

Олимпиада по математическа лингвистика – 2012 г.

Национален кръг – София, 5–6 май

Индивидуално състезание

**Задача 1 (40 т.).** Дадени са аритметични примери в пределите на 200, написани с думи на езика сора:

$$\text{unji} + \text{gulji} = \text{galmuy}$$

$$\text{galmuy} + \text{yagi} = \text{mijgəl-bagu}$$

$$\text{tinji} + \text{tamji} = \text{mijgəl-mənləy}$$

$$\text{gəlji} + \text{bəkori gəlji} = ?$$

$$\text{bəkori mijgəl-yagi} + \text{mijgəl-tudru} = \text{yakori mijgəl-bəy}$$

$$\text{gulji} \times \text{bagu} = \text{mijgəl-bagu}$$

$$\text{bəkori mijgəl-mənləy} \times \text{yagi} = ?$$

$$\text{unji} \times \text{unji} = \text{mijgəl-unji}$$

$$\text{bəkori tamji} \times \text{bagu} = \text{unjikori mijgəl-unji}$$

$$\text{gəlji} \times \text{mijgəl} = ?$$

$$\text{mijgəl-bəy} \times \text{mijgəl-bəy} = \text{tamjikori tinji}$$

- (а) Напишете тези сметки с арабски цифри.
- (б) Напишете на езика сора числата, заменени с въпросителни.
- (в) Напишете на езика сора числата 228 и 138.

(Езикът сора е от езиковото семейство мунда. Говори се от над 200 хил. души в източна Индия. Буквите **ɟ**, **ʝ** и **ɽ** означават съгласни, а **a**, **ə**, **i** и **ɔ** – гласни звукове.)

**Задача 2 (40 т.).** Дадени са изречения на езика нупе и техните преводи на български език:

1. **bagi dede də guṣe o.** *Лошият мъж е в пещерата.*
2. **emitso le nyəgbəntucizi nə sa yigbecizi guṇ nə ye a.**  
*Стопанинът не вижда смелчаците, които изцеряват крадците.*
3. **məfwəci wa ede nə ta esa bo a nə za.**  
*Месарят търси парцала, който не е на стола.*
4. **emileguṇci la takada nə bə cigbəṇ o nə so a.**  
*Съседът не крие хартията, която е на дървото.*
5. **bocizi sa emileguṇci dede guṇ.** *Докторите изцеряват лошия съсед.*
6. **eṣitso la kpatizi nə fi guṣe bo a nə so.**  
*Гражданинът крие кутиите, които не са в пещерата.*
7. **zasə bologi bə ʃempa ziko bo.** *Прекрасната снимка е на черната стена.*
8. **litafizi fi kpati nə ta esako o nə bo.** *Книгите са в кутията, която е на масата.*
9. **bagi wa daguṇ nə də məkanta woṇciṅko bo nə za a.**  
*Мъжът не търси стълбата, която е в голямото училище.*
10. **takadazi kpe takuṇ o.** *Хартиите са на камъка.*
11. **nyəgbəntuci le takada nə bə egba o nə ye a.**  
*Смелчакът не вижда хартията, която е на оградата.*
12. **bagi bə zasəzi nə kpe ezuṇ wəṅjiṅ bo nə ya.**  
*Мъжът доставя снимките, които са на жълтия покрив.*
13. **nyəgbəntuci də ezi bo a.** *Смелчакът не е в града.*

(а) Преведете на български:

14. **zasəzi fi kpati bo a.**
15. **eṣitso la zasə nə ta esa wəṅjiṅ bo nə so a.**
16. **ede bə ʃempa o.**
17. **bagizi le emitso dede nə wa yigbeci za a nə ye.**

(б) Преведете на нупе:

18. *Съседът не търси черния парцал, който е на голямото дърво.*
19. *Докторът крие мъжа, който не доставя хартията.*
20. *Смелчакът изцерява месарите, които са в града.*
21. *Кутиите са на прекрасния стол, който не е в училището.*

(в) От каква информация бихте имали нужда, за да можете да преведете на нупе изречението:

22. *Парцалите не са на стената.*

(На езика нупе, за по-кратко наричан също нупе-нупе-тако, говорят около 100 хиляди души в Нигерия. Буквите **ə**, **c**, **j**, **ʃ**, **y**, **z** се четат съответно като *ъ*, *ч*, *дж*, *ш*, *й*, *ж*; **ɔ** е отворено *о*, а **ɲ** означава, че предходната гласна е носова. Съчетанията **kp** и **gb** означават слято произнесени *кп* и *гб*, които в този език се смятат за по един звук. Както в повечето африкански езици, в нупе има тонове, означавани обикновено с надбуквени знаци, но тук те са пропуснати освен в думата **bò**, за да се различава тя от **bo**.)

**Задача 3 (40 т.).** В днешно време повечето национални пощенски служби използват пощенски кодове. Пощенският код е последователност от букви или цифри, която се добавя към пощенския адрес с цел улесняване сортировката (включително и автоматическата) на кореспонденцията.

Пощенският код, който се използва в САЩ, се нарича ЗИП-код. Основният ЗИП-код е петцифрен и представлява код на местността. Разширеният ЗИП-код се състои от девет цифри, като последните четири са код на пощенска зона вътре в дадена местност и позволяват пратките да бъдат упътени по-точно и бързо до получателя. Понякога за уточняване мястото на доставката се добавя и „точка на доставяне“ (например когато няколко семейства живеят в една къща).

За подпомагане на автоматичната сортировка се използват няколко баркодни символики, като например POSTNET (използва се за пощенския код), PLANET (използва се от пощите за улесняване на сортировката и включва вид на кореспонденцията, код на абоната, код на кореспонденцията). Кодът се кодира като последователност от къси и дълги вертикални черти. Баркодът завършва с контролна сума.

(а) Като имате пред вид, че при кодирането на всяка от цифрите в баркодните символики POSTNET и PLANET се използват по две черти от единия вид, определете общия брой черти, необходим за кодирането на всяка цифра, така че да бъдат кодирани еднозначно.

(б) Ето няколко сведения за всеки от баркодовете А–D, дадени без определен ред. Посочете навсякъде буквата и вида на баркода и определете всички части на баркодовете:

1. кодът на пощенската зона е 3600, контролна сума 7;
2. кодът на абоната е 56934, писмото е стандартно (42);
3. писмото е първи клас (40);
4. кодът на пощенската зона е 2499.

A.



B.



C.



D.



(в) Нарисувайте баркодовете 97086-12340 и 50627-16292 в подходящата символика.

(г) Напишете пълния код на баркод Е. От кой вид е?

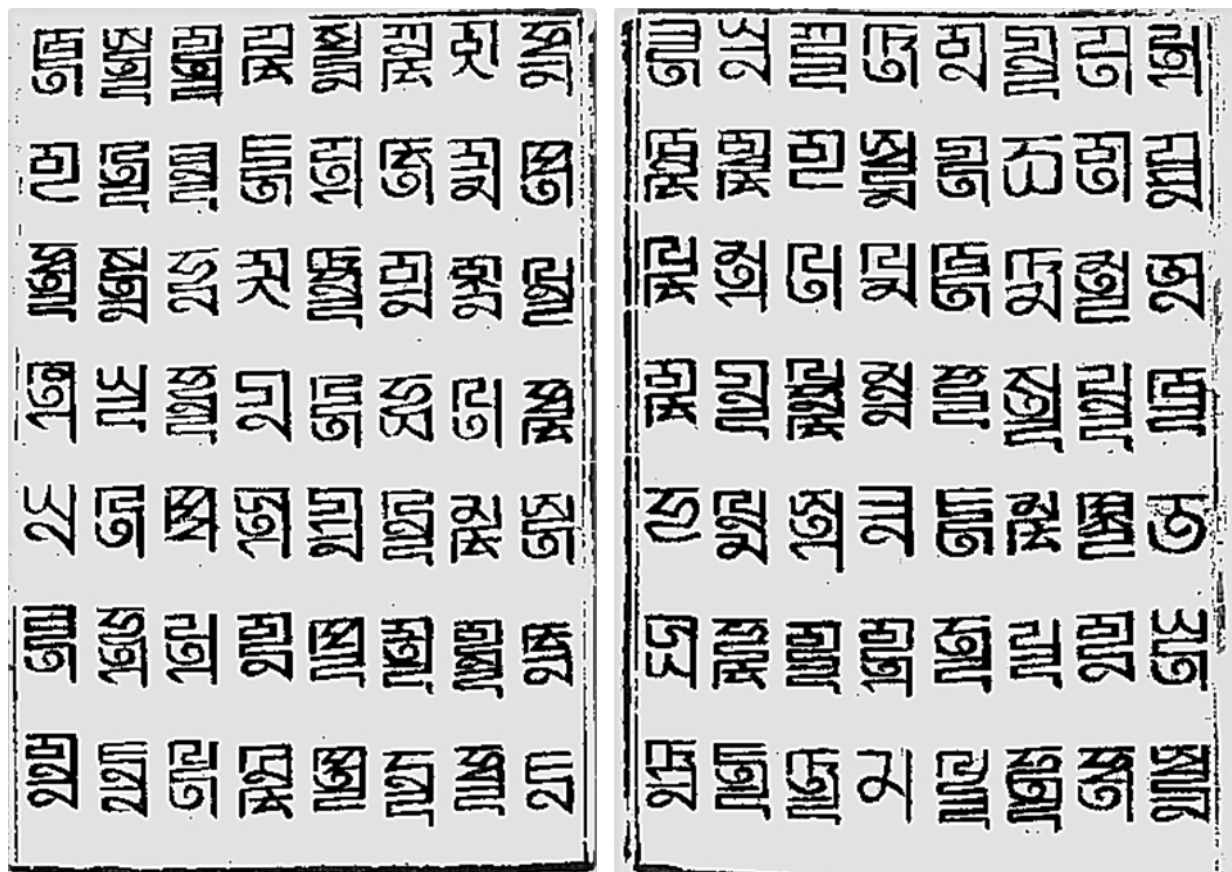
Е.



Олимпиада по математическа лингвистика – 2012 г.  
Национален кръг – София, 5–6 май. Отборно състезание

Когато монголският хан Кубилай основал в Китай династията Юан (1271–1368 г.), той наредил на тибетския лама<sup>1</sup> Дрогон Чогял Пагпа да създаде единна писменост, с която да може да се пише на всички основни езици в империята. Така се появило монголското квадратно писмо (писмото на Пагба-лама). То излязло от употреба скоро след падането на династия Юан, но от някои класически китайски текстове са запазени преписи с него.

„Байдзясинг“ („Стоте фамилии“) е поема от времето на династията Сунг (960–1279 г.), която изброява над 400 класически китайски фамилни имена. Първоначално е била написана с йероглифи, но през времето на династията Юан е била преписана и с писмото на Пагба-лама. Ето две последователни страници от текста (от един ръкопис от 1340 г.).



А ето и 20 реда (9–28) от поемата, написани с латинска транскрипция.<sup>2</sup> Двете страници от ръкописа отговарят на част от този откъс. Някои имена са пропуснати.

<sup>1</sup> Т.е. будистки свещеник.

<sup>2</sup> Транскрипцията отразява произношението от времето на династията Юан, а не съвременното. Буквата **g** показва, че предходният звук се образува с извит назад връх на езика, а **h** – че се произнася с придихание. Буквосъчетанието **ny** се чете като *нь*, а **ng** – като *н* в *гонг*. **x** е нещо като беззвучно *й* или много меко *ш*. ' е гърленият взрив (съгласният звук, който се чува в средата на отрицанието *ъ-ъ*). Знакът <sup>˘</sup> над гласна буква я превръща в полугласна.

	a	b	c	d	e	f	g	h
9	Fi	Lem	Drxim	Sĩa	Lue	Ho	Yi	Thang
10	Dxing	'In	Lo	Pi	Haũ	'U	'An	Srang
11	Yaũ	Yiu	Sri	Fu	Bue	Pen	Dzi	Khang
12	U	Yiu	Ngĩuan	Pu	Ku	Mung	Bing	Hõang
13	_____	Fu	Sring	Tai	Dam	Sung	_____	Bang
14	Xiung	Ki	_____	Khiu	_____	Triu	Tung	Liang
15	Du	_____	Lam	_____	Zi	_____	_____	Giang
16	Kĩa	Lu	Lxiũ	Ngue	Kĩang	Dung	_____	Kũaũ
17	_____	Sring	Lim	Xĩaũ	Trung	Ziu	Khiũ	Laũ
18	Kaũ	_____	Tshai	Den	Fan	Hu	_____	Faũ
19	Ngĩu	Wan	Tri	Ko	_____	Kõan	Lu	Maũ
20	Kĩing	_____	_____	Wu	Kan	Xiaĩ	'Ing	Tsung
21	Ting	Sĩuan	Pue	Dxing	'ũ	Sren	Hang	Hung
22	Paũ	Triu	_____	Sri	Tshue	Kĩi	Nriũ	Kiung
23	Dring	Xii	Xĩing	_____	Bue	Liu	Ngĩung	'Ung
24	Siun	Yang	_____	Xiue	Trin	Khiu	Kĩa	Fung
25	Nyue	Yi	Drĩu	Kin	Ki	Ping	Mue	Ziung
26	Tsing	Dõan	Fuũ	Wu	'U	Tsiaũ	Pa	Kiung
27	Wu	Ngue	Sran	Ku	Trhĩa	Hiũ	Fu	Bung
28	Dziũan	Trhi	Pan	Ngĩang	Tshiũ	Drĩung	Yi	Kiung

(а) Възстановете 18-те пропуснати фамилни имена (на позиции 13а, 13g, 14c, 14e, 15b, 15d, 15f, 15g, 16g, 17a, 18b, 18g, 19e, 20b, 20c, 22c, 23d, 24c).

(б) Вдясно виждате едно поле 3×3, преснето от една страница на друг (датиращ от 1418 г.) препис на „Байдзясинг“ с писмото на Пагба-лама. Шест от имената са пропуснати. Възстановете ги.



Олимпиада по математическа лингвистика – 2012 г.

Национален кръг – София, 5–6 май

Индивидуално състезание. Решения

**Задача 1.** Бройната система на сора е с основи 20 и 12. Числата от 1 до 11 са *bəy*, *bagu*, *yagi*, *unji*, *mənləy*, *tudru*, *gulji*, *tamji*, *tinji*, *gəłji*, *gəłmuy*.  $12+\gamma$  е *mijgəl-γ* (поне за  $1\leq\gamma\leq 7$ );  $\delta\times 20$  е *δkoři* с известни съкращения при малките коефициенти (*bəkoři* < *bəy·koři* 20, *bakoři* < *bagu·koři* 40, *yakoři* < *yagi·koři* 60).

(а, б)

$4 + 7 = 11$	$4 + 7 = 11$	$unji + gulji = gəłmuy$
$11 + 3 = 14$	$11 + 3 = (12+2)$	$gəłmuy + yagi = mijgəl-bagu$
$9 + 8 = 17$	$9 + 8 = (12+5)$	$tinji + tamji = mijgəl-mənləy$
$10 + 30 = 40$	$10 + (1\times 20+10) = 2\times 20$	$gəłji + bəkoři gəłji = bakoři$
$55 + 18 = 73$	$2\times 20+(12+3) + (12+6) = 3\times 20+(12+1)$	$bakoři mijgəl-yagi + mijgəl-tudru = yakoři mijgəl-bəy$
$7 \times 2 = 14$	$7 \times 2 = (12+2)$	$gulji \times bagu = mijgəl-bagu$
$37 \times 3 = 111$	$1\times 20+(12+5) \times 3 = 5\times 20+11$	$bəkoři mijgəl-mənləy \times yagi = mənləkoři gəłmuy$
$4 \times 4 = 16$	$4 \times 4 = (12+4)$	$unji \times unji = mijgəl-unji$
$48 \times 2 = 96$	$2\times 20+8 \times 2 = 4\times 20+(12+4)$	$bakoři tamji \times bagu = unjikoři mijgəl-unji$
$10 \times 12 = 120$	$10 \times 12 = 6\times 20$	$gəłji \times mijgəl = tudrukoři$
$13 \times 13 = 169$	$(12+1) \times (12+1) = 8\times 20+9$	$mijgəl-bəy \times mijgəl-bəy = tamjikoři tinji$

(в)  $228 = 11\times 20+8 gəłmukoři tamji$ ;  $138 = 6\times 20+(12+6) tudrukoři mijgəl-tudru$ .

**Задача 2.** Изречението на нупе започва с подлога, а глаголят се състои от две части, които обграждат допълнението (т.е. строежът на изречението е ПГ<sup>1</sup>ДГ<sup>2</sup>). Съществителното може да се следва от определение-прилагателно и(ли) от относително изречение, което се загражда с *nə* от двете страни. Отрицанието се изразява с *a* в края на изречението (значи в относителното изречение – преди второто *nə*). Множественото число има окончание *-zi*.

Ако глаголят изразява пространствено отношение, първата половина отразява числото на подлога, а втората е *bo*, което обаче се замества от *o* непосредствено след съществително, освен ако се намира пред отрицанието *a* (тогава пак е *bo*).

<i>e ...</i>	<i>sa ...</i>	
<b>də</b>	<b>fi</b>	<i>β</i>
<b>ta</b>	<b>kpe</b>	<i>na</i> (водоравна повърхност)
<b>bə</b>	?	<i>na</i> (отвесна повърхност)

(а, б)

- Снимките не са в кутията.
- Гражданинът не крие снимката, която е на жълтия стол.
- Парцалът е на стената.
- Мъжете виждат лошия стопанин, който не търси крадеца.
- emileguŋci wa ede ziko nə bə cigbəŋ woŋciŋko bo nə za a.**
- boci la bagi nə bə takada ya a nə so.**
- nyəgbəntuci sa məfwəci zi nə fi ezi o nə guŋ.**
- kpatizi kpe esa bologi nə də məkanta bo a nə bo.**

(в) Как се казва ‘*sa na* (отвесна повърхност)’ (въпросителната в табличката по-горе). {Отговорът е, че при този глагол числата не се различават, така че *Парцалите не са на стената* е **edezi bə šempa bo a**.}

**Задача 3.**

(а) Нека броят черти, необходим за шифроването на всяка цифра, е  $x$ . За всяка от десетте цифри кодировката трябва да бъде уникална. Знаем, че и двете баркодни символики използват две черти от единия вид. Тогава  $\binom{x}{2} = 10 \Rightarrow x=5$ .

Други закономерности:

- всеки баркод започва и приключва с по една дълга рамкова черта;
- всяка цифра в символиката POSTNET се кодира с 2 дълги и 3 къси черти, а в символиката PLANET се кодира с 3 дълги и 2 къси черти; кодировката на цифрите в символиката PLANET се получава от кодировката им в символиката POSTNET, като заменим късите черти с дълги и обратно;

	1	2	3	4	5	6	7	8	9	0
POSTNET										
PLANET										

- кодът POSTNET се състои от:
  - код на местността (град, район в града) – 5 цифри
  - код на пощенската зона вътре в местността – 4 цифри
  - точка на доставяне (не е задължителна!) – 2 цифри
  - контролна сума – 1 цифра
 ⇒ общо 10 или 12 цифри;
- кодът PLANET се състои от:
  - клас на кореспонденцията – 2 цифри
  - код на абоната – 5 цифри
  - код на кореспонденцията – 4 или 6 цифри
  - контролна сума – 1 цифра
 ⇒ общо 12 или 14 цифри;
- алгоритъм за изчисляване на контролната сума: събират се всички цифри без последната; делим сумата на 10 с остатък; контролната сума = 10 – остатък.

(б) Брой цифри в дадените баркодове: А – 14; В – 10; С – 12; D – 12.

1 → 92027-3600-31-(7) → POSTNET → С

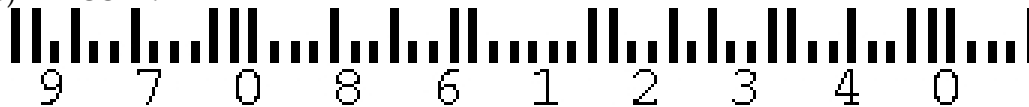
2 → 42-56934-1130-(2) → PLANET → D

3 → 40-16354-225636-(3) → PLANET → А

4 → 30313-2499-(6) → POSTNET → В

(в)

97086-1234-(0) → POSTNET



50627-1629-(2) → POSTNET



(г) 40-24567-247585-(1) → PLANET

- писмото е първи клас – 40
- кодът на абоната е 24567
- кодът на кореспонденцията е 247585
- контролна сума – 1

Олимпиада по математическа лингвистика – 2012 г.

Национален кръг – София, 5–6 май

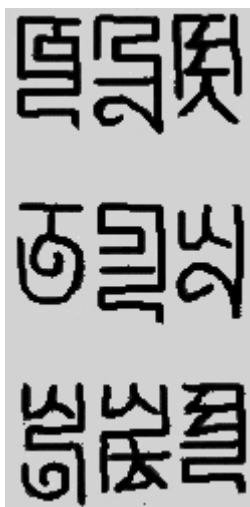
Отборно състезание. Решение

Пише се отгоре надолу в колонки, които вървят отдясно наляво. Всеки знак представлява една сричка (едно китайско фамилно име) и се състои от букви, които също се разполагат отгоре надолу. Всяка буква отговаря на един звук (който може да се означава с повече от една буква в латинската транскрипция). Гласната **a** не се пише с отделна буква, а се подразбира, ако в сричката няма друга гласна. За някои звукове обаче има по две букви в зависимост от мястото им в сричката (например двете **ǎ**-та в **Kǎǎi** се пишат с различни букви, по което впрочем се познава къде точно е **a**-то).

Откъсът от ръкописа е от 11e **Bue** до 25d **Kin**. Римувачите се срички (последните във всеки ред) се подреждат по диагоналите от четвъртия знак в най-дясната колонка на дясната страница надолу и наляво, после от първия в шестата колонка отдясно и т.н. – вижда се, че те образуват големи групи от знаци с еднакви долни краища.

(a) 13a. **Kǐi** – 13g. **Mǎi** – 14c. **Srǐu** – 14e. **Xǐang** – 15b. **Ngǐǎn** – 15d. **Min** – 15f. **Kǐue** – 15g. **Ma** – 16g. **Yan** – 17a. **Mue** – 18b. **Xǐa** – 18g. **Ling** – 19e. **Tsam** – 20b. **Fang** – 20c. **Giǔ** – 22c. **Tso** – 23d. **Hǎa** – 24c. **'U**.

(б)



11h. <b>Khang</b>	10g. <b>'An</b>	9f. <b>Ho</b>
12a. <b>U</b>	10h. <b>Srang</b>	9g. <b>Yi</b>
12b. <b>Yǐu</b>	11a. <b>Yaǔ</b>	9h. <b>Thang</b>