第六届国际理论, 数理及应用语言学奥林匹克竞赛

保加利亚, 阳光海滩, 2008年8月4 — 9日

个人赛题目

解答规则

1. 毋需抄题. 将不同问题的解答分述于不同的答题纸上. 每张纸上注明题号、座位号和姓名. 否则答题纸可能被误放或遗失.
2. 解答需详细论证. 无解释之答案, 即便完全正确, 也会被处以低分.

题 #1 (20 分). 下面是以所谓的 Listuguj 拼写法拼写的米克玛克语的单词, 它们的语音转写及汉语翻译:

1	***tmi'gn***	dəmīgən	斧头
2	***an'stawteg***	anəstawtek	不安全
3	***gjiansale'wit***	əkciansalēwit	天使长
4	***mgumie'jo'tlatl***	əmkumiējōdəladəl	钉蹄铁 *(给马)*
5	***amqwanji'j***	amx^{w}ancīc	匙
6	***e'jnt***	ējənt	印第安代理人
7	***tplutaqan***	ətpəludaɣan	法律
8	***ge'gwising***	gēg^{w}isink	躺在顶部
9	***lnu'sgw***	lənūsk^{w}	印第安女人
10	***g'p'ta'q***	gəbədāx	上方
11	***epsaqtejg***	epsaxteck	炉

(a) 转写这些词:

12	***gsnqo'qon***	愚蠢 (名词)
13	***tg'poq***	泉水
14	***gmu'jmin***	树莓
15	***emtoqwatg***	崇拜
16	***te'plj***	山羊

(b) 用 Listuguj 拼写法拼写:

17	ətpədēsən	南方
18	əmteskəm	蛇
19	alaptək	环顾四周
20	gəlamen	所以, 因此

NB: 米克玛克语是一种阿尔冈昆语. 在加拿大, 约有8000人使用该语言.

In the 转写 ə ≈ *o* in *abbot*, [c] = *ch* in *church*, [j] = *j* in *judge*, [x] = *ch* in Scottish *loch*, ɣ is the same sound but voiced; [^{w}] shows that the preceding consonant is pronounced with rounded lips. 标记 ¯ 表示长元音.

—Bozhidar Bozhanov

题 #2 (20 分). The following are four excerpts from Old Norse poems composed around 900 C.E. All of them are written using the meter named *dróttkvætt* (lit. 'court meter'):

I

1 **ók at ísarnleiki**
2 **Jarðar sunr, en dunði ⋯**

II

1 **þekkiligr með þegnum**
2 **þrymseilar hval deila.**
3 **en af breiðu bjóði**
4 **bragðvíss at þat lagði**
5 **ósvífrandi ása**
6 **upp þjórhluti fjóra.**

III

1 **áðr gnapsólar Gripnis**
2 **gnýstœrandi fœri**
3 **rausnarsamr til rimmu**
4 **ríðviggs lagar skíðum.**

IV

1 **háði gramr, þars gnúðu,**
2 **geira hregg við seggi,**
3 **(rauð fnýsti ben blóði)**
4 **bryngǫgl í dyn Skǫglar,**
5 **þás á rausn fyr ræsi**
6 **(réð egglituðr) seggir ⋯**

One of the main principles of *dróttkvætt* is alliteration. The first line of each distich (pair of lines) contains two words beginning with the same sound, and the first word of the second line begins with this sound, too: e. g., **rausnarsamr**, **rimmu** and **ríðviggs** (III:3–4). All vowels are considered to alliterate with one another and with **j**: e. g., **ók**, **ísarnleiki** and **Jarðar** (I:1–2). But this is not the only rule.

The texts given above have been handed down in more than one manuscript. Sometimes different words are found in corresponding parts of the text, and the scholars have to decide which of the variants is original. Different considerations may motivate the conclusion. Sometimes the rules of versification help to recognize some of the variants as false. For example, in line I:2 we find not only **dunði**, but also **dulði** and **djarfi**. **dulði** can be rejected because of the structure of the verse, but both **dunði** and **djarfi** fit into the line, and one needs other reasons to choose between these words. In line III:1 **Gripnis** and **Grímnis** occur in the manuscripts, but **Grímnis** doesn't fulfill the requirements of the verse.

(a) Describe the rules which are observed in a distich of *dróttkvætt*.

***

(b) Given is a stanza in which 13 words are omitted:

V

1 [a] **(þreifsk reiddra øxa**
2 [b] **; knǫttu spjǫr** [c] **)**
3 [d] **bitu seggi**
4 [e] **þjóðkonungs ferðar,**
5 **þás (** [f] **hǫlða)**
6 [g] [h] [i]
7 **(hǫr vas** [j] **of** [k] **)**
8 [l] **(flugbeiddra** [m] **).**

The following list contains (in alphabetical order) all 13 omitted words and two words which do not belong in stanza V:

andskoti, Gauta, glymja, hlaut, hugfyldra, hœgra, ríks, rymr, sigr, smíði, svartskyggð, sverð, svírum, sǫngr, vigra

Fill in the gaps in stanza V.

NB: Old Norse is a North Germanic language which was in use approximately between 700 and 1100 C.E.

æ ≈ English *a* in *cat*, **œ** = French *eu* or German *ö* (these letters stand for long vowels). **ø** is read as a short **œ**; **y** = French *u* or German *ü*, **ǫ** is an open *o*. **au** and **ei** are pronounced as a single syllable. **ð** and **þ** = English *th* in *this* and *thin* respectively. **x** = **k**+**s**. 标记 ´ 表示长元音. All samples of poetry in the problem are given in a normalized orthography and conform to the rules of the genre. —*Alexander Piperski*

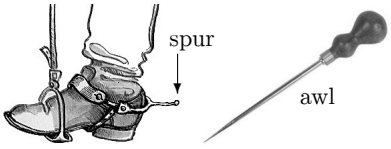
题 #3 (20 分). The following are words and compounds in two languages of New Caledonia -- Drehu 语 andCemuhî -- andtheir English translationsgiven out of order:

Drehu 语	中文
drai-hmitrötr, gaa-hmitrötr, i-drai, i-jun, i-wahnawa, jun, ngöne-gejë, ngöne-uma, nyine-thin, uma-hmitrötr	sanctuary, bunch of bananas, calendar, bone, church, coast, awl, Sunday, skeleton, wall

Cemuhî	中文
a-pulut, ba-bwén, ba-jié, bé-ôdu, bé-tii, bé-wöli, bé-wöli-wöta, tii, wöta	bed, animal, fork, cup, pencil, coast, to write, twilight, spur

And here are several words translated from Drehu 语 into Cemuhî:

Drehu 语	***gaa***	***ngöne-gejë***	***nyine***	***thin***
Cemuhî	***a***	***ba-jié***	***bé***	***wöli***



(a) 找出正确的对应关系.

(b) What do you think the words ***wahnawa*** and ***drai*** mean in Drehu 语, and ***wöli*** and ***pulut*** in Cemuhî?

(c) In Drehu 语 ***tusi*** is `book' and ***bii*** is `bee'. 翻译 Drehu 语: ***i-bii***, ***tusi-hmitrötr***.

NB: Drehu is spoken by over 10 000 people on Lifu Island to the east of New Caledonia. Cemuhî is spoken by approx. 2000 people on the east coast of New Caledonia. Both languages belong to the Austronesian family.

In Drehu ***ë*** ≈ *a* in *aspen*, ***ö*** = French *eu* or German *ö*, ***hm*** and ***hn*** are specific unvoiced consonants; ***dr*** and ***tr*** ≈ *d* and *t* in *word* and *art*, uttered with the tip of the tongue turned back; ***j*** and ***th*** = English *th* in *this* and *thin* respectively; ***ng*** = *ng* in *hang*; ***ny*** ≈ *ni* in *onion*.

A sanctuary is the principal, most sacred part of a church. —*Ksenia Gilyarova*

题 #4 (20 分). The following are words in Copainalá Zoqueandtheir English translations:

mis nakpatpit	with your cactus	**k2m2ŋdaPm**	shadows
nakpat	a cactus	**P2s ncapk2sm2šeh**	as ifabove my sky
mokpittih	onlywith the corn	**capšeh**	likea sky
pokskuky2sm2taPm	above the chairs	**pahsungotoya**	for the squash
pokskuy	a chair	**pahsunšehtaPmdih**	justlikesquashes
peroltih	onlya kettle	**t2ckotoyatih**	onlyfor the tooth
koc2ktaPm	mountains	**kumguky2sm2**	above the town
komg2sm2tih	rightabove the post	**kumgukyotoyataPm**	for the towns
P2s ŋgom	my post	**cakyotoya**	for the vine
k2m2ŋbitšeh	as ifwith the shadow	**mis ncay**	your vine

(a) 翻译成中文:

caky2sm2tih
k2m2ŋšeh
P2s mok
mis nd2ctaPm
pahsunbit
perolkotoyašehtaPm

(b) 翻译成 Copainalá Zoque:

for the chair
with my kettle
justlikea mountain
posts
above the shadows
your town

NB: The Copainalá Zoque language is of the Mixe-Zoque linguistic family. 在墨西哥南部的 Chiapas 省, 约有10 000人使用该语言.

2 ≈ *u* in *but*; **c** ≈ *ts* in *hats* (pronounced as a single consonant), **nc** ≈ *nds* in *hands*, **š** = *sh*, **ŋ** = *ng* in *hang*, **y** = *y* in *yay!*; **P** is a specific consonant (the so-called glottal 塞音).

—*Ivan Derzhanski*

题 #5 (20 分). The following are sentences in Inuktitutandtheir English translations:

1.	***Qingmivit takujaatit.***	Your dogsaw you.
2.	***Inuuhuktuup iluaqhaiji qukiqtanga.***	The boy shot the doctor.
3.	***Aanniqtutit.***	You hurt yourself.
4.	***Iluaqhaijiup aarqijaatit.***	The doctor cured you.
5.	***Qingmiq iputujait.***	You speared the dog.
6.	***Angatkuq iluaqhaijimik aarqisijuq.***	The shamancured a doctor.
7.	***Nanuq qaijuq.***	The polar bear came.
8.	***Iluaqhaijivit inuuhuktuit aarqijanga.***	Your doctor cured your boy.
9.	***Angunahuktiup amaruq iputujanga.***	The hunterspeared the wolf.
10.	***Qingmiup ilinniaqtitsijiit aanniqtanga.***	The dog hurt your teacher.
11.	***Ukiakhaqtutit.***	You fell.
12.	***Angunahukti nanurmik qukiqsijuq.***	The huntershot a polar bear.

(a) 翻译成中文:

13. ***Amaruup angatkuit takujanga.***
14. ***Nanuit inuuhukturmik aanniqsijuq.***
15. ***Angunahuktiit aarqijuq.***
16. ***Ilinniaqtitsiji qukiqtait.***
17. ***Qaijutit.***
18. ***Angunahuktimik aarqisijutit.***

(b) 翻译成 Inuktitut:

19. The shamanhurt you.
20. The teacher saw the boy.
21. Your wolf fell.
22. You shot a dog.
23. Your doghurt a teacher.

NB: Inuktitut (Canadian Inuit) belongs to the Eskimo-Aleut family of languages. 在加拿大北部, 约有35 000人使用该语言.

The letter ***r*** denotes a `Parisian' *r* (pronounced far back in the mouth), and ***q*** stands for a *k*-like sound made in the same place.

A shaman is a priest, sorcerer and healer in some cultures. —*Bozhidar Bozhanov*

编者: Alexander Berdichevsky, Bozhidar Bozhanov, Svetlana Burlak, Ivan Derzhanski (主编), Ludmilla Fedorova, Dmitry Gerasimov, Ksenia Gilyarova, Ivaylo Grozdev, Stanislav Gurevich, Adam Hesterberg, Boris Iomdin, Ilya Itkin, Renate Pajusalu, Alexander Piperski, Maria Rubinstein, Todor Tchervenkov.

中文文本: 刘闽晟.

祝你好运!

第六届国际理论, 数理及应用语言学奥林匹克竞赛

保加利亚, 阳光海滩, 2008年8月4 — 9日

个人赛解答

题 #1. 规则:

1. The apostrophe indicates length if it follows a vowel, and is read as ə if it follows a consonant.
2. The letter ***w*** stands for a rounding of the lips after a consonant and for the sound [w] otherwise.
3. ə is pronounced, though not written, between any consonant and a following sonorant consonant ([l m n]).
4. ə is also pronounced before a consonant cluster at the beginning of a word.
5. ***p t j g gw q qw*** are pronounced as 浊辅音 s (b d j g g^w γ $γ^w$) at the beginning of a word or between vowels and as voiceless consonants (p t c k k^w x x^w) at the end of a word or next to another consonant.

答案:

(a) 12 əksənxōγon, 13 ətkəbox, 14 gəmūjəmin, 15 emtoγ^{w}atk, 16 dēbəlc;

(b) 17 ***tp'te'sn***, 18 ***mtesgm***, 19 ***alapt'g***, 20 ***glamen***.

题 #2. **(a)** 规则:

1. Number of syllables. Each line contains 6 syllables.
2. Alliteration. See the statement of the problem.
3. Internal rhyme. Let us denote the vowels (and diphthongs) in each line by V_1, V_2, ⋯, V_6. At least one consonant immediately following V_5 must immediately follow V_n (n = 1, 2 or 3). Also, in even lines $V_n = V_5$.

For instance, cf. lines IV, 1–6 (alliteration is marked in boldface, internal rhyme by underlining):

IV

1 háði **g**ramr, þars **g**núðu,
2 **g**eira hregg við seggi,
3 (rauð fnýsti **b**en **b**lóði)
4 **b**ryngǫgl í dyn Skǫglar,
5 þás á **r**ausn fyr **r**æsi
6 (**r**éð egglituðr) seggir ⋯

(b) Leftover words: **hœgra**, **smíði**.

V

1 $_a$[**r**íks] (þreifsk **r**eiddra øxa
2 $_b$[**r**ymr] ; knǫ́ttu spjǫr $_c$[glymja])
3 $_d$[**s**vartskyggð] bitu **s**eggi
4 $_e$[**s**verð] þjóðkonungs ferðar,
5 þás ($_f$[**h**ugfyldra] **h**ǫlða)
6 $_g$[**h**laut] $_h$[andskoti] $_i$[Gauta]
7 (hǫr vas $_j$[sǫngr] of $_k$[svírum])
8 $_l$[sigr] (flugbeiddra $_m$[vigra]).

题 #3. The modifier follows its head in both languages.

(a)	***jun***	bone	
	i-jun	skeleton	(multitude of bones)
	i-wahnawa	bunch of bananas	(multitude of bananas)
	i-drai	calendar	(multitude of days)
	drai-hmitrötr	Sunday	(holyday)
	gaa-hmitrötr	sanctuary	(holyplace)
	uma-hmitrötr	church	(holyhouse)
	ngöne-uma	wall	(house border)
	ngöne-gejë	coast	(water border)
	nyine-thin	awl	(tool to poke)
	tii	to write	
	bé-tii	pencil	(tool to write)
	bé-wöli	fork	(tool to poke)
	wöta	animal	
	bé-wöli-wöta	spur	(tool to poke animal)
	bé-ôdu	cup	(tool to drink)
	ba-jié	coast	(water border)
	ba-bwén	twilight	(night border)
	a-pulut	bed	(place to sleep)

(b) ***wahnawa*** 'banana', ***drai*** 'day'; ***wöli*** 'to poke', ***pulut*** 'to sleep'.

(c) ***i-bii*** 'swarm of bees (multitude of bees)', ***tusi-hmitrötr*** 'Bible (holybook)'.

题 #4. The noun suffixes seen in this problem are:

1. **-k2sm2** '上方', **-kotoya** 'for', **-pit** `with';
2. **-šeh** 'like, as if';
3. **-taPm** plural;
4. **-tih** 'only (just, right)'.

After a nasal consonant (**m**, **n**, **ŋ**) the 塞音 s **p**, **t**, **k** become voiced (**b**, **d**, **g** respectively). If **k** comes after **y**, the two sounds exchange places.

The possessive pronouns are **P2s** `my' and **mis** `your'; if the noun begins with a 塞音, this consonant becomes voiced and the corresponding nasal appears before it.

(a)	**caky2sm2tih**	rightabove the vine
	k2m2ŋšeh	likea shadow
	P2s mok	my corn
	mis nd2ctaPm	your teeth
	pahsunbit	with the squash
	perolkotoyašehtaPm	as iffor the kettles

(b)	for the chair	**pokskukyotoya**
	with my kettle	**P2s mberolpit**
	justlikea mountain	**koc2kšehtih**
	posts	**komdaPm**
	above the shadows	**k2m2ŋg2sm2taPm**
	your town	**mis ŋgumguy**

题 #5. The Inuktitut sentences have the following general structure:

	X-***(q)***		V--	'X V (himself).'
	X-***(q)***	Y-***(r)mik***	V-***si***--	'X V a Y.'
X-***up***	Y-***(q)***		V--	'X V the Y.'

where X and Y are nouns and V is the verb. If a noun gets the ending ***-q*** when it is either a definite object or a subject of a sentence that doesn't have a definite object, it also gets ***-r*** before the ending ***-mik*** when it is an indefinite object (***nanu-q*** — ***nanu-r-mik***; ***iluaqhaiji*** — ***iluaqhaiji-mik***). To say `your', ***-(q)*** is replaced by ***-it***, ***-up*** by ***-vit***.

The verb receives the following suffixes:

- ***-j*** following a vowel or ***-t*** following a consonant;
- an ending for the persons of the subject and the definite object, if there is one:
  - in the first two schemata: ***-u-tit*** '2', ***-u-q*** '3';
  - in the third schema: ***-a-it*** '2/3', ***-a-nga*** '3/3', ***-a-atit*** '3/2'.

一个没有宾语的及物动词当作反身动词处理.

(a)
13. The wolf saw your shaman.
14. Your polar bear hurt a boy.
15. Your hunter cured himself.
16. You shot the teacher.
17. You came.
18. You cured a hunter.

(b)
19. ***Angatkuup aanniqtaatit.***
20. ***Ilinniaqtitsijiup inuuhuktuq takujanga.***
21. ***Amaruit ukiakhaqtuq.***
22. ***Qingmirmik qukiqsijutit.***
23. ***Qingmiit ilinniaqtitsijimik aanniqsijuq.***

第六届国际理论, 数理及应用语言学奥林匹克竞赛

保加利亚, 阳光海滩, 2008年8月4 — 9日

团体赛题目

在编译广韵字典时 (1007--1011), 汉语相当同质化. 由于汉字不是表音文字, 该字典使用了一套简单的系统, 通过两个汉字来表示一个汉字的发音, 而读者理应知道前者的发音 (它们是常用字). 这套系统叫做*反切*.

后来, 虽然汉语方言分化, 但许多古代的反切转写仍可使用, 只不过在不同的方言中有不同 (且更复杂) 的使用方法.

下面是一些反切转写. 每个汉字的广东话读音亦给出.

	汉字	=	转写	
1.	倦 $\mathbf{kyn}^{2}$	= 渠 $\mathbf{k^{h}œy}^{21}$	⋆ 卷 $\mathbf{kyn}^{3}$
2.	求 $\mathbf{k^{h}au}^{21}$	= 巨 $\mathbf{kœy}^{2}$	⋆ 鳩 $\mathbf{kau}^{53}$
3.	住 $\mathbf{cy}^{2}$	= 持 $\mathbf{c^{h}i}^{21}$	⋆ 遇 $\mathbf{y}^{2}$
4.	病 $\mathbf{piŋ}^{2}$	= 皮 $\mathbf{p^{h}ei}^{21}$	⋆ 命 $\mathbf{miŋ}^{2}$
5.	掉 $\mathbf{tiu}^{2}$	= 徒 $\mathbf{t^{h}ou}^{21}$	⋆ 弔 $\mathbf{tiu}^{3}$
6.	鳩 $\mathbf{kau}^{53}$	= 居 $\mathbf{kœy}^{53}$	⋆ 求 $\mathbf{k^{h}au}^{21}$
7.	僖 $\mathbf{hei}^{53}$	= 許 $\mathbf{hœy}^{35}$	⋆ 其 $\mathbf{k^{h}ei}^{21}$
8.	朗 $\mathbf{loŋ}^{13}$	= 盧 $\mathbf{lou}^{21}$	⋆ 黨 $\mathbf{toŋ}^{35}$
9.	韶 $\mathbf{siu}^{21}$	= 市 $\mathbf{si}^{13}$	⋆ 昭 $\mathbf{ciu}^{53}$
10.	帳 $\mathbf{cœŋ}^{3}$	= 知 $\mathbf{ci}^{3}$	⋆ 亮 $\mathbf{lœŋ}^{2}$
11.	愀 $\mathbf{c^{h}iu}^{35}$	= 親 $\mathbf{c^{h}an}^{3}$	⋆ 小 $\mathbf{siu}^{35}$
12.	舞 $\mathbf{mou}^{13}$	= 文 $\mathbf{man}^{2}$	⋆ 甫 $\mathbf{p^{h}ou}^{35}$
13.	謏 $\mathbf{siu}^{35}$	= 先 $\mathbf{sin}^{53}$	⋆ 鳥 $\mathbf{niu}^{13}$
14.	臼 $\mathbf{k^{h}au}^{13}$	= 其 $\mathbf{k^{h}ei}^{21}$	⋆ 九 $\mathbf{kau}^{35}$
15.	斜 $\mathbf{c^{h}e}^{21}$	= 似 $\mathbf{c^{h}i}^{13}$	⋆ 嗟 $\mathbf{ce}^{53}$
16.	薨 $\mathbf{kau}^{3}$	= 古 $\mathbf{ku}^{35}$	⋆ 候 $\mathbf{hau}^{2}$

(a) 解释古代反切转写是如何应用于现代广东话的.

(b) 在编译广韵时反切转写是如何工作的? 在广东话中, 上述转写只有一个可以应用这条简单的老规则来得到正确的结果. 哪一个呢?

在多数当代汉语方言中 (包括广东话和普通话) 不存在浊辅音, 除了响音 ($\mathbf{l}$, $\mathbf{m}$, $\mathbf{n}$, $\mathbf{ŋ}$). 在广韵编译时汉语存在其他浊辅音, 它们后来并成对应的清音: 浊擦音变成清擦音 (e. g., $\mathbf{z} > \mathbf{s}$), 浊塞音变成送气或不送气清塞音 (e. g., $\mathbf{d} > \mathbf{t}$ or $\mathbf{t^{h}}$). 浊音在吴语中得到了保留. 比如, 汉字 徒 在吴语中发作 [$\mathbf{du}^{21}$], 在广东话中发作 [$\mathbf{t^{h}ou}^{21}$], 在普通话中发作 [$\mathbf{t^{h}u}^{35}$].

(c) 上一节中的哪些汉字在编译广韵时首辅音为浊辅音? 这些浊辅音在广东话中变得送气或者不送气取决于什么条件?

(d) 在古汉语中存在四种声调, 但是只有三种出现在此题中. 解释这三种音调是如何演变出广东话的六种音调.

下面是另外一些转写, 但只给出了其普通话读音:

17. 遷 **ɕan**^{5} = 張 **ɕaŋ**^{5} ⋆ 連 **lian**^{35}
18. 良 **liaŋ**^{35} = 呂 **ly**^{214} ⋆ 章 **ɕaŋ**^{5}
19. 遵 **cun**^{5} = 將 **ƙiaŋ**^{51} ⋆ 倫 **lun**^{35}
20. 蕭 **ɤ́iao**^{5} = 蘇 **su**^{5} ⋆ 彫 **tiao**^{5}
21. 嵌 **ƙ^{h}ian**^{5} = 口 **k^{h}ou**^{214} ⋆ 銜 **ɤ́ian**^{35}
22. 先 **ɤ́ian**^{5} = 蘇 **su**^{5} ⋆ 前 **ƙ^{h}ian**^{35}
23. 巉 **ɕ^{h}an**^{35} = 鋤 **ɕ^{h}u**^{35} ⋆ 銜 **ƙ^{h}ian**^{35}
24. 婻 **ɤ́iŋ**^{51} = 胡 **xu**^{35} ⋆ 頂 **tiŋ**^{214}
25. 弗 **ɕ^{h}an**^{214} = 初 **ɕ^{h}u**^{5} ⋆ 限 **ɤ́ian**^{51}
26. 趡 **c^{h}uei**^{214} = 千 **ƙ^{h}ian**^{5} ⋆ 水 **ʂuei**^{214}
27. 初 **ɕ^{h}u**^{5} = 楚 **ɕ^{h}u**^{214} ⋆ 居 **ƙy**^{5}
28. 釧 **ɕ^{h}uan**^{51} = 尺 **ɕ^{h}** ^{214} ⋆ 絹 **ƙyan**^{51}
29. 卷 **ƙyan**^{214} = 居 **ƙy**^{5} ⋆ 轉 **ɕuan**^{214}
30. 處 **ɕ^{h}u**^{51} = 昌 **ɕ^{h}aŋ**^{5} ⋆ 據 **ƙy**^{51}
31. 俜 **p^{h}iŋ**^{5} = 普 **p^{h}u**^{214} ⋆ 丁 **tiŋ**^{5}
32. 蚪 **tou**^{214} = 當 **taŋ**^{5} ⋆ 口 **k^{h}ou**^{214}

(e) 暂时忽略音调, 给出在普通话中使用古代反切转写的规则.

下面是些汉字及其广东话和普通话的读音:

		广东话	普通话
33.	唐	**t^{h}oŋ**^{21}	**t^{h}aŋ**^{35}
34.	謨	**mou**^{21}	**mo**^{35}
35.	踐	**c^{h}in**^{13}	**ƙian**^{51}
36.	少	**siu**^{35}	**ʂao**^{214}
37.	夔	**k^{h}wai**^{21}	**k^{h}uei**^{35}
38.	你	**nei**^{13}	**ni**^{214}
39.	暫	**caam**^{2}	**can**^{51}

		广东话	普通话
40.	采	**pin**^{2}	**pian**^{51}
41.	帝	**tai**^{3}	**ti**^{51}
42.	透	**t^{h}au**^{3}	**t^{h}ou**^{51}
43.	被	**p^{h}ei**^{13}	**pei**^{51}
44.	囂	**hiu**^{53}	**ɤ́iao**^{5}
45.	枌	**fan**^{21}	**fen**^{35}

(f) 描述音调和浊首辅音是如何在普通话中演变的. 可以总结出哪些在普通话中读反切转写的音调的规则?

(g) 一些初始的辅音和音调组合在现代普通话中极为罕见. 哪些呢?

下面是另外一些汉字及其广东话和普通话的读音. 一些音调被移除:

		广东话	普通话
46.	罿	**t^{h}uŋ**⋯⋯	**t^{h}uŋ**^{35}
47.	載	**coi**^{3}	**cai**⋯⋯
48.	米	**mai**⋯⋯	**mi**^{214}

		广东话	普通话
49.	眠	**min**^{21}	**mian**⋯⋯
50.	蛸	**siu**⋯⋯	**ɤ́iao**^{5}
51.	亂	**lyn**⋯⋯	**luan**^{51}

(h) 判断出遗失的音调是哪些.

(i) 给出下面的转写在广东话中的读法:

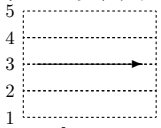
52. 梯 ? = 土 **t^{h}ou**^{35} ⋆ 雞 **kai**^{53}
53. 嘯 ? = 蘇 **sou**^{53} ⋆ 弔 **tiu**^{3}
54. 浪 ? = 魯 **lou**^{13} ⋆ 當 **toŋ**^{53}
55. 憊 ? = 蒲 **p^{h}ou**^{21} ⋆ 拜 **paai**^{3}

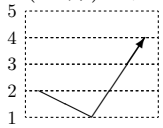
(j) 给出下面的转写在普通话中的读法. 一些转写本不可读, 但这道题包含了足以读出它们的信息:

56.	賽 ? = 先 **x́ian**^{5}=13A=22X	⋆ 代 **tai**^{51}
57.	簡 ? = 古 **ku**^{214}=16A	⋆ 限 **x́ian**^{51}=25B
58.	賞 ? = 書 **ʂu**^{5}	⋆ 兩 **liaŋ**^{214}
59.	倌 ? = 普 **p^{h}u**^{214}=31A	⋆ 乃 **nai**^{214}
60.	泫 ? = 胡 **xu**^{35}=24A	⋆ 畎 **k̑^{h}yan**^{214}
61.	犬 ? = 苦 **k^{h}u**^{214}	⋆ 泫 =60X
62.	下 ? = 胡 **xu**^{35}=24A	⋆ 駕 **k̑ia**^{51}
63.	捍 ? = 下 =62X	⋆ 赧 **nan**^{214}
64.	紂 ? = 除 **ç^{h}u**^{35}	⋆ 柳 **liou**^{214}
65.	囊 ? = 奴 **nu**^{35}	⋆ 當 **taŋ**^{5}=32A=54B
66.	鰓 ? = 蘇 **su**^{5}=20A=22A=53A	⋆ 來 **lai**^{35}

NB: 普通话是中国的官方语言, 基于北京方言. 约八亿五千万人使用该语言. 九千万人使用吴语 (上海话), 七千万人使用广东话 (粵语).

每个汉语方言都有固定数目的音调 (每个音节发音的旋律). 本题使用了语言学家赵元任提出的系统, 其用数字 1 (最低) 到 5 (最高) 来标记音高的五级, 并将音调转写成连续的音级:

a^{3} 

, **a**^{214} 

. All the tones you need are present in this problem.

标记 $^{\mathbf{h}}$ indicates that the preceding 塞音 consonant is aspirated (pronounced with a puff of air). **x** = *ch* in Scottish *loch*, **ŋ** = *ng* in *hang*. **c** ≈ *ts* in *hats* (pronounced as a single consonant), **ʂ** and **ç** are hard consonants similar to English *sh* in *shut* and *ch* in *chuck*, **x́** and **k̑** are soft consonants similar to *sh* in *sheet* and *ch* in *cheat*. **œ** and **y** = French *eu* and *u* (German *ö* and *ü*).

如果你不想写汉字, 你可以使用转写的序号并指明具体的汉字来指代它们: X (被转写的), A (第一个用于转写的汉字) or B (第二个用于转写的汉字).

注意普通话的 28A 汉字的读法不包含元音. —*Todor Tchervenkov*

中文文本: 刘闽晟.

祝你好运!

第六届国际理论, 数理及应用语言学奥林匹克竞赛

保加利亚, 阳光海滩, 2008年8月4 — 9日

团体赛解答

汉语音节由三部分组成: 声母 (首辅音, 可能不存在如 3B), 韵母 (后面的所有音) 和声调. 广东话音调可以认为存在两种不同的性质: 音高 (高 or 低) 和轮廓 (升, 平 or 降).

	升	平	降
高	35	3	53
低	13	2	21

(a) 若要在广东话中使用反切转写, A 的声母和声调音高将于 B 的韵母及声调轮廓组合. But if A's (and X's) tone is 低, X's onset, if a 塞音, must always be 送气 if B's (and X's) tone is 升 (13) or 降 (21), and un 送气 if it is 平 (2).

(b) Certainly the onset was from the A character, and the rhyme from B. But the aspiration rule is strange. Probably it was not part of the original fanqie system. Maybe the tone came from only one of the two characters? That has to be B, because the old rule should give correct results in only one 转写.

Thus the original simple rule for fanqie was: A's onset is combined with B's rhyme and tone. Only 转写 11 can be read now using this rule.

(c) Looking at the syllables with a sonorant onset, we see that they are always in a low tone (13, 2 or 21). Assuming that all 浊辅音 s evolved alike in Cantonese, we may conclude that what is in a low tone now, had a voiced onset earlier. This is also true of the character of the example from Wu. What is said in **(d)** supports this idea.

Thus the characters whose onsets were voiced are: 1X and 1A, 2X (=6B) and 2A, 3X and 3A, 3B (if it had an onset at all), 4X and 4A, 5X and 5A, 7B (=14A), 9X and 9A, 14X, 15X and 15A, 16B.

Voiced 塞音 s became 送气 if the tone was 升 or 降, and un 送气 if it was 平.

(d) The contours of the Cantonese tones correspond to the three tones of Classical Chinese; tone height is an innovation brought about by the evolution of the 浊辅音 s.

因此我们可以解释为什么反切转写在广东话中这样读. The X character has the same tone height as A because it got its onset from A, and height in Cantonese is determined by the voicing of the onset in Classical Chinese. But if the onset was a voiced 塞音, it could evolve in different ways in X and A, because its aspiration was determined by the tone contour, which X got from B, and it could differ from A's contour.

(e) In Mandarin onsets and rhymes are not combined in such a straightforward way as in Cantonese. It can be noted that after **ɤ́** (**k̑**, **k̑ʰ**) we always find **i** or **y**, whereas **x** (**k**, **kʰ**), **s** (**c**, **cʰ**) and **ʂ** (**ç**, **çʰ**) are never followed by these vowels.

We already know that the onset came from A and the rhyme from B. When the constraint above came into being,

- **i** was lost and **y** became **u** after **ʂ** (**ç**, **çʰ**);
- **x** (**k**, **kʰ**) and **s** (**c**, **cʰ**) became **ɤ́** (**k̑**, **k̑ʰ**) before **i** or **y**.

在普通话中使用反切转写同样需要应用如上规则. However,

- if A's onset is **x́** (**ƙ**, **ƙ^{h}**) and B's rhyme starts with neither **i** nor **y**, we can't determine what X's onset is;
- if B's onset is **ʂ** (**ɕ**, **ɕ^{h}**) and A's onset is none of these, we can't determine what X's rhyme is.

(f) On the basis of the tone of the Cantonese syllable we can determine whether the onset was voiced or not in Classical Chinese. In Mandarin the tones developed as follows:

- 升: 51 if the onset was voiced but not a sonorant, 214 otherwise;
- 平: 51 (always);
- 降: 5 if the onset was voiceless, 35 otherwise.

We see that the contour is not preserved here. Voiced 塞音 s became 送气 if the tone was 降, and un 送气 if it was 平 or 升.

In 反切转写 s read 在普通话中 the tones work as follows:

	5, 35	214	(F, H−)^{51}	(H+, L)^{51}
5	5	214	214, 51	51
L^{35}	35	214	214, 51	51
(F, H+)^{35}	35	51	51	51
L^{214}	35	214	214, 51	51
(F, H±)^{214}	5	214	214, 51	51
L^{51}	35	214	214, 51	51
H+^{51}	5	214	214, 51	51
(F, H−)^{51}	5, 35	214, 51	214, 51	51

Here L stands for a sonorant, F for a 擦音, H− for an un 送气 and H+ for an 送气 塞音. Thus most of the time X's tone in Mandarin can't be derived unambiguously from A's and B's tones, though in some cases it can.

(g) Syllables with a sonorant onset and tone 5 or with an un 送气 onset and tone 35 should not exist in Mandarin (if they do, then the rules must have had exceptions).

(h) 46: **21**, 47: **51**, 48: **13**, 49: **35**, 50: **53**, 51: **2**.

(i) 52 **t^{h}ai**^{53}, 53 **siu**^{3}, 54 **loŋ**^{2}, 55 **paai**^{2}.

(j) 56 **sai**^{51}, 57 **ƙian**^{214}, 58 **ʂaŋ**^{214}, 59 **p^{h}ai**^{214}, 60 **x́yan**^{51}, 61 **ƙ^{h}yan**^{214}, 62 **x́ia**^{51}, 63 **xan**^{51}, 64 **ɕou**^{51}, 65 **naŋ**^{35}, 66 **sai**^{5}.