ORMC Advanced 2: Linguistics II Notes and Solutions

Sunny Liang

This document contains links to solutions as written by the original creators of the problems, as well as additional resources related to the problems. I have also included many of my own comments, notes, and observations.

1 Tenji Karaoke (NACLO 2009, Patrick Littel)

Original solution: `https://naclo.org/resources/problems/2009/N2009-AS.pdf`

2 What's the time in Tallinn? (NACLO 2014, Babette Newsome)

Original solution: `https://naclo.org/resources/problems/2014/N2014-AS.pdf`

(a)

(b) We have 1:00 and 2:00 from (a) and (b) respectively. So, "kell on" corresponds to "o'clock", "üks" is 1, and "kaks" is 2.

(c) This is probably the trickiest part that students may need the most guidance on. 2:15 is "veerand kaks" and we know that kaks is 2. Combined with (e), we can see that veerand must mean "quarter" or "fourth". This seems odd to us. In English we say "quarter past one" whereas in Estonian one says "quarter towards to two". Alas, convention is merely invention.

(d) Following the previous examples of the minutes coming first, "pool" must mean "half". Then, because of what we found in (c), "neli" must mean 4, **not** 3.

(e) If "veerand" is a quarter, and "kolmveerand" corresponds to 45 minutes, then "kolm" must be 3. "Üksteist" is 11, again, due to (c), but also because "üks" means 1. Interestingly, "teist" is related to "teine" which means "second" or "another". So "üksteist" is literally just "one, another one".

(f) "Minutit" sounds like "minute" because it is. Thus, "viis" must be 5. Since we know "üks" means 1, "läbi" must indicate that we're referring to the previous hour-mark not the next one, so "läbi" probably means "past".

3 CCG (NACLO 2014, J. Kummerfeld, A. Blackwell, P. Littell)

Original solution: `https://naclo.org/resources/problems/2014/N2014-OS.pdf`

4 Transcendental Algebra (IOL 2003, Ksenia Guiliarova)

Original solution: `https://ioling.org/booklets/iol-2003-indiv-sol.en.pdf`

Comprehensive resource: `https://tck.mn/transalg/`

It helps if you know about some elements of sentence construction: subject (the thing that the sentence is about), verb (what the subject is doing), and predicate (the thing being acted on).

To give you an idea of how to solve this problem, here's how you might analyze the sample sentences:

1. It's clear from that the two terms joined by addition mean father and son respectively. The numerator of the first term seems to have symbols for a man, woman, boy, and girl, yet the denominator is missing the man. Hence, applying our usual understanding of division, the woman, boy, and girl cancel out, leaving "father." We can also deduce that the adjacent multiplication of man*woman*boy*girl means "family" and boy*girl means something like "children" or "siblings."

   Addition clearly corresponds to "and".

   The exponentiated mouth-spray-like symbol that is applied to both father and son means "talking" or "speech". This also gives us a vague idea of what exponentiation might syntactically mean which will be reinforced by later examples. It seems that the base corresponds to the subject of a sentence and the exponent corresponds to the verb, that is, an event, action, or state of the subject.

2. From the previous problem, we saw that figures with dots over them are usually people (I mean, they even look like people), so the capital I with a dot over it is probably just a generic person, which is confirmed in later examples.

   The $>$ is a greater than sign modifying the person, hence, "giant".

   The hammer probably means "work" by process of elimination; also, it's a hammer.

   The t probably has something to do with time, which is, again, confirmed by later examples. Hence, the negative sign must indicate negation.

   The multiplication by n indicates plurality, analogous to the convention in math where n denotes an arbitrary natural number.

3. In the numerator we have children minus parents to indicate a family without parents, and then we divide by the negative parents to indicate that we are referring specifically to the children. This is the subject.

   The exponentiated pen must be the verb, that is, "writing".

   What are they writing? Oh, equals letter, they're writing a letter. So equals indicates the predicate.

4. The sg. here is just short for singular, which is important because we've seen that this language has plurals.

   Putting together everything we've learned so far, you should be able to see how the first term corresponds to "it wasn't us" because we have a person (the I with the dot), more than one of them (times n), and it was not these people (negative), who wrote (exponentiated pen) about you (equals sign, other person).

   So what does the negative t mean? Well, if t is for time, then subtracting time should indicate past tense. And indeed, the only examples written in past tense have the minus t, whereas the others are all present tense.

   So what do the subscripts mean? Well if "us" has subscript 1 and "you" as subscript 2, then perhaps the subscript indicates the perspective of narration, that is, first-person (I, us, we), second-person (you, y'all), third-person (they, them, he, she, etc.). Sentence 8 also demonstrates this.

5. The one new element here is the radical in the exponent. Interesting how this is the only sentence written in passive voice...

6. It's a heart ♡.

7. To be honest, I don't really know how the concentric semi-circle is related to food or eating, but there's only one verb in the whole sentence so there's nothing else it can be. I'm open to ideas though.

8. You got this ☺

5 Counting in the Roon (NACLO 2023, Riley Kong)

`https://naclo.org/resources/problems/2023/N2023-PS.pdf`

Ignoring the years for now, notice how "nuru" is 2 and "rimenuru" is 7, suggesting that "rim" is 5.

Notice that "arzus safur nuru" is 32, but we already know "safur" is 10 and "nuru" is 2. So, "arzus" must be 20, the base. Since "arzus" becomes "aresoyosier", we might deduce that "ares" is also 20 and "yosier" is 1 (which I've decided to give away in the hints).

Note, from the incomplete table, that "kior" is 3. Also notice that "arzus (20) safur (10)" became "ares (20) o yosier (1) safur (10)" which then became "ares (?) kior (3)". This, combined with my hint that Roon in 2012 is base-10, suggests that "ares" becomes 10 in 2012, even though "safur" already means 10. Thus, "safur" merely serves a vestigial role for the numbers from 10 to 19 and "ares" is used otherwise.

In the first line, "nuru (2) ŋokor (?)" is somehow the same as "yosier (1) + rim (5)". So "ŋokor" might be 3.

In the second line, we have "onem (6) × fak" = "ares nuru (20) beberin fiak" and we know that "fak" = "fiak". The only multiple of 6 in the twenties is 24. Therefore, "fak/fiak" is 4 and "beberin" merely indicates addition.

6 EAN-13 (IOL 2011, Hugh Dobbs)

Note that the original problem and solution uses the term "subcodes" a lot, whereas I chose to use "dialect" to refer to the ten possible starting digits. I think that saying "there are ten subcodes" is a bit confusing here since each digit only has three encodings which are all basically the same and "subcode" feels like it has a much more specific meaning than "dialect".

Original solution: `https://ioling.org/booklets/iol-2011-indiv-sol.en-us.pdf`

Useful resource: `http://www.gomaro.ch/Specifications/EAN13e.htm`

Parity structure: `https://en.wikipedia.org/wiki/International_Article_Number#Encoding_of_the_digits`

This problem is impossibly difficult to do without guidance or prior knowledge.

It helps to simplify notation, such as using 1's and 0's to represent black and white bars respectively.

The first key is to realize that the start, middle, and end consist respectively of 3, 5, and 3 alternating black and white bars (in binary: 101, 01010, 101); this is strongly-suggested by the template in (b).

If you count the remaining bars, there are 84 bars left which correspond neatly with the 12 digits we have to encode (7 bars per digit). This periodicity of 7 is also pretty much given away by the template in (b) as well as the first example with all zeroes.

From the all-zero example (and my hint), we can see that there are multiple ways of encoding the same digit: two encodings on the left and one on the right. If you write out all three encodings you might notice that they are all derivative of each other. To see what I mean, look at the first table in the *useful resource* linked above. Set B is just Set A but inverted and reversed; Set C is just Set A inverted, which is equivalent to Set B reversed.

Notice the following:

- Set A has odd parity, whereas Set B and C have even parity
- Set A and B start with 0 and end with 1, whereas Set C starts with 1 and ends with 0
- Set A and B are always on the left of the median, Set C is always on the right

Thus, we can get all three encodings from any one. And now we can take advantage of the second sample barcode and the limited information from part (a) to get the encodings for as many digits as we can.

Also, by comparing the leading digit with the positions of parities on the left-side, we can determine the the parity positions of each dialect. See second table in *useful resource* or a better table in *parity structure*.

This triple-encoding strategy has two advantages:

1. You can tell if the barcode is upside down if the first digit appears to be encoded using Set B instead of A (because a digit encoded using Set C when flipped upside-down will just look like it's from Set B). You can also just look at the parity of the first digit since all 10 dialects use even parity (Set A) for the first digit.

2. By comparing the parity of the left-side encodings to the leading digit, the scanner can tell if the barcode was read correctly. This is a nice back-up to the checksum feature.

Understanding all of this isn't that hard. Reverse-engineering it from a few examples is Herculean.

You don't need to know how the checksum works to solve this problem but it's pretty cool to learn anyways: `https://en.wikipedia.org/wiki/ISBN#ISBN-13_check_digit_calculation`