appear in the names of the labeled stations, though it appears in the name of the subway system (METROPOLITEN).

(D) Huevos y Pimientos (1/1)

D1.

a. This sentence can be translated in two different ways:

pepinos y pimientos rojos
pimientos rojos y pepinos

b. This one can only be translated this way:

pomelos y pimientos rojos

D2.

a. TRUE- the eggs are unambiguously green
b. FALSE- the ham is unambiguously green

D3. jamon y huevos verdes

The "trick" is to reverse the word order to preserve the ambiguity.

2010 Solutions

(E) Texting, Texting, One Two Three (1/2)

From examining repeated elements and letters, we can work out most, but not all, of the character codes for the letters, along with SPACE being 1, the SHIFT sequence that creates a capital letter being 33, and the END MESSAGE sequence being 331 (SHIFT + SPACE, a sequence that otherwise wouldn't be used).

Lowercase 'z' doesn't appear in the plaintext, but knowing that uppercase 'Z' is 3323444 and "shift" is 33 we can conclude that lowercase 'z' is 23444.

The system we find is a "variable-length", rather than "fixed-length", code system. Although some of the codes are much longer than three digits, overall most codes are much shorter, because very common characters (like e, t, "space", etc.) are given very short codes whereas only fairly rare letters are given the longer codes.

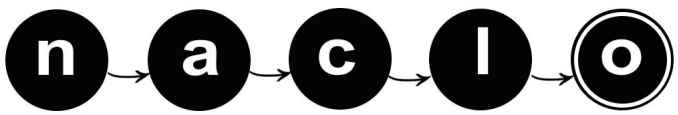
a	31	**n**	42
b	2341	**o**	32
c	242	**p**	342
d	233	**q**	23442
e	21	**r**	44
f	244	**s**	43
g	341	**t**	22
h	231	**u**	241
i	41	**v**	2342
j	23443	**w**	344
k	2343	**x**	23441
l	232	**y**	343
m	243	**z**	23444

Two letters remain, however, 'r' and 'x', neither of which appear in the plaintext. To determine their values, we have to work out the overall logic of the system.

Looking at the numerical codes, we notice that they aren't random: there are frequently repeated initial subcodes, and a lot of gaps. For example, many codes begin in 23-, 234-, and 34-, but none begin in, for example, 1-.

But why shouldn't a code begin with 1? If you consider the use of such a device, what would happen if a letter code began with one? What would happen is that, since 1 is "space", the device wouldn't know whether that 1 was intended as a space or as the first number of a longer code.

Looking further, we can see that *none* of the codes begins with another letter's code. That is, since 'a' is 31, no other letters' codes have 31- as their first two numbers, since 'b' is 2341, no other codes have these as their first four numbers, etc.

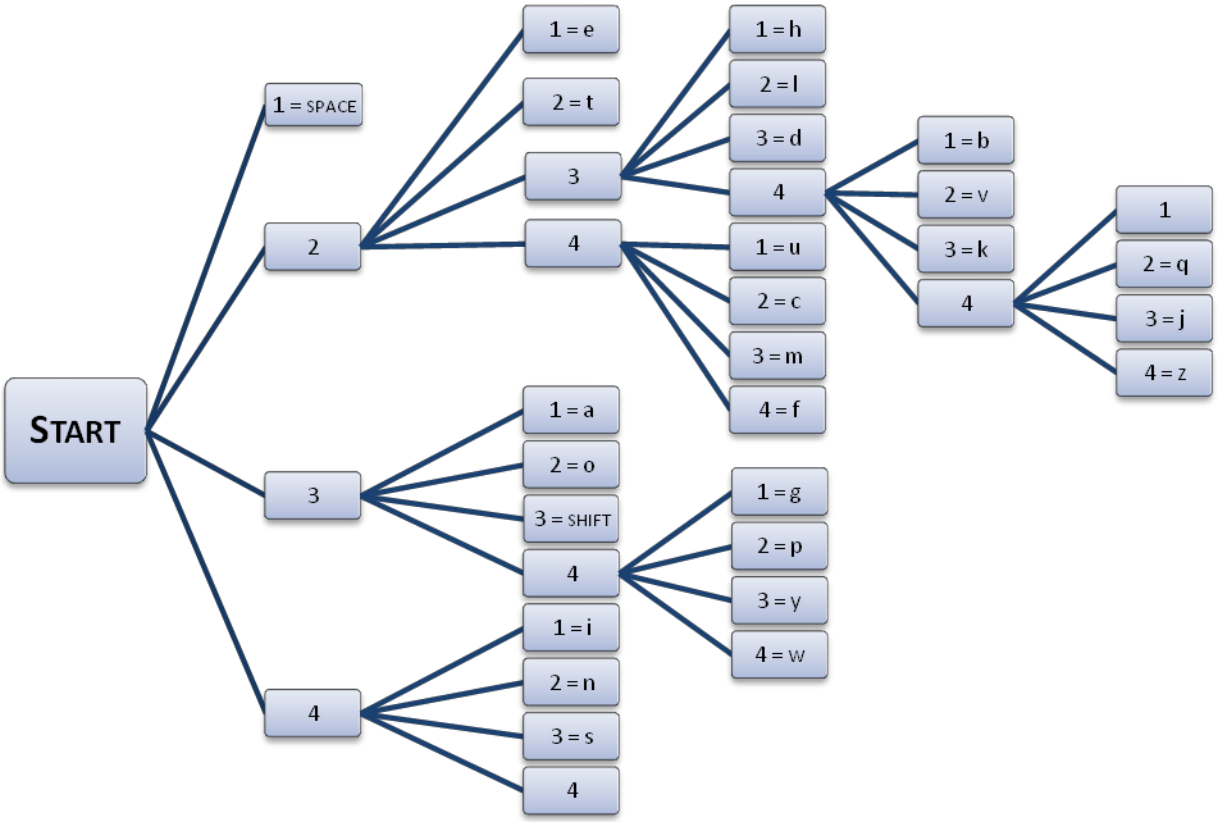


2010 Solutions

(E) Texting, Texting, One Two Three (2/2)

"Fixed-length" code systems, like the original three-number code system, always know when the user has keyed in a complete code. But since this system has "variable-length" codes, it needs some system to tell it whether some sequence, of whatever length, is a complete code or just the first part of a longer one. In this case, it knows when a code is complete because no beginning part of a valid code is a valid code.

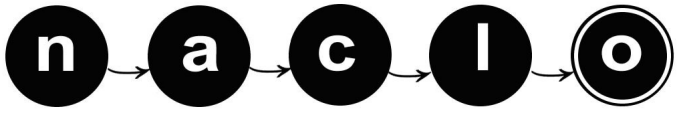
It's especially clear if we draw a "tree" of the codes: only those nodes that don't have further "branches" are assigned characters. Assigning "31" to "a" is fine, because there aren't any "311", "312", etc. to confuse the system. On the other hand, we can't assign "34" to anything because then it would prevent "341", "342", etc. from being entered.



Looking carefully at our tree, there are exactly two "free" nodes – that is, ones that don't already have a character assigned and that don't have any "branches": "44" and "23441". These are where "r" and "x" have to go – if they go anywhere else, the internal logic of the system is compromised.

Since frequent letters (like "e","t","a","o","i","n","s") get short codes, and rare letters (like "q","j","z") get long codes, "r" must be "44" while "x" is "23441".

Now we have all 26 letters, SPACE, SHIFT, and the END sequence, and can encode and decode any message for this device.



2010 Solutions

(F) Türkış Delıt (1/1)

The two suffixes in the problem have the following meaning:

consonant (ç or c) + vowel is "-er" in English and means "maker of something"

s + vowel + z is "-less" in English and means "without"

Whether the first consonant of the first suffix ("-er") is **ç or c** depends on the previous sound:

if the last sound of the stem is a voiced consonant or a vowel, the first consonant of the suffix is **c** (also voiced)

if the last sound of the stem is voiceless consonant, the first consonant of the suffix is **ç** (also voiceless)

The vowel depends on the last vowel in the stem:

if last vowel in the stem is a front, unrounded vowel (**e, i**), the suffix vowel is **-i**

if last vowel in the stem is a front, rounded vowel (**ö, ü**), the suffix vowel is **-ü**

if last vowel in the stem is a back, unrounded vowel (**a, ı**), the suffix vowel is **-ı**

if last vowel in the stem is a back, rounded vowel (**o, u**), the suffix vowel is **-u**

To summarize – the vowel in the suffix is the narrow vowel of the same type as the vowel preceding it. This is called *vowel harmony*

F1.

ikbalsiz, takatsiz – the vowels in the suffixes of these two words do not conform to the rules of vocal harmony and we can assume they are not of Turkish origin.

F2.

milkman - sütçü

speechless – sözsüz

F3.

linguist – dilci

mute - dilsiz

molder – kalıpçı

loose - kalıpsız

(G) Tangkhul Tangle (1/2)

Note that all but one of the Tangkhul sentences (sentence g) consistof two words. The two words consist of recurring components. For the first word, these are i, na, ā, ni, and thum. For the second word, these are masik, thāi, rā, ser, ngāi, ngarok, hāira, ei, lā, and ra. The word in the one word sentence (g) is drawn from components in the second set. In exactly one of the English sentences (sentence 5), the pronoun standing for the person doing the action of the verb is enclosed in parenthesis, showing that it is not present in the Tangkhul original. From this we may infer that:

1. sentences g and 5 match.
2. the final word in in each Tangkhul sentence is the verb.
3. The first words in each two-word Tangkhul sentence must be a pronoun.

The pronouns vary in person and number. First person includes the speaker (I, we), second person includes the one being addressed (you), and third person refers to some other entity (he, she, it, they).

	sg	dl	pl	tot
1st	0	1	1	2
2nd	1	0	1	2
3rd	1	2	1	4
tot	2	3	3	8

Exactly one of the Tangkhul pronouns occurs twice (āni, in b and h).This must be 3rd dual. Therefore, these sentences must match sentences 6 and 7. The component ā occurs four times, like third person; the component ni occurs three times, like dual. Working in this direction, it is possible to establish the following equivalences:

1st	i
2nd	na
3rd	ā

sg	(unmarked)
dl	ni
pl	thum

(G) Tangkhul Tangle (2/2)

This establishes the following matches between the Tangkhul and English: a=4, c=1, d=9, e=2, f=8, i=3. Tangkhul sentences a and c both contain masik; the English equivalents both contain "pinch/pinched". d, e, and f all contain thāi; the English equivalents all contain "see/saw." Tangkhul sentences i and g both contain "rā"; the English equivalents both contain "come". Thus, the first part of the final word in the Tangkhul sentences is the verb root ("pinch", "see", or "come"). It follows that b=6 and h=7. Now that the sentences are matched, it is possible to determine the meanings of the verb suffixes:

ser	all
ngarok	reciprocal (X one another)
ngāi	desiderative (want to X)
hāira	perfective (have Xed)
ei	past
ra	future
lā	interrogative

These are not problematic, except for ser. Transitive verbs are verbs which take object and intranstivie verbs are verbs which do not. If ser is suffixed to the transitive verb masik "pinch", we get to "pinch all". That is, it quantifies over the object. However, if it is suffixed to the intranstive verb rā "see", we get "call come". That is, it quantifies over the subject. Given this observation, and the above equivalences, it is possible to provide the correct translations for G2 and G3.