

Восьмая Международная олимпиада по лингвистике

Стокгольм (Швеция), 19–24 июля 2010 г.

Задачи индивидуального соревнования

Правила оформления решений задач

1. Не переписывайте условий. Решайте каждую задачу на отдельном листе (или листах), проставляя на каждом листе номер решаемой задачи, номер Вашего места и Вашу фамилию. Только в этом случае гарантируется точная оценка Вашей работы.
2. Полученные Вами ответы нужно обосновывать. Даже абсолютно верный ответ оценивается низко, если он приведён без всякого обоснования.

Задача №1 (20 баллов). Даны глаголы будухского языка в трёх формах:

форма 1: запретительное наклонение, I класс (мужской род)	форма 2: будущее время, I класс (мужской род)	форма 3: будущее время, II класс (женский род)	
<i>amaxax</i>	<i>axaxa</i>	<i>axaxa</i>	спать
<i>čötorhucu</i>	<i>čörhucura</i>		обменивать
<i>čimeoçi</i>		<i>čiroçira</i>	нести, вести
<i>hümoçonxu</i>	<i>hüçonxuna</i>	<i>hürçonxuna</i>	догонять
	<i>osura</i>	<i>orsura</i>	класть
<i>womoltu</i>	<i>woltula</i>		привязывать
?	<i>harkira</i>		натравливать
?	<i>jölküla</i>	<i>jölküla</i>	катить
?	<i>qalqala</i>		лежать
?	<i>quroçura</i>	<i>quroçura</i>	останавливать
?	<i>sonçona</i>	<i>sonçona</i>	вздрагивать
<i>amolqol</i>	?	<i>alqola</i>	садиться
<i>emensi</i>	?		гасить
<i>hömörçü</i>	?		толкать
<i>čumaraqar</i>		?	догонять
<i>hamoloçu</i>		?	глотать
<i>imankan</i>		?	оставаться
<i>jemeçi</i>		?	переходить

Заполните пустые клетки (затенённые заполнять не надо).

▲ Будухский язык относится к нахско-дагестанской языковой семье. На нём говорят около 5 000 человек в Азербайджане.

ö и ü ≈ ë и ю в словах *тётя* и *тюль* соответственно; i ≈ ы.

č, ç, ɔ, h, j, k, q, š, t, w, x — согласные звуки.

—Иван Держанский

Задача №2 (20 баллов). Даны числительные языка деху в алфавитном порядке и их значения в порядке возрастания:

*caatr nge caako, caatr nge caangömen, caatr nge caaqaihano,
ekaatr nge ekengömen, kõniatr nge kõniko, kõniatr nge kõnipi,
kõniatr nge kõniqaihano, lueatr nge lue, lueatr nge luako, lueatr nge luepi*

26, 31, 36, 42, 50, 52, 73, 75, 78, 89

(a) Установите правильные соответствия.

(b) Запишите цифрами:

*kõniatr nge eke + caatr nge luepi = ekaatr nge ekako
luengömen + luako = ekeqaihano*

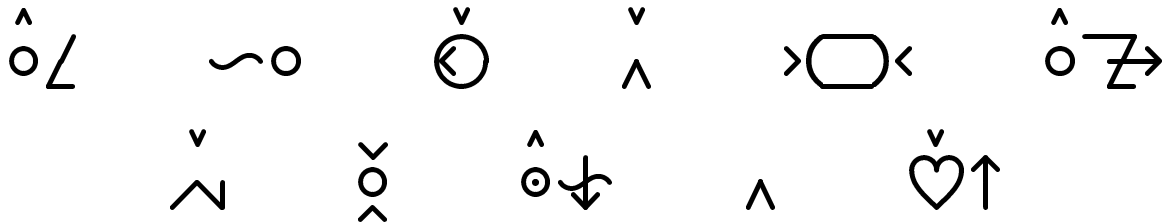
(c) Запишите на языке деху: 21, 48, 83.

⚠ Язык деху относится к австронезийской языковой семье. На нём говорят около 10 000 человек на острове Лифу, к востоку от Новой Каледонии. *c* ≈ *ч*; *ng* = *ng* в английском слове *hang*; *ö* ≈ *ё* в слове *тётя*; *q* — глухой звук, парный к английскому *w*; *tr* — согласный, схожий с русским *m*, но произносимый с загнутым назад кончиком языка.

—*Ксения Гилярова*

Задача №3 (20 баллов). Блиссимволика — универсальная система символов, разработанная австралийцем австрийского происхождения Чарльзом К. Блиссом (1897–1985), по мнению которого она должна быть понятна всем людям, независимо от их родного языка.

Даны слова, записанные блиссимволикой, и их переводы на русский язык в перепутанном порядке:



талия; деятельный; больной, нездоровый; губы; деятельность; дуть; западный; весёлый;
плакать; слюна; дышать.

(a) Установите правильные соответствия.

(b) Укажите, что означают следующие символы, если известно, что два из них имеют одно и то же значение :



(c) Запишите блиссимволикой:

воздух; туловище; подниматься; восток; печальный.

—*Александр Пиперски*

Задача №4 (20 баллов). Одним из самых значительных достижений генетики стала расшифровка генетического кода — создание мРНК–полипептидного словаря. Полипептиды (белки) являются строительным материалом для всех живых организмов. Молекулы полипептидов — это цепочки, которые состоят из аминокислот (обозначаемых как *Arg*, *Leu*, *Phe* и т. д.). Именно порядок аминокислот в полипептиде определяет его свойства. Когда клетки синтезируют полипептиды, они следуют инструкциям, записанным в молекулах матричной рибонуклеиновой кислоты (мРНК), цепочках, которые состоят из четырех нуклеотидов (обозначаемых как U, C, A, G).

Если клетка использует как матрицу следующую цепочку мРНК:

```
AUGUCGAGAAGUCACACCCACCCUCCGAAUUCUAGCCUCAAGAAUCUAGCUCGUGGCCGGAUCUAUACACGAU
GAAUGAGGUGGUGUCUUGUGUGCGAGUUAUUCUAAAUGAACCCGCUAGAUGGGUCAUGCGCCGGACGUAGGAUU
GUUUCAGGCACCCACUAUUCUGUACGUCCAAAUAGAUAAGUUGCCUCA,
```

то будут синтезированы следующие полипептиды:

- *Met-Ser-Arg-Ser-His-Thr-Pro-Pro-Ser-Glu-Ser-Ser-Leu-Lys-Asn-Leu-Ala-Arg-Gly-Arg-Ile-Tyr-Thr-Arg*
- *Met-Arg-Trp-Cys-Leu-Val-Cys-Glu-Leu-Phe*
- *Met-Asn-Arg*
- *Met-Gly-His-Ala-Pro-Asp-Val-Gly-Leu-Phe-Gln-Ala-Pro-Thr-Ile-Leu-Tyr-Val-Gln-Ile-Asp-Lys-Val-Ala-Ser*

(a) Клетка использует следующую цепочку мРНК:

```
AUGUUAACGUUCUAAAUGUGGGGGGACACCAG
```

Какой полипептид (какие полипептиды) она синтезирует?

(b) Клетка синтезировала следующий полипептид:

Met-Lys-Cys-Ile

Какую цепочку (какие цепочки) мРНК она могла использовать?

(c) Пары нуклеотидов иногда называют **корнями** и разделяют на две группы: сильные корни и слабые корни. Примерами сильных корней являются CU, GU, AC, GG. Примерами слабых корней являются AU, UA, UG, AA. Классифицируйте все остальные корни.

⚠ Представленные здесь данные несколько упрощены.

—Александр Бердичевский

Задача №5 (20 баллов). Даны слова на двух диалектах романшского языка и их переводы на русский язык. В таблице имеются пропуски:

сурсельвский	энгадинский	
<i>tut</i>	<i>tuot</i>	всё
<i>ura</i>	<i>ura</i>	время
?	<i>uolm</i>	вяз
<i>stumi</i>	?	желудок
<i>dunna</i>	<i>duonna</i>	женщина
<i>num</i>	<i>nom</i>	имя
<i>num</i> s	<i>nom</i> s	имена
?	<i>cuort</i>	короткий
<i>mund</i>	?	мир
<i>insumma</i>	<i>insomma</i>	наконец
<i>numer</i>	<i>nomer</i>	номер
<i>fuorcla</i>	?	перевал
?	<i>plomba</i>	пломба
?	<i>muossar</i>	показать
<i>buglia</i>	<i>buoglia</i>	пюре
<i>discuors</i>	<i>discuors</i>	разговор
<i>puolpa</i>	<i>puolpa</i>	сушёное мясо
<i>angul</i>	<i>angul</i>	угол
<i>fuorma</i>	<i>fuorma</i>	форма
<i>flur</i>	<i>flur</i>	цветок
<i>culant</i>	?	щедрый

- (a) Заполните пропуски.
- (b) Как на сурсельвском диалекте ‘труд’, *lavur* или *lavuor*? А на энгадинском?
- (c) На энгадинском ‘цветы’ – *fluors*, а ‘родители’ – *genituors*. Можно было бы думать, что и на сурсельвском эти слова выглядят так же, но на самом деле там говорят *flurs* и *geniturs*. Как это можно объяснить?
- (d) Переведите на оба диалекта: ‘вязы’, ‘углы’.

⚠ Романшский язык относится к ретороманской подгруппе романских языков. Это один из четырёх официальных языков Швейцарии наряду с немецким, французским и итальянским. На нём говорят около 35 000 человек в кантоне Граубюнден.

—Борис Иомдин

Редакторы: Александр Бердичевский, Божидар Божанов, Светлана Бурлак, Дмитрий Герасимов, Ксения Гилярова, Станислав Гуревич, Иван Держанский, Борис Иомдин, Алексей Назаров, Ренате Паюсалу, Александр Пиперски (отв. ред.), Мария Рубинштейн, Людмила Фёдорова, Адам Хестерберг, Тодор Червенков.

Русский текст: Александр Бердичевский, Ксения Гилярова, Иван Держанский, Борис Иомдин, Александр Пиперски.

Желаем успеха!

Восьмая Международная олимпиада по лингвистике

Стокгольм (Швеция), 19–24 июля 2010 г.

Решения задач индивидуального соревнования

Задача №1. Правила:

- форма 1: $-mV$ - после первого гласного, причём V зависит от гласного в следующем слоге (a перед a , o перед o или u , e перед i , $ö$ перед $ü$);
- форма 2:
 - $-a$, если основа оканчивается на $-aR$ или $-oR$,
 - $-Ra$, если основа оканчивается на $-i$, $-u$ или $-ü$,

где R — l или n , если в корне присутствует один из этих согласных, иначе r ;

- форма 3: форма 2 с $-r$ - после первого гласного, если только не следует непосредственно R .

Ответы:

форма 1	форма 2	форма 3
<i>hamerki</i>	<i>harkira</i>	
<i>jömölkü</i>	<i>jölküla</i>	<i>jölküla</i>
<i>qamalqal</i>	<i>qalqala</i>	
<i>qumoroŋu</i>	<i>quroŋura</i>	<i>quroŋura</i>
<i>somŋon</i>	<i>sonŋona</i>	<i>sonŋona</i>

форма 1	форма 2	форма 3
<i>amolqol</i>	<i>alqola</i>	<i>alqola</i>
<i>emensi</i>	<i>ensina</i>	
<i>hömörçü</i>	<i>hörçüra</i>	
<i>čumaraqar</i>		<i>čuraqara</i>
<i>hamoloŋu</i>		<i>haloŋula</i>
<i>imankan</i>		<i>inkana</i>
<i>jemeçi</i>		<i>jerçira</i>

Задача №2.

- 1–4: *caa* 1, *lue* 2, *köni* 3, *eke* 4;
 - 5, 10, 15: $\beta\text{-}pi = 5\beta$ ($1 \leq \beta \leq 3$);
 - 6–9, 11–14, 16–19: $\alpha\text{-}ngömen = 5 + \alpha$, $\alpha\text{-}ko = 10 + \alpha$, $-e\text{-}ko > -ako$
 $\alpha\text{-}qaihana = 15 + \alpha$ ($1 \leq \alpha \leq 4$);
 - 20, 40, 60, 80: $\gamma\text{-}atr = 20\gamma$ ($1 \leq \gamma$); $caa\text{-}atr > caatr$, $eke\text{-}atr > ekaatr$
 - 21–39, 41–59, ...: $\Gamma\text{-}nge\ \Delta = \Gamma + \Delta$ ($\Gamma = 20\gamma$, $1 \leq \Delta \leq 19$).
- (a) *caatr nge caako*: **31**, *caatr nge caangömen*: **26**, *caatr nge caaqaihana*: **36**, *ekaatr nge ekengömen*: **89**, *köniatr nge köniko*: **73**, *köniatr nge köniŋi*: **75**, *köniatr nge köniqaihana*: **78**, *lueatr nge lue*: **42**, *lueatr nge luako*: **52**, *lueatr nge luepi*: **50**.
- (b) *köniatr nge eke*: **64** + *caatr nge luepi*: **30** = *ekaatr nge ekako*: **94**
luengömen: **7** + *luako*: **12** = *ekeqaihana*: **19**
- (c) 21: *caatr nge caa*, 48: *lueatr nge köningömen*, 83: *ekaatr nge köni*.

Задача №3. ☐: существительное, ☐^v: прилагательное, ☐[^]: глагол (если в слове более одного символа, значок ставится над самым левым).

Стрелки (∧, ∨, <, >) используются для указания на отдельные части символов.

(a)

	часть речи	состав	значение
☐ [^] ∟	глагол	рот + нос	дышать
∩☐	существительное	вода + рот	слюна
☐ ^v	прилагательное	круг (солнце) + указатель	западный
∧	прилагательное	деятельность	деятельный
>☐<	существительное	туловище + 2 указателя	талиа
☐ [^] ∟→	глагол	рот + (воздух + наружу)	дуть
∧	прилагательное	больной, нездоровый	больной, нездоровый
∩∩	существительное	рот + 2 указателя	губы
☐ [^] ∩↓	глагол	глаз + (вода + вниз)	плакать
∧	существительное	деятельность	деятельность
∩↑	прилагательное	сердце + вверх	весёлый

(b)

	часть речи	состав	значение
∟	существительное	нос	нос
∩	существительное	вода	вода, жидкость
☐	существительное	туловище + указатель	шея
∧	глагол	деятельность	делать, действовать
>☐	существительное	глаз с бровью + указатель	бровь
☐ [^]	существительное	голова с шеей + указатель	шея

(c)

	часть речи	состав	значение
∟	существительное	воздух	воздух
☐	существительное	туловище	туловище
↑	глагол	вверх	подниматься
☐	существительное	круг (солнце) + указатель	восток
∩↓	прилагательное	сердце + вниз	печальный

Задача №4. Четыре полипептида в условии состоят из 24, 10, 3 и 25 аминокислот, а цепочка мРНК содержит $195 = ((24 + 10 + 3 + 25) + 3) \times 3$ нуклеотидов. Кажется вероятным, что три нуклеотида (триплет) обозначают одну аминокислоту или являются разделителем полипептидов (на самом деле символом останковки синтеза). Всего существует $4^3 = 64$ возможных триплета (все, кроме двух, представлены в условии) и лишь 20 разных аминокислот. Значит, некоторые триплеты имеют одинаковое значение.

	...U...	...C...	...A...	...G...
U...	UUU → <i>Phe</i> UUC → <i>Phe</i> UUA → <i>Leu</i> UUG → <i>Leu</i>	UCU → <i>Ser</i> UCC → <i>Ser</i> UCA → <i>Ser</i> UCG → <i>Ser</i>	UAU → <i>Tyr</i> UAC → <i>Tyr</i> UAA → STOP UAG → STOP	UGU → <i>Cys</i> UGC → <i>Cys</i> UGA → STOP UGG → <i>Trp</i>
C...	CUU → <i>Leu</i> CUC → <i>Leu</i> CUA → <i>Leu</i> CUG → <i>Leu</i>	CCU → <i>Pro</i> CCC → <i>Pro</i> CCA → <i>Pro</i> CCG → <i>Pro</i>	CAU → <i>His</i> CAC → <i>His</i> CAA → <i>Gln</i> CAG → <i>Gln</i>	CGU → <i>Arg</i> CGC → <i>Arg</i> CGA → <i>Arg</i> CGG → <i>Arg</i>
A...	AUU → <i>Ile</i> AUC → <i>Ile</i> AUA → <i>Ile</i> AUG → <i>Met</i>	ACU → <i>Thr</i> ACC → <i>Thr</i> ACA → <i>Thr</i> ACG → ?	AAU → <i>Asn</i> AAC → <i>Asn</i> AAA → <i>Lys</i> AAG → <i>Lys</i>	AGU → <i>Ser</i> AGC → <i>Ser</i> AGA → <i>Arg</i> AGG → <i>Arg</i>
G...	GUU → <i>Val</i> GUC → <i>Val</i> GUA → <i>Val</i> GUG → <i>Val</i>	GCU → <i>Ala</i> GCC → <i>Ala</i> GCA → <i>Ala</i> GCG → <i>Ala</i>	GAU → <i>Asp</i> GAC → <i>Asp</i> GAA → <i>Glu</i> GAG → <i>Glu</i>	GGU → <i>Gly</i> GGC → <i>Gly</i> GGA → <i>Gly</i> GGG → ?

Все цепочки мРНК начинаются с AUG → *Met*.

AUG UUA ACG UUC UAA AUG UGG GGG GGA CAC CAG
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 (a) *Met-Leu-?Thr-Phe* STOP *Met-Trp-?Gly-Gly-His-Gln*. Цепочка содержит два триплета, которые в условии не даны, так что мы не можем быть уверены в своем ответе. Мы, однако, получим ему подтверждение, когда решим задачу до конца.

(b) *Met-Lys-Cys-Ile* ← AUG $\left\{ \begin{matrix} \text{AAA} \\ \text{AAG} \end{matrix} \right\} \left\{ \begin{matrix} \text{UGU} \\ \text{UGC} \end{matrix} \right\} \left\{ \begin{matrix} \text{AUU} \\ \text{AUC} \\ \text{AUA} \end{matrix} \right\}$ ($1 \times 2 \times 2 \times 3 = 12$ возможностей).

(c) Корень XY называется сильным, если XYA, XYG, XYS и XYU кодируют одну и ту же аминокислоту (UC, CC, CG, GC). Корень называется слабым, если это не так (UU, CA, AG, GA).

Задача №5.

сурсельвский	энгадинский	
<i>uo</i>	<i>uo</i>	перед сочетанием <i>l</i> или <i>r</i> с другим согласным
<i>u</i>	<i>u</i>	перед <i>l</i> или <i>r</i> без другого согласного
<i>u</i>	<i>o</i>	перед <i>m</i>
<i>u</i>	<i>uo</i>	перед другим согласным

	сурсельвский	энгадинский	
(a)	<i>uolm</i>	<i>uolm</i>	вяз
	<i>stumi</i>	<i>stomi</i>	желудок
	<i>cuort</i>	<i>cuort</i>	короткий
	<i>mund</i>	<i>muond</i>	мир
	<i>fuorcla</i>	<i>fuorcla</i>	перевал
	<i>plumba</i>	<i>plomba</i>	пломба
	<i>mussar</i>	<i>muossar</i>	показать
	<i>culant</i>	<i>culant</i>	щедрый

(b) *lavur* на обоих диалектах.

(c) В сурсельвском (в отличие от энгадинского) первое правило не применяется к формам мн. числа. Это может означать, что оно не действует, если первый согласный относится к основе, а второй к окончанию, или что гласный определяется до добавления окончания, или что гласный в форме мн. числа уподобляется гласному из формы ед. числа.

(d) ‘вязы’: *uolms* (на обоих диалектах).

‘углы’: *anguls* (сурсельвский), *anguols* (энгадинский).

Восьмая Международная олимпиада по лингвистике

Стокгольм (Швеция), 19–24 июля 2010 г.

Задача командного соревнования

Перед вами некоторые слова и их толкования из толкового словаря монгольского языка (*Mongol qelnij товч тайлбар тол'*, Ulaanbaatar, 1966) в латинской транслитерации:

1. **asaq**: носоq, gal genel garaq
2. **bal**: zögijn cecgijn šüüseer bolovsruulaq čiqer amttaj ötgön züjl
3. **bor**: qar cagaan qojor qol'col'dson öngö
4. **büleen**: zöög, qaluun biš, qüjten biš
5. **cagaan**: jumny cas met öngö
6. **cas**: žiqüün cagt agaart usan talstuud bij bolž cav cajm ungaril širqgüüdeer buuq agaaryn tundas
7. **čiqer**: tuscgaj manžingas jalgaruulan avdag cagaan öngötej bögөөд amlag težeelijн talst bodis
8. **davs**:
 - (1) gašunduu qurc amttaj talst bodis, qoolond amt oruulaqад qereglene
 - (2) ustörөгч atom n' tömөрlөгijn atomaar soligdson qimijn bodis
9. **gal**: šataž bajгаа bodisoos garsan qaluun
10. **ideq**: am'tny jumnyг qool болгон qeregleq
11. **kal'ci**: qimijn үндсен маqbод, qөngөн cagaan tömөрlөг
12. **kilogramm**: qünd qөngөнij gemžүүр, neg mјangan grammtaj tencүү
13. **kofë**:
 - (1) kofëjn mod gedeg qaluun орны modny böөрөнqij үр
 - (2) ene үреер čanasan und
14. **manan**: usny uur дүүрсен tungalag bus агаар
15. **mös**: qөldsön us
16. **nocoq**: asaq, šataq
17. **nojton**: quurajn esreg utga, ustaj
18. **nүүrs**: mod šataqад bij болоq šataq qatuu züjl
19. **өtgön**: šingenij esreg utga
20. **šaraq**: ideenij zүjlijг gald түлž болгоq
21. **šataq**: gal nocoq
22. **šingen**: ötgön gedgijn esreg utga

23. **süü:** am'tny qöqnöös garaq cagaan šaranguj öngötej šingen züjl
24. **talst:** tals büqij qatuu bodis
25. **tülš:** gald tüleqed zoriulž beltgesen tülee, argal, nüürs zereg jum
26. **und:** uuq jum, undaan
27. **us:** ustörögč qüčiltörögč qojoryn qimijn cever nijlel boloq öngögüj, tungalag, šingen züjl
28. **ustaj:** us büqij
29. **utaa:** jum šataqad garaq nüürsnij narijn širqeg büqij qööröq züjl
30. **uur:** šingen züjljin qalaqad garaq nojton qij
31. **uuq:** šingen jumyg balgaž zalgiq
32. **qaluun:** bodisyn qödölgöönij tusgaj negen qelber bögööd bodisyn öčüüqen quv' molèkul, atomyn qödlöqöd bij boloq ilč
33. **qar:** cagaany esreg, qöö, nüürsnij öngö
34. **qatuu:** zöölön gedgijn esreg utga
35. **qij:** gazryn agaар mandlyg bij bolgogč agaар бije, agaар bodis
36. **qojor:** neg deer negijg nemsen too
37. **qöldmöl:**
 - (1) qöldsön jum
 - (2) qöldöösön amtlag idee
38. **qöngön:** qünd gedgijn esreg utga
39. **qöö:** jumand togtson utaa
40. **qool:** ideq težeelijн züjl
41. **qüjten:** qaluun gedegtej esergüiceldsen utga, jumny serüün žiqüünij n'
42. **qünd:** čanar qöngöngüj, žintej
43. **quuraj:** nojton gedgijn esreg utga
44. **žin:**
 - (1) qünd qöngönij qemžee; neg žin n' 16 lan bögööd 600 grammtaj tencene
 - (2) qünd qöngönij bagcaa

* * *

(a) Переведите на русский язык:

čiqertej kofé, mјangan žin, neg kilogramm, ötgön manan, qaluun us, qojor utga, quuraj süü, qüjten us, süü uuq, süün qöldmöl, süüteј kofé, undny us.

(b) Переведите как можно больше монгольских слов из текста задачи.

—Борис Иомдин

Русский текст: Борис Иомдин.

Желаем успеха!