

fa

بیست و یکمین المپیاد جهانی زبان شناسی

برازیلیا (برزیل)، 23-31 ژوئیه 2024

مسأله مسابقه گروهی

آمارواژگان — مجموعه از روش‌هاست که برای ارزیابی میزان نزدیکی زبان‌ها بین خود و سایر زبان‌ها بر اساس واژگان استفاده می‌شود. طبق قاعده این روش‌ها برای فهرست بلند لغات نوشته شده به دست کارشناسان استفاده می‌شوند. کارشناسان برای هر جفت مشخص می‌کنند که آیا آنها به یک منبع وصل می‌شوند. ولی گاهی اوقات زبان‌شناسان روش‌های آمارواژگان را برای فهرست لغاتی که اتوماتیک تهیه شده‌اند استفاده می‌کنند. یکی از این روش‌ها بر اساس مفهوم کلاس‌های کنستانت بنا شده است که توسط زبان‌شناس شوروی-اسرائیلی آ.ب. دولگوپولسکی در 1964 معرفی شده است.

$pb\phi\beta fv$	P.	$kgxyqg\chi\psi$	K.	$j\zeta$ (در اول ریشه)	Y.	$m\eta$	M.
$td\theta\delta\tau d$	T.	$rr\tau al\text{†}\ddot{k}[\text{†}l$	R.	wM (در اول ریشه)	W.	$n\eta n\eta$	N.
$sz\int z\int z\zeta z\zeta z c\int$	S.					$\text{†}\ddot{d}\ddot{g}$	Q.

H. $h\int h\int \text{†}h\text{†}?$ صدادار و $z\zeta wM$ (به غیر از شروع ریشه)

کلاس‌های کنستانت دولگوپولسکی

پایین تر شما قسمتی از فهرست لغات نشانه‌گذاری شده برای چند زبان همخانواده دنیا را می‌یابید. نشانه‌گذاری با اعداد زیرنویس داده شده است. طبق این لیست درختانی ساخته شده‌اند که با زبان‌های همخانواده با کمک دو گزینه ساده شده ای با نام الگوریتم StarlingNJ و با هر کلمه شاخص ثابت دارای زیرنویس متناسب هستند. درختان و شاخص‌های ثابت بالای صفحه با لیست نشانه‌گذاری دستی ساخته شده‌اند و پایین صفحه لیست با نشانه‌گذاری اتومات. برای هر لیست دو درخت بر اساس الگوریتم آ و ب داده شده. توجه کنید که در برخی شرایط چند درخت ممکن است که متناسب با یک لیست لغات هستند در چنین شرایطی یک درخت اتفاقی انتخاب شده است. برای هر گره از هر درخت فاصله آمارواژگان مشخص شده است. هرچقدر معانی شان بیشتر باشد زبان‌ها نزدیک‌ترند. بنابراین ”فاصله آمارواژگان معکوس“ اصطلاح دقیق‌تر از ”فاصله آمارواژگان“ خواهد بود. برای سادگی در این مسأله از عبارت ”فاصله آمارواژگان“ استفاده می‌کنیم.

هم شاخصه ثابت هم فاصله آمارواژگان تا دو علامت بعد از ممیز رند شده‌اند. اگر علامت سوم بعد از ممیز کمتر از 5 است، معنی به طرف پایین رند می‌شود در غیر این صورت بالا رند می‌شود. 2.836 تا 2.84 رند می‌شود، 0.705 تا 0.71، و 0.703 تا 0.70. رند شدن تنها برای معانی ای که برای کاربر قابل مشاهده است بکار برده می‌شود. به عبارت دیگر کامپیوتر در پروسه اعمال الگوریتم‌ها معنای رند نشده را ”می‌بیند“.

توجه داشته باشید که بعضی لغات احتمالاً یا قطعاً از زبان دیگر وام گرفته شده‌اند. برای مثال کلمه $jok:i$ ’نمک‘ زبان کادیوتو از گوارانی $juki$ گرفته شده است، $\text{?}a:n\text{?}$ ’سال‘ زبان ایبای (مساگرانده) از $\text{?}ano\text{?}$ اسپانیایی. گاهی برای یک معنی از طریق ممیز چند مترادف داده شده‌است. کلماتی با معنی ’پا‘ به زبان زبان و جوز را می‌توان به عنوان مثال داد.

در اطلاعات داده شده ذیل تمام پیشوندها با علامت ’= $’$ و پسوندها با علامت ’= $-$ ‘ جدا شده‌اند. بعضی لغات بدون پیشوند استفاده نمی‌شوند. آنها با علامت ’= $’$ شروع می‌شوند.

اطلاعات با الفبا آوایی بین‌المللی آوانویسی شده‌اند. ’= $’$ تکیه اصلی، ’= 1 ‘ تکیه فرعی (ضعیف تر از اصلی)، ’= \circ ‘ صدای بلند، ’= \bullet ‘ صدای خیلی کوتاه، $X = \overline{XY}$ و Y مانند یک صدا تلفظ می‌شوند، ’= \acute ‘ لحن بالا، ’= \grave ‘ لحن پایین، ’= \hat ‘ لحن رو به پایین، ’= $^?$ ‘ حرف صامت از نای (که قبل آن نای کمی منقبض می‌شود)، ’= \circ ‘ صامت گرینشی (با انقباض کوتاه نای تلفظ می‌شود)، ’= \circ ‘ صامت بیصدا، ’= \circ ‘ صدا خیشومی شده (از بینی ادا می‌شود)، ’= \circ ‘ صدای جیرجیر (صدای زیر، جیرجیر)، ’= \circ ‘ جریان هوا از بینی قبل از صامت را مشخص می‌کند، ’= \circ ‘ صامت حلقوی (با بیرون دادن نفس تلفظ می‌شود)، ’= \circ ‘ صامت شفوی (با لبان گرد شده تلفظ می‌شود)، ’= \circ ‘ صامت نرم. $\theta, \theta, y, \theta, \theta, \Lambda, \theta, \mu, \upsilon, \theta, \text{†}, I, \epsilon, \theta, \alpha$ صدادار هستند.

علامت‌های دیگر صامت را مشخص می‌کنند. Δ دانش هر یک از زبان‌های ذکر شده در مسئله، مزیتی در حل آن ندارد.

بخش ۱. خانواده گواکورویی (آرژانتین، برزیل، پاراگوئه)

کادیوئو	موکوی (چاکو)	پیلاگا	توبا (شرقی)	
lol:adi ₃	naweyelek ₂	'lo=?ok ₁	l=?ok ₁	ابر
n=ol:edi ₂	norek ₁	'd=ole?2	nodek ₁	آتش
nij:ogo-d̄zegi ₃	naʎin ₂	'nijaq ₁	njaq ₁	ماهی
=ak:ilo ₂	=qaik ₁	'qajk ₁	=qajk ₁	سر
=el:owadi ₁	=alawat ₁	=a'la:t ₁	=alawat ₁	کشتن
ep:enaj ₃	ʃirajyo ₂	?a'woʃojk ₁	?awoʃojk ₁	ماه
=m:iq:o ₁	=mik ₁	'mik ₁	=mik ₁	بینی
jok:i ₁	?we ₁	ol'ʎek ₂	towe ₁	نمک
wet:iga ₂	qa?1	'qa?1	qa?1	سنگ
=ok:el:i ₃	=o?ley-akʌn-akʌt ₂	=a'tʃ-aʃat ₁	=atʃ-akʌt ₁	زبان

	الگوریتم B	الگوریتم A	
0.50 ابر 0.50 آتش 0.50 ماهی 0.75 سر 1.00 کشتن 0.50 ماه 1.00 بینی 0.67 نمک 0.75 سنگ 0.50 زبان شاخصه ثابت:			دستی
0.50 ابر 0.50 آتش 0.75 ماهی 0.75 سر 1.00 کشتن 0.50 ماه 1.00 بینی 0.25 نمک 0.75 سنگ 0.50 زبان شاخصه ثابت:			خودکار (تومات)

بخش ۲. خانواده نوبی ها (مصر، سودان)

بیرگید	دبری	کادارو	دیلینگ	کنوزی	دنگولوی	
fila:l-e ₁	wur-i ₂	wur-i ₂	hur ₂	be:₁	'be:₁	کشتن
ma:l ₂	nən-to ₁	nən-tu ₁	nən-ti ₁	an-at-ti ₁	u'n-at-t ₁	ماه
eji ₁	ɔtu ₁	ɔto ₁	ɔti ₁	essi ₁	'ess ₁	آب
te:-n ₁	ti ₁	ti ₁	ti ₁	tir ₁	'tir ₁	دادن
azze-n ₃	kɛŋ ₂	kɛn ₂	ken ₂	sere:₁	'sɛrɛ:₁	خوب
kurr-i ₃	irf-o ₂	irf-o ₂	irf-i ₂	turug ₁	'turug ₁	باد
ur=dill-e ₁	til-tu ₁	til-tu ₁	tel-ti ₁	si:r ₂	'dil-ti ₁	مو
tu:₁	to ₁	to ₁	te-te ₂	tu:₁	'tu:₁	دل
ne:r-i ₁	jer-i ₁	dwallɛli ₂	jer ₁	ne:r ₁	'ne:r ₁	خوابیدن
?i:zi ₂	ɛngal-to ₃	aju ₂	ɛj ₂	masi ₁	'masi ₁	خورشید

		الگوریتم B	الگوریتم A			
0.50	کشتن			دستی		
0.83	ماه					
1.00	آب					
1.00	دادن					
0.50	خوب					
0.50	باد					
0.83	مو					
0.83	دل					
0.83	خوابیدن					
0.50	خورشید					
0.33	کشتن			خودکار (تومات)		
0.50	ماه					
0.50	آب					
0.67	دادن					
0.50	خوب					
0.50	باد					
0.83	مو					
1.00	دل					
0.50	خوابیدن					
0.50	خورشید					

(A) (2 امتیاز) صامت **B** مثل حرف **r** در زبان فرانسه، با قسمت پشتی زبان تلفظ می شود (= غ). این به کدام کلاس دولگوپولسکی مربوط می شود و شما چگونه به این پی بردید؟

(B) (2 امتیاز) درخت نوبی ها در گوشه بالا سمت چپ تنها یکی از دو امکان موجود برای این ترکیب الگوریتم و نوع نشانه گذاری است. درخت دیگری ممکن را بکشید.

(C) (2 امتیاز) درخت نوبی ها در گوشه پایین سمت چپ تنها یکی از دو امکان موجود برای این ترکیب الگوریتم و نوع نشانه گذاری است. درخت دیگری ممکن را بکشید.

(D) (2 امتیاز) فاصله آمارواژگان 0.49 در ریشه درخت نوبی ها در گوشه سمت راست بالا مانند سایر فاصله ها در این مسئله بعد ممیز تا دو علامت رند شده اند. فاصله دقیق چیست؟

بخش ۳. خانواده ماتاگویان (آرژانتین، بولیوی، پاراگوئه)

ماکا	نیواکله (چیشامنه هلاوس)	نیواکله (شیچام هلاوس)	مانخوی	ایوجایاها	وینهایک	وجوز	ویچی (ریوادویا)	ویچی (برمهو پایینی)	
fe [?] t ₂	ʔitax ₁	ʔitax ₁	'ʔeit'e ₁	'hwat ₂	ʔi:tax ₁	ʔitah ₁	ʔitox ₁	ʔitox ₁	آتش
sehets ₋₁	saxetʃ ₋₁	saxetʃ ₋₁	ʃi [?] jus ₋₁	si [?] jus ₋₁	ʔwa:hat ₁	wahat ₁	wahat ₁	ʔwahat ₁	ماهی
=f'i [?] ʔ ₅	=φo [?] ʔ ₄	=φo [?] ʔ ₄	=ka [?] la [?] ʔ ₂	= [?] sat ₃	=pa:k [?] o [?] ʔ ₁	=patʃ [?] o ₁ , =kala ₂	=qol ₂	=patʃ [?] u ₁	پا
iweli [?] ʔ ₃	jina [?] t ₁	jina [?] t ₁	ʔa [?] nat ₁	ʔi [?] n [?] at ₁	ʔina:t ₁	wah ₂	ʔinot ₁	ʔinot ₁	آب
tis-ix ₅	=xut-ej ₄	=xut ₄	= [?] haj [?] ʔ ₃ , = [?] wɛn ₂	= [?] wɛhn-a [?] m ₂	= [?] wɛŋ-o [?] ʔ ₁	= [?] wɛŋ-o ₁	=wɛŋ-u ₁	= [?] wɛŋ-u ₁	دادن
t=ejk'un-ej ₂	ʔis ₁	ʔis ₁	'ʔeis ₁	'ʔes ₁	ʔis ₁	ʔis ₁	ʔis ₁	ʔis ₁	خوب
t'unik'i ₆	ʔaβi [?] m ₅	ʔaβi [?] m ₅	'hlahwuu [?] ʔ ₄	'hlahwu [?] ʔ ₄	=ja:ʔ ₂ , =x ^w ox ^w ₃	ʔihwok ^w ₁	ʔinwək ₁	ʔinwok ^w ₁	باد
naxka-k ₃	ji [?] kla [?] ʔ ₁	ʔa [?] kxi-juk ₂	ʔa [?] la-k ₁	ʔa [?] la [?] ʔ ₁	ha [?] la [?] ʔ ₁	ha [?] la ₁	hal ₁	ha [?] lo ₁	درخت
=ʔewkux-its ₅	=je [?] s ₄	=jate [?] t ₃	= [?] wole-j ₁	= [?] wole ₁	= [?] wo:le-ç ₁ , hi:lenax ₂	= [?] wole-j ₁	=wule-j ₁	= [?] wule-j ₁	مو
=lan ₁	=klan ₁	=klan ₁	=lan ₁	=la [?] ʔan ₁	=la:ŋ ₁	=lan ₁	=lon ₁	=lon ₁	کشتن

	الگوریتم B	الگوریتم A	
<p>تشخیصه بیت:</p> <p>0.78 آتش</p> <p>1.00 ماهی</p> <p>0.33 پا</p> <p>0.78 آب</p> <p>0.44 دادن</p> <p>0.89 خوب</p> <p>0.33 باد</p> <p>0.78 درخت</p> <p>0.67 مو</p> <p>1.00 کشتن</p>	<p>پایینی (برمهو ویچی)</p> <p>ویچی (ریواداویا)</p> <p>وجوز</p> <p>وینهایک</p> <p>ایوجایاها</p> <p>مانخوی</p> <p>هلاوس (شیچام نیواکله)</p> <p>هلاوس (چیشامنه نیواکله)</p> <p>ماکا</p>	<p>پایینی (برمهو ویچی)</p> <p>ویچی (ریواداویا)</p> <p>وجوز</p> <p>وینهایک</p> <p>ایوجایاها</p> <p>مانخوی</p> <p>هلاوس (شیچام نیواکله)</p> <p>هلاوس (چیشامنه نیواکله)</p> <p>ماکا</p>	دستی
<p>تشخیصه بیت:</p> <p>0.78 آتش</p> <p>0.44 ماهی</p> <p>0.33 پا</p> <p>0.56 آب</p> <p>0.67 دادن</p> <p>0.89 خوب</p> <p>0.22 باد</p> <p>0.67 درخت</p> <p>0.67 مو</p> <p>1.00 کشتن</p>	<p>پایینی (برمهو ویچی)</p> <p>ویچی (ریواداویا)</p> <p>وجوز</p> <p>وینهایک</p> <p>ایوجایاها</p> <p>مانخوی</p> <p>هلاوس (شیچام نیواکله)</p> <p>هلاوس (چیشامنه نیواکله)</p> <p>ماکا</p>	<p>پایینی (برمهو ویچی)</p> <p>ویچی (ریواداویا)</p> <p>وجوز</p> <p>وینهایک</p> <p>ایوجایاها</p> <p>مانخوی</p> <p>هلاوس (شیچام نیواکله)</p> <p>هلاوس (چیشامنه نیواکله)</p> <p>ماکا</p>	خودکار (انومات)

بخش ۴. خانواده مغولی (جمهوری خلق چین، مغولستان، روسیه)

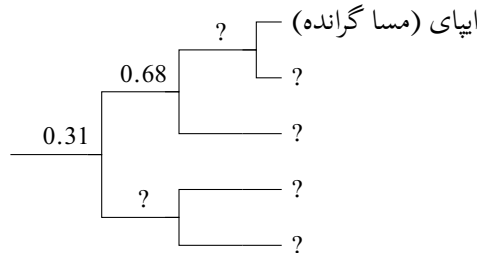
(E) (10 امتیاز) فهرست بعدی را بررسی کنید. شاخص ثبات متناسب با درختان با نشانه گذاری دستی یا اتوماتیک را حساب کنید. برای اینکه به شما کمک کنیم ما هر دو شاخص ثابت را برای معنی 'همه' آماده کرده‌ایم. در 0.36 و 0.40 به صورت رندوم داده شده‌اند.

بونا	شیرا یوگور	اردوس	خالخا	کالمیک	خوشوت	الیود	نیوبارگوت	بوریات (خوری)	خامنیگان (منچو)	داگور (هایلار)	
hanə ₋₂	tʃ ^h uq ₅	pyyyte ₄ , xamuk ^h ₋₁	pux ₃ , pugt ₄ , xamäg ₋₁	tsuk ₅ , xamak ₋₁	lug ₅	tsug ₅	bygd ₄	bʊxi ₋₃	bʊlt ₂	hə: ₁	همه
arasun ₄	χalsən ₁	turusu ₃	xəɫt ^h ɔs ₂	dursn ₃	xalis ₁	xolts ₂	xalʃhu: ₁	χəltəhən ₂	qalihən ₁	hails ₁	پوست درخت
kele ₁	ketesən ₂	ketysy ₂	gitis ₂ , xiwɫɟij ₋₁	gesn ₂	gets ₂	ge:s ₂	gedy: ₂	gedehen ₂	getəhən ₂	ke:li ₁	دل
bendzər ₂	ʃu:n ₁ , peltʃər ₂	ʃuβu: ₁	ʃuwu ₁	ʃowun ₁	ʃuwu: ₁	ʃuvu: ₁	ʃuwu: ₁	ʃubu:n ₁	eiwan ₁	dəgi ₋₁	پرند
χal ₁	qal ₁	qal ₁	gal ₁	gal ₁	gal ₁	gal ₁	gal ₁	gal ₁	gal ₁	gali ₁	آتش
mor ₄	mør ₄	tʃam ₋₁	tsam ₋₁	xa:l-kə ₃	dzam ₋₁	dzam ₋₁	zam ₋₁	χargi ₂ , zam ₋₁	qargōi ₂	terg-u:l ₁	راه
dabsuŋ ₂	ta:psən ₂	taβusu ₂	tawsä ₂	dawsn ₂	daws ₂	daws ₂	dawuhu: ₂	dabhan ₂	dawhən ₂	hata: ₁	نمک
mba ₋₁	umpa ₋₁	usu-tʃ ^h i-la ₋₄	siɫi ₋₃	us-tʃi ₋₄ , ø:m ₋₅	umba ₋₁	sele ₋₃	umb ₋₁	t ^h amar ₋₂	umba ₋₁	unpa-du ₁	شنا کردن
sə ₁	q ^h usun ₁	usun ₁	ʊsč ₁	usn ₁	us ₁	usn ₁	u:ha ₁	uhan ₁	ʊxʊn ₁	ʊsɔ ₁	آب
ki ₁	k ^h i: ₁	k ^h i: ₁	salɟxi ₂	salʃkn ₂	salki ₂	salʃxin ₂	halxi ₂	halxin ₂	halkin ₂	kei ₁	باد

بخش ۵. خانواده یومانی (مکزیک، ایالات متحده آمریکا)

(F) (8 امتیاز) فهرست بعدی را بررسی کنید. ذیل درختی داده شده که بر اساس همان لیست بنا شده است. برخی داده‌ها (اسامی زبان‌ها و فاصله آمارواژگان) جا افتاده است. جای خالی را پر کنید. ذکر کنید که آیا درخت در نشانه گذاری به دست کارشناسان یا اتوماتیک بنا شده است و همچنین آیا به کمک الگوریتم آ یا ب ساخته شده است.

موجاوه	کوکوپا	یاواپای	تیپای (جامل)	ایپای (مسا گرانده)	(مسا)
wena=wen-a ₁	'xʌ=ʔut ₂	'tʃkr=ot-i ₂	lə=ʔuj ₁	mə=put-k ₃	کوتاه
ʔitʃ=i=jer ₁	'ʃa ₂	'ʔ=tʃ=sa ₂	aʔ='ʃa ₂	ʔa:ʃa:₂	پرنده
ɲ=a=s=ak ₁	'ɲ=j=a:k ₁	'tʃ=j=a:k-a ₁	'ak ₁	aq ₁	استخوان
i=ro:z-v-k ₁	's=ʔar ₂	'ru-β-i ₁	's=ʔa:j ₃	sa:j ₃	خشک
kʷi:kʷay ₁	ʔi='ma:tʃ ₂	'kʷe:='θo-β-a ₃	'kʷak ₄	kukʷa:j-p ₁	گوشت
maʎaqe ₁	'm=puk ₂	'mlq ₁	i:='puk ₂	i:='puk ₂	گردن
i=ju:-k ₁	'wi:₂	'ʔu:₁	'wi:w ₂	ə=wu:w ₂	دیدن
i:='ʔar ₁	'ʃ=juʎ ₂	'β=hé ₃	ʃə='juʎ ₂	xə=juʎ ₂	دم
havik-k ₁	'x=wak ₁	'hʷâk-i ₁	xə='wak ₁	xə=wak ₁	دو
hu:ðe ₁	'mat-'ka:m ₂	'ʔ=tʃʰur-a ₃	mat-'wam ₂	ʔa:n ₁	سال



(G) (20 امتیاز) تعدادی درخت برای زبان‌های یومانی با فاصله آمارواژگان در ریشه (یعنی در آخر درخت سمت چپ) بوجود آمده بودند:

۱. 0.20
۲. 0.23
۳. 0.24

هر یک از این درختان را بکشید. برای هر درخت ذکر کنید که نشانه گذاری به دست کارشناسان یا اتوماتیک بنا شده است و همچنین آیا به کمک الگوریتم آ یا ب ساخته شده است.

(H) (3 امتیاز) دو فاصله ذکر شده در تکلیف (G) تا دو علامت بعد از ممیز رند شده‌اند: معنی 0.23 از طریق رند کردن 0.225 بدست آمده است. دیگر کدام فاصله رند شده و کدام معنی اش رند نشده است؟

(I) (4 امتیاز) توضیح دهید که چطور شاخص ثبات حساب می‌شود.

(J) (5 امتیاز) توضیح دهید که چطور فاصله آمارواژگان حساب می‌شود.

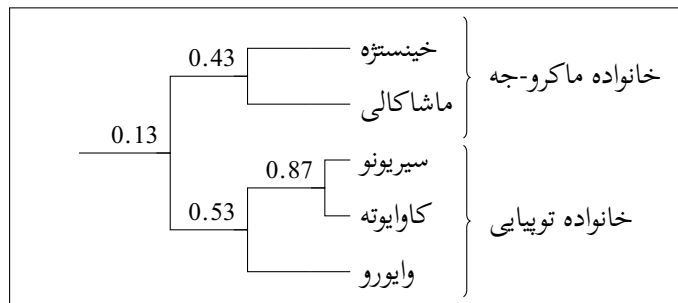
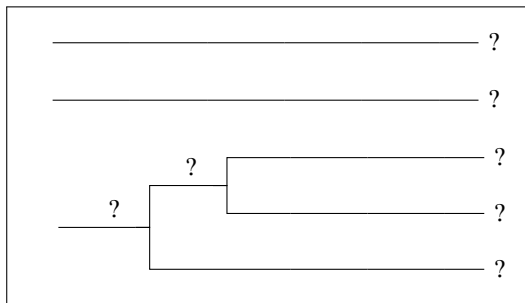
(K) (4 امتیاز) تفاوت میان الگوریتم آ و ب را توضیح دهید.

بخش ۶. خانواده ماکرو-جه و خانواده توپیایی (برزیل، بولیوی)

(L) (28 امتیاز) خانواده ماکرو-جه و خانواده توپیایی خانواده بزرگ زبان‌های جنوب آمریکا هستند. برخی زبان‌شناسان معتقدند که آنها دوردور مرتبط هستند. فهرست‌های بعدی را بررسی کنید.

E	Δ	Γ	B	A	
= 'pɛ	mĩβ̃m= 'tɛaj	kɥp= 'pɛ	hʷĩ= 'kʰʌ	e= 'e-ke	پوست درخت
=rɛ'wɛk	= 'tɛj	=ã'ũn	= 'tʰigi	'e=rje	دل
=ru'i	= 'hɛβ̃p	=d̃z= a'u	=ka'nbɾo	e= 'ruki	خون
=ra'pi	mũ=...= 'haβ̃p	=po'kʷa	=rɔ= 'kʰɔ̃	'e=raĩ	سوزاندن
= 'kap	= 'tuβ̃p	= 'd̃z= ap	= 'tʰwəmi	e= 'kira	چربی
= 'pi	=po'ta	= 'βi	= 'hʷaji	'e=i	پا
= 'pɔ	= 'nĩβ̃m	= 'βo	=nĩ'kʰɔ̃	'e=o	دست
=po'ij	=β̃p'tɔj	=po'ti	=wi'tʰĩ	e= 'usi	سنگین
=pi'ʔa	=tɛiβ̃pkĩ'nāj	=pi'a	= 'nba	'e=ja	جگر
=pia'u	= 'tiβ̃p	=pa'gop	= 'ndiwi	e= 'jasu	جدید
=ra'pɔ	mĩβ̃m=nĩβ̃m=tɛa'tiə	kɥp=kɥjo'pɛ	=ja'ɾe	e= 'rao	ریشه
= 'pit	= 'tɛaj	= 'pɛ	= 'kʰʌ	'e=i	پوست
= 'raj	=nā:= 'kiβ̃p	=d̃z= o'kʷaj	= 'nbi	e= 'rokoĩ	دم
= 'sĩŋ	=β̃p'douɥ	=d̃zi'ra	=ja'kʰa	'e=ʃĩ	سفید
=pe'pɔ, =ji'wa	=nĩ'māuɥ	=pe'o	=ja'ɾa	e= 'heo	بال

ذیل دو درخت داده شده که بر اساس همان لیست‌ها بنا شده است. برخی داده‌ها (اسامی زبان‌ها و فاصله آمارواژگان) جا افتاده است. جای خالی را پر کنید. برای هر درخت ذکر کنید که نشانه گذاری به دست کارشناسان یا اتوماتیک بنا شده است و همچنین آیا به کمک الگوریتم آیا ب ساخته شده است.



E	Δ	Γ	B	A
?	?	?	?	?

▲ نشانه گذاری دستی و شاخصه‌های ثابت در این تکلیف عمدا داده نشده‌اند.

(M) (10 امتیاز) پروسه‌های خودکار که براساس کلاس‌های دولگوپولسکی بنا شده‌اند می‌توانند نتایج اشتباه بدهند. در مثال بررسی شده پروسه خودکار بیشتر بین سیرینو و یک زبان ماکرو-جه خاص (خینستزه) شباهت تشخیص می‌دهد تا بین سیرینو و سایر زبان‌های توپایی. گزینه پروسه خودکار مد نظر خود را پیشنهاد دهید که می‌توانست در رابطه با مطالب زبان‌های داده شده ماکرو-جه و توپایی کارآمد باشد و کوتاه آن را توضیح دهید.

این تکلیف فقط در شرایط برابر بین تیم‌ها با امتیاز بالا بررسی خواهد شد. \triangle

طراحان مسأله از الخاندرو ویدال، ماریا کونوشنکو، ایلیا گرونوف و یامتو سویا بابت مشورت در زبان‌های خاص تشکر می‌کنند.
—آندری نیکولین، میلنا ونوا

ویراستاران: بوریس ایومدین، هیو دابز، ایوان درژانسکی (ویراستار تخصصی)، ماریا روبینشتاین، لیام مک نایت، آندری نیکولین (سر ویراستار)، الیسیا وارنر، میلنا ونوا، یان پتر، آکسی پگوشف، الکساندر پیپرسکی، استانیسلاو گورویچ.

متن فارسی: بهناز مسلمیان.

موفق باشید!