

## Kahekümne esimene rahvusvaheline lingvistikaolümpiaad

Brasília (Brasília), 23.–31. juuli 2024

## Meeskonnavõistluse ülesanne

Leksikostatistika on grupp meetodeid, mida kasutatakse selleks, et hinnata, kui lähedased on keeled üksteisele sõnavara poolest. Neid meetodeid kohaldatakse tavaliselt pikkadele sõnaloeteludele, mida spetsialistid käsitsi kommenteerivad, et näidata, kas antud sõnapaar on eeldatavalt samast allikast pärit. Mõnikord kasutavad keeleteadlased siiski leksikostatistilisi meetodeid automaatselt märgendatud sõnaloeteludele. Üks selline meetod põhineb *konsonandiklasside* mõistel, mille nõukogude-iisraeli keeleteadlane Aharon Dolgopolski võttis kasutusele aastal 1964.

P.	p b ɸ β f v	K.	k g x ɣ q ɕ χ ɰ	Y.	j ç (tüve alguses)	M.	m ɱ
T.	t d θ ð t̪ d̪	R.	r r̥ ɽ ɻ ʁ ʂ ʃ ʣ	W.	w ɱ (tüve alguses)	N.	n ɲ ɳ ɰ
S.	s z ʃ ʒ ʂ ʣ ʒ ʣ					Q.	ʈ ɖ ʈ
H.	h ʕ ɦ ʕ ʔ h ɦ ʔ, vokaalid ja j ç w ɱ (välja arvatud tüve alguses)						

## Dolgopolski konsonandiklassid

All leiate märkustega varustatud sõnaloetelude väljavõtted maailma erinevatest keelkondadest. Märkused on esitatud alaindeksitega. Nende loetelude põhjal on konstrueeritud keelepuud, kasutades nn *StarlingN7* algoritmi kahte lihtsustatud meetodit, ning igale sõnale on määratud *stabiilsusindeks*. Ülemises osas esitatud puud ja stabiilsusindeksid põhinevad käsitsi märgendatud sõnaloeteludel ja alumises osas esitatud puud ja stabiilsusindeksid põhinevad automaatselt märgendatud loeteludel. Iga sõnaloetelu jaoks on olemas kaks konstrueeritud puud, vastavalt algoritmi mõlemale versioonile: Algoritm A ja Algoritm B. Pange tähele, et mõnikord on mitu võimalikku puud, mis vastavad ühele sõnaloetelule; sellisel juhul on üks puu valitud juhuslikult. Iga sõlmpunktile on määratud leksikostatistiline kaugus. Mida suurem on kaugus, seda lähedasemalt on keeled omavahel suguluses. Täpsem mõiste oleks siis „vastupidine leksikostatistiline kaugus“, mitte „leksikostatistiline kaugus“. Et asja lihtsamaks teha, kasutame siin ülesandes mõistet „leksikostatistiline kaugus“.

Nii stabiilsusindeksid kui ka leksikostatistilised kaugused on ümardatud kahe kümnendkohani. Kui kolmas number pärast koma on väiksem kui 5, ümardatakse allapoole; vastasel juhul ümardatakse ülespoole. Näiteks on 2,836 ümardatud 2,84 peale, 0,705 0,71 peale ja 0,703 0,70 peale. Ümardamine kehtib ainult nende väärtuste kohta, mida näidatakse inimestest lugejatele. Seega arvuti „näeb“ algoritmide rakendamise käigus ümardamata väärtusi.

Pange tähele, et mõned sõnad on teadaolevalt või arvatavasti laenatud teistest keeltest. Näiteks kadiveu keele **jok:i** ‘sool’ sõna on laenatud guaranikeelsest **juki** ning ’iipay (Mesa Grande murre) keele **?a:nj** ‘aasta’ sõna on laenatud hispaaniakeelsest **’ano**.

Mõnel juhul on sõnade loeteludes ühe tähenduse jaoks esitatud mitu sünonüümi, mis on üksteisest eraldatud komaga. Üks näide on ‘jalg’ vehhose keeles.

Allpool esitatud andmetes on kõik eesliited eraldatud märgiga „=“ ja kõik järelliited märgiga „-“. Mõnda sõna kasutatakse ainult eesliitega koos. Need algavad märgiga „=“.

Andmed on transkribeeritud rahvusvahelise foneetilise tähestiku (IPA) abil. <sup>ˈ</sup> = pearõhk, <sub>ˌ</sub> = kaasarõhk (nõrgem kui pearõhk), ː = pikk häälik, ˘ = ülilühike häälik,  $\widehat{XY}$  = X ja Y hääldatakse ühe häälikuna, ˊ = kõrge toon, ˋ = madal toon, ˆ = langev toon, ˚ = preglotaliseeritud häälik (eelneb lühike õhuvoolu katkestus kurgus), ˗ = ejetiiv (konsonant, mille hääldamisel kurk korraks sulgub), ˚ = helitu häälik, ˚ = nasaliseerunud häälik (nina kaudu hääldatav häälik), ˚ = kärisev häälik (madala, ragiseva kõlaga),

$^n\circ$  tähendab, et enne kaashäälikut läheb osa õhust läbi nina,  $\circ^h$  = aspireeritud konsonant (hääldusega kaasneb lisahõngus),  $\circ^w$  = labialiseeritud kaashäälik (hääldatakse ümarate huultega),  $\circ^j$  = palataliseeritud (pehmendatud) häälik (hääldamisel liigub keel kõva suulae poole). **ɑ, æ, ɛ, ɪ, ɔ, ʊ, ɘ, ə, ʌ, ɒ, ɔ, ʊ, ɐ, ø** on täishäälikud. Muud erimärgid tähistavad kaashäälikuid.

⚠ Mõne ülesandes kasutatud keele oskamine ei anna ülesande lahendamiseks mingeid eeliseid.

### I osa. Guaikuru keelkond (Argentina, Brasiilia, Paraguay)

	toba (idamurre)	pilaga	mokovi (chaco murre)	kadiveu
pilv	l=ʔok <sub>1</sub>	'lo=ʔok <sub>1</sub>	naweyelek <sub>2</sub>	lol:adi <sub>3</sub>
tuli	nodek <sub>1</sub>	'd=oleʔ <sub>2</sub>	norek <sub>1</sub>	n=ol:edi <sub>2</sub>
kala	njaq <sub>1</sub>	'nijaq <sub>1</sub>	naʎin <sub>2</sub>	nij:ogo-ḍʒegi <sub>3</sub>
pea	=qajk <sub>1</sub>	'qajk <sub>1</sub>	=qaik <sub>1</sub>	=ak:ilo <sub>2</sub>
tapma	=alawat <sub>1</sub>	=a'la:t <sub>1</sub>	=alawat <sub>1</sub>	=el:owadi <sub>1</sub>
kuu (taevakeha)	ʔawoʂojk <sub>1</sub>	ʔa'woʂojk <sub>1</sub>	ʃirajyo <sub>2</sub>	ep:enaj <sub>3</sub>
nina	=mik <sub>1</sub>	'mik <sub>1</sub>	=mik <sub>1</sub>	=m:iq:o <sub>1</sub>
sool	towe <sub>1</sub>	ol'yek <sub>2</sub>	ʔwe <sub>1</sub>	jok:i <sub>1</sub>
kivi	qaʔ <sub>1</sub>	'qaʔ <sub>1</sub>	qaʔ <sub>1</sub>	wet:iga <sub>2</sub>
keel	=atʃ-aʂat <sub>1</sub>	=a'tʃ-aʂat <sub>1</sub>	=oʔley-aʂan-aʂat <sub>2</sub>	=ok:el:i <sub>3</sub>

	A-algoritm	B-algoritm	
käsitsi	<p>leksikostatistiline kaugus</p>		Stabiilsusindeksid: pilv 0,50 tuli 0,50 kala 0,50 pea 0,75 tapma 1,00 kuu (taevakeha) 0,50 nina 1,00 sool 0,67 kivi 0,75 keel 0,50
automatiseeritud			Stabiilsusindeksid: pilv 0,50 tuli 0,50 kala 0,75 pea 0,75 tapma 1,00 kuu (taevakeha) 0,50 nina 1,00 sool 0,25 kivi 0,75 keel 0,50

## II osa. Nuubia keelkond (Egiptus, Sudaan)

	dongolawi	kenuzi	dilingi	kadaru	debri	birgidi
tapma	'bɛ:₁	be:₁	hur₂	wur-i₂	wur-i₂	fila:l-e₁
kuu (taevakeha)	u'n-at-t₁	an-at-ti₁	nɔn-ti₁	nɔn-tu₁	nɔn-to₁	ma:l₂
vesi	'ɛss₁	essi₁	ɔti₁	ɔto₁	ɔtu₁	ɛji₁
andma	'tir₁	tir₁	ti₁	ti₁	ti₁	te:n₁
hea	'sɛrɛ:₁	sere:₁	ken₂	kɛn₂	kɛŋ₂	azze-n₃
tuul	'turug₁	turug₁	irf-i₂	irf-o₂	irf-o₂	kurr-i₃
juuksed	'dil-ti₁	si:r₂	tel-ti₁	til-tu₁	til-tu₁	ur=dill-e₁
kõht	'tu:₁	tu:₁	te-te₂	to₁	to₁	tu:₁
magama	'nɛ:r₁	ne:r₁	ɟer₁	dwallɛli₂	ɟer-i₁	ne:r-i₁
päike	'masil₁	masil₁	ɛɟ₂	aju₂	ɛŋgal-to₃	ʔi:zi₂

	A-algoritm	B-algoritm	
käsitsi			tapma 0,50 kuu (taevakeha) 0,83 vesi 1,00 andma 1,00 hea 0,50 tuul 0,50 juuksed 0,83 kõht 0,83 magama 0,83 päike 0,50
automatiseeritud			tapma 0,33 kuu (taevakeha) 0,50 vesi 0,50 andma 0,67 hea 0,50 tuul 0,50 juuksed 0,83 kõht 1,00 magama 0,50 päike 0,50

- (A) (2 punkti) Kaashäälik **ɟ** hääldub nagu prantsuse *r* ehk keele tagaosas. Millisesse Dolgopolski klassi see kuulub ja kuidas te teada saite?
- (B) (2 punkti) Nuubia keelte puu üleval vasakul on ainult üks kahest võimalikust puust selle algoritmi ja märkustetüübi kombinatsiooni puhul. Joonistage teine võimalik puu.
- (C) (2 punkti) Nuubia keelte puu all vasakul on ainult üks kahest võimalikust puust selle algoritmi ja märkustetüübi kombinatsiooni puhul. Joonistage teine võimalik puu.
- (D) (2 punkti) Leksikostatistiline kaugus 0,49 (mis on määratud Nubia puu juurele üleval paremal) on ümardatud kahe kümnendkohani, nagu ka mõned teised selle ülesande kaugused. Mis on täpne kaugus?

III osa. Mataguai keelkond (Argentina, Boliivia, Paraguay)

	vitši (Bermejo alamjooksu murre)	vitši (Rivadavia murre)	vehhose	'weenhayeki	iyojwa'aja'	manjui	nivakle (shichaam lhavos)	nivakle (chisham-nee lhavos)	maka
tuli	ʔitox <sub>1</sub>	ʔitox <sub>1</sub>	ʔitah <sub>1</sub>	ʔi:tax <sub>1</sub>	'hwat <sub>2</sub>	'ʔeite <sub>1</sub>	ʔitax <sub>1</sub>	ʔitax <sub>1</sub>	feʔt <sub>2</sub>
kala	ʔwahat <sub>1</sub>	wahat <sub>1</sub>	wahat <sub>1</sub>	ʔwa:hat <sub>1</sub>	si'ʔjus <sub>-1</sub>	ʃi'ʔjus <sub>-1</sub>	saxetʃ <sub>-1</sub>	saxetʃ <sub>-1</sub>	sehets <sub>-1</sub>
jalg	=patʃ <sub>u</sub> <sub>1</sub>	=qol <sub>2</sub>	=patʃ <sub>o</sub> <sub>1</sub> , =kala <sub>2</sub>	=pa:k'oʔ <sub>1</sub>	=sat <sub>3</sub>	=ka'laʔ <sub>2</sub>	=φoʔ <sub>4</sub>	=φoʔ <sub>4</sub>	=f'iʔ <sub>5</sub>
vesi	ʔinot <sub>1</sub>	ʔinot <sub>1</sub>	wah <sub>2</sub>	ʔina:t <sub>1</sub>	ʔi'n'at <sub>1</sub>	ʔa'ʔnat <sub>1</sub>	jinaʔt <sub>1</sub>	jinaʔt <sub>1</sub>	iweliʔ <sub>3</sub>
andma	=ʔweŋ-u <sub>1</sub>	=weŋ-u <sub>1</sub>	=ʔweŋ-o <sub>1</sub>	=ʔweŋ-oʔ <sub>1</sub>	=wɛhn-aʔm <sub>2</sub>	=hajʔ <sub>3</sub> , =wɛn <sub>2</sub>	=xut <sub>4</sub>	=xut-ej <sub>4</sub>	tis-ix <sub>5</sub>
hea	ʔis <sub>1</sub>	ʔis <sub>1</sub>	ʔis <sub>1</sub>	ʔis <sub>1</sub>	'ʔes <sub>1</sub>	'ʔeis <sub>1</sub>	ʔis <sub>1</sub>	ʔis <sub>1</sub>	t=ejk'un-ej <sub>2</sub>
tuul	ʔinwok <sup>w</sup> <sub>1</sub>	ʔinwək <sub>1</sub>	ʔihwok <sup>w</sup> <sub>1</sub>	=ja:ʔ <sub>2</sub> , =x <sup>w</sup> ox <sup>w</sup> <sub>3</sub>	'hlahwuʔ <sub>4</sub>	'hlahwu <sub>4</sub>	ʔaβiʔm <sub>5</sub>	ʔaβiʔm <sub>5</sub>	t'unik'i <sub>6</sub>
puu	haʔlo <sub>1</sub>	hal <sub>1</sub>	haʔla <sub>1</sub>	haʔlaʔ <sub>1</sub>	ʔa'ʔlaʔ <sub>1</sub>	ʔa'ʔla-k <sub>1</sub>	ʔaʔkxi-juk <sub>2</sub>	jiʔklaʔ <sub>1</sub>	naxka-k <sub>3</sub>
juuksed	=ʔwule-j <sub>1</sub>	=wule-j <sub>1</sub>	=ʔwole-j <sub>1</sub>	=ʔwo:le-ç <sub>1</sub> , hi:lənax <sub>2</sub>	=ʔwole <sub>1</sub>	=ʔwole-j <sub>1</sub>	=ʃateʔtj <sub>3</sub>	=jeʔs <sub>4</sub>	=ʔewkux-its <sub>5</sub>
tapma	=lon <sub>1</sub>	=lən <sub>1</sub>	=lan <sub>1</sub>	=la:ŋ <sub>1</sub>	=laʔan <sub>1</sub>	=lan <sub>1</sub>	=klan <sub>1</sub>	=klan <sub>1</sub>	=lan <sub>1</sub>

	A-algoritm	B-algoritm	
käsitsi			Stabiilsusindeksid: tuli 0,78 kala 1,00 jalg 0,33 vesi 0,78 andma 0,44 hea 0,89 tuul 0,33 puu 0,78 juuksed 0,67 tapma 1,00
automatiseeritud			Stabiilsusindeksid: tuli 0,78 kala 0,44 jalg 0,33 vesi 0,56 andma 0,67 hea 0,89 tuul 0,22 puu 0,67 juuksed 0,67 tapma 1,00

**IV osa. Mongoli keelkond** (Hiina Rahvavabariik, Mongoolia, Venemaa)

(E) (10 punkti) Vaadake järgnevat sõnaloetelu. Arvutage stabiilsusindeks, nii käsitsi tehtud kui ka automatiseeritud märkuste jaoks.

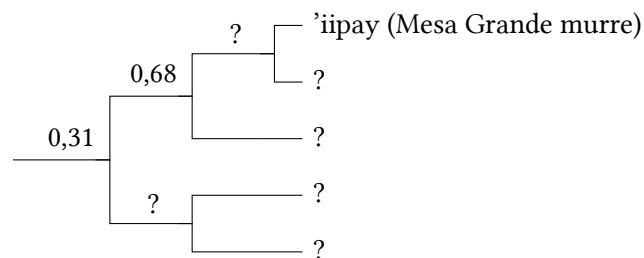
Et teid pisut aidata, oleme sõna 'kõik' jaoks mõlemad stabiilsusindeksid juba välja arvutanud. Need on – suvalises järjekorras – 0,36 ja 0,40.

	daguuri (hailari muure)	hamnigani (mandžuu- ria murre)	burjaadi (hori murre)	uusbarga	ööldi	hošuudi	kalmõki	halha	ordose	šira- juguri	bonani
kõik	hɔ:₁	bʊlt₂	bʊxi:₃	bygd₄	tsug₅	lug₅	tsuk₅, xamak₋₁	pux₃, pugt₄, xamǎg₋₁	pyyyte₄, xamukʰ₋₁	tʃʰuq₅	hanə₂
puukoor	hails₁	qalihon₁	χoltɔhɔn₂	xalʰhu:₁	xolts₂	xalis₁	dursn₃	xɔɣtʰɔs₂	turusu₃	χalsən₁	arasun₄
kõht	ke:li₁	gətəhən₂	gedehen₂	gedy:₂	ge:s₂	gets₂	gesn₂	gitis₂, xiwɣij₁	ketysy₂	ketesən₂	kele₁
lind	dəgi₁	eiwan₁	ʃubu:n₁	ʃuwu:₁	ʃuvu:₁	ʃuwu:₁	ʃowun₁	ʃuwu₁	ʃuβu:₁	ʃu:n₁, peltʃɔr₂	bendzer₂
tuli	gali₁	gal₁	gal₁	gal₁	gal₁	gal₁	gal₁	gal₁	qal₁	qal₁	χal₁
rada	terg-u:l₁	qargʊi₂	χargi₂, zam₁	zam₁	dzam₁	dzam₁	xa:-lɔə₃	tsam₁	tʃam₁	mør₄	mor₄
sool	hata:₁	dawhɔn₂	dabhan₂	dawuhu:₂	daws₂	daws₂	dawsn₂	tawsǎ₂	taβusu₂	ta:psən₂	dabsuŋ₂
ujuma	unpa-du₁	ɔmba₁	tʰamar₂	umb₁	sele₋₃	umba₁	us-tɛi₋₄, ɔ:m₋₅	siɣi₋₃	usu-tʃʰi-la₋₄	umpa₁	mba₁
vesi	ɔsɔ₁	ɔxɔn₁	uhan₁	u:ha₁	usn₁	us₁	usn₁	ʊšɔ₁	usun₁	qʰusun₁	sə₁
tuul	kein₁	halkin₂	halxin₂	halxi₂	salʰxin₂	salkʲi₂	salʲkn₂	saɣxi₂	kʰi:₁	kʰi:₁	ki₁

**V osa. Juma keelkond** (Mehhiko, Ameerika Ühendriigid)

(F) (8 punkti) Vaadake järgnevat sõnaloetelu. All näete puud, mis konstrueeriti sama loetelu alusel. Mõningad andmed (keelte nimed ja leksikostatistilised kaugused) on puudu. Täitke lüngad. Pange kirja, kas puu on käsitsi tehtud või automaatne, samuti, kas see on tehtud algoritmi A või B järgi.

	mohave	kokopa	javapai	tiipai (jamuli)	'iipay (Mesa Grande murre)
lühike	wena=wen-a <sub>1</sub>	'xʌ=ʔut <sub>2</sub>	'tʃkr=ot-i <sub>2</sub>	lə=ʔuŋ <sub>1</sub>	mə=put-k <sub>3</sub>
lind	ʔitʃ=i=jer <sub>1</sub>	'ʃa <sub>2</sub>	'ʔ=tʃ=sa <sub>2</sub>	aʔ='ʃa <sub>2</sub>	ʔa:=ʃa:2
luu	ŋ=a=s=ak <sub>1</sub>	'ŋ=j=a:k <sub>1</sub>	'tʃ=j=a:k-a <sub>1</sub>	'ak <sub>1</sub>	aq <sub>1</sub>
kuiiv	i=ro:-v-k <sub>1</sub>	'ʃ=ʔar <sub>2</sub>	'ru-β-i <sub>1</sub>	's=ʔa:j <sub>3</sub>	sa:j <sub>3</sub>
liha	k <sup>w</sup> i:k <sup>w</sup> ay <sub>1</sub>	ʔi='ma:tʃ <sub>2</sub>	'k <sup>w</sup> e:='θo-β-a <sub>3</sub>	'k <sup>w</sup> ak <sub>4</sub>	kuk <sup>w</sup> a:j-p <sub>1</sub>
kael	maʎaqe <sub>1</sub>	'm=puk <sub>2</sub>	'mlq <sub>1</sub>	i:='puk <sub>2</sub>	i:=puk <sub>2</sub>
nägema	i=ju:-k <sub>1</sub>	'wi:2	'ʔu:1	'wi:w <sub>2</sub>	ə=wu:w <sub>2</sub>
saba	i:=ʔar <sub>1</sub>	'ʃ=juʎ <sub>2</sub>	'β=hé <sub>3</sub>	ʃə='juʎ <sub>2</sub>	xə=juʎ <sub>2</sub>
kaks	havik-k <sub>1</sub>	'x=wak <sub>1</sub>	'h <sup>w</sup> âk-i <sub>1</sub>	xə='wak <sub>1</sub>	xə=wak <sub>1</sub>
aasta	hu:ðe <sub>1</sub>	'mat-'ka:m <sub>2</sub>	'ʔ=tʃ <sup>n</sup> ur-a <sub>3</sub>	mat-'wam <sub>2</sub>	ʔa:n <sup>j</sup> .1



(G) (20 punkti) Juma keelte jaoks on loodud veel mõned puud, kus puude juurte vahel on järgmised leksikostatistilised kaugused (iga puu vasakpoolses otsas):

1. 0,20
2. 0,23
3. 0,24

Joonistage neist igaühe kohta puu. Pange iga puu kohta kirja, kas see on käsitsi tehtud või automaatne, samuti, kas see on tehtud algoritmi A või B järgi.

(H) (3 punkti) Kaks osaülesandes (G) esitatud kaugust on ümardatud kahe kümnendkohani: 0,23 on ümardatud 0,225 pealt. Milline teine kaugus on ümardatud ning mis on selle täpne väärtus?

(I) (4 punkti) Selgitage, kuidas arvutatakse stabiilsusindeksid.

(J) (5 punkti) Selgitage, kuidas arvutatakse leksikostatistilised kaugused.

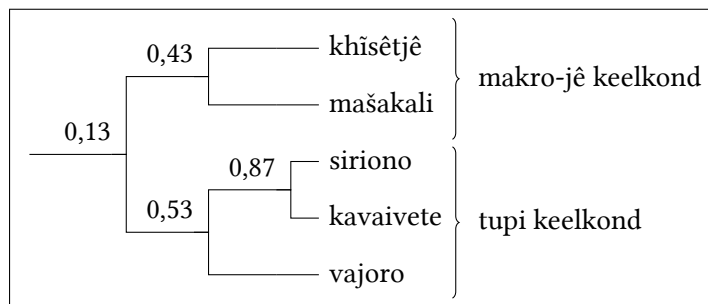
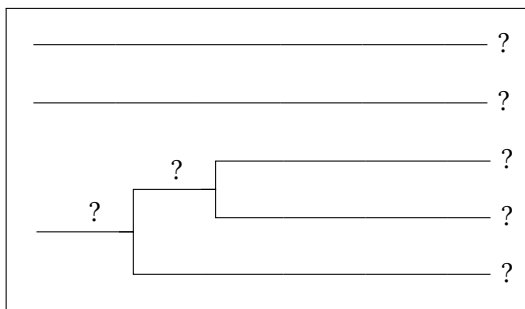
(K) (4 punkti) Selgitage A- ja B-algoritmide vahelist erinevust.

**VI osa. Makro-jê keelkond ja tupi keelkond (Brasiilia, Boliivia)**

(L) (28 punkti) Makro-jê ja tupi on kaks suuremat keelkonda Lõuna-Ameerikas. Mõned keeleteadlased usuvad, et nad on omavahel kaugemalt suguluses. Vaadake järgnevaid sõnaloetelusi.

	A	B	Γ	Δ	E
puukoor	e='e-ke	h <sup>w</sup> ĩ='k <sup>h</sup> λ	kup='pe	mĩβ̃m='tɛaj	'pe
kõht	'e=rje	't <sup>h</sup> igi	=ã'ün	'tæj	=rɛ'wek
veri	e='ruki	=ka <sup>n</sup> bɔ	=d̃z=a'u	'heβ̃p	=ru'i
põletama	'raĩ	=rɔ='k <sup>h</sup> λɔ	=po'k <sup>w</sup> a	mũ=...'haβ̃p	=ra'pi
rasv	e='kira	't <sup>h</sup> wəmi	'd̃z=ap	'tuβ̃p	'kap
jalg	'e=i	'h <sup>w</sup> aji	'βi	=pɔ'ta	'pi
käsi	'e=o	=ɲĩ'k <sup>h</sup> λa	'βo	'ɲĩβ̃m	'pɔ
raske	e='usi	=wi't <sup>h</sup> ĩ	=po'ti	=β̃p'tɛj	=pɔ'ij
maks	'e=ja	' <sup>n</sup> ba	=pi'a	=tɛiβ̃pkĩ'nãj	=pi'ʔa
uus	e='jasu	' <sup>n</sup> diwi	=pa'gop	'tiβ̃p	=pia'u
juur	e='rao	=ja'ɕe	kup=kujɔ'pe	mĩβ̃m=ɲĩβ̃m='tɛa'tiɔ	=ra'pɔ
nahk	'e=i	'k <sup>h</sup> λ	'pe	'tɛaj	'pit
saba	e='rokoĩ	' <sup>n</sup> bi	=d̃z=o'k <sup>w</sup> aj	=nã:'kiβ̃p	'raj
valge	'e=fi	=ja'k <sup>h</sup> a	=d̃zi'ra	=β̃p'douɥ	'sĩɲ
tiib	e='heo	=ja'ɕa	=pe'o	=ɲĩ'mãuɥ	=pe'pɔ, =ji'wa

All näete kahte puud, mis konstrueeriti samade loetelude alusel. Mõningad andmed (keelte nimed ja leksikostatistilised kaugused) on puudu. Täitke lüngad. Pange iga puu kohta kirja, kas see on käsitsi tehtud või automaatne, samuti, kas see on tehtud algoritmi A või B järgi.



A	B	Γ	Δ	E
?	?	?	?	?

⚠ Käsitsi tehtud märkused ning stabiilsusindeksid on siin osaülesandes meelega ära jäetud.

(M) (10 punkti) Dolgopolski klassidel põhinevad automaatsed protsessid võivad viia valede tulemusteni. Selles näites tuvastab automaatne protsess rohkem sarnasusi siriono ja ühe teatud makro-jê keele (khîsêjtê keele) vahel kui siriono ja teiste tupi keelte vahel. Tehke ettepanek modifitseeritud automaatse protsessi kohta, mis annaks makro-jê ja tupi sõnastike puhul õige klassifikatsiooni, ja kirjeldage seda *lühidalt*.

⚠ Seda osaülesannet hinnatakse ainult sel juhul, kui parimate tulemustega võistkonnad lõpetavad viigiga.

Autorid tänavad Alejandro Vidali, Maria Konošenkot, Ilja Gruntovit ja Jamthô Suyat, kes lahkelt vastasid küsimustele üksikute keelte kohta. —Andrei Nikulin, Milena Veneva

---

**Toimetajad:** Ivan Deržanski (tehn. toim.), Hugh Dobbs, Stanislav Gurevitš, Boriss Iomdin, Eimear McKnight, Andrei Nikulin (vast. toim.), Aleksejs Peguševs, Jan Petr, Aleksandr Piperski, Marija Rubinštein, Milena Veneva, Elysia Warner.

**Eesti tekst:** Axel Jagau.

Edu!