pt(B)

Vigésima primeira Olimpíada Internacional de Linguística

Brasília – DF (Brasil), 23–31 de julho de 2024

Problema da prova por equipes

A lexicoestatística é um grupo de métodos criados para avaliar quão distante é a relação entre um conjunto de línguas com base em seu vocabulário. Normalmente tais métodos são aplicados a longas listas de palavras anotadas manualmente por especialistas, que indicam se, de acordo com seu juízo, um determinado par de palavras provém de uma mesma fonte. Porém às vezes linguistas aplicam métodos lexicoestatísticos a listas de palavras anotadas por meio de procedimentos automatizados. Um desses procedimentos baseia-se no conceito de *classes consonantais*, introduzido pelo linguista soviético e israelense Aharon Dolgopolsky em 1964.

P.	p b ɓ ɸ β f v	K.	k g x ɣ q ɢ χ ʁ	Y.	j ç (no início da raiz)	M.	m ɱ
T.	t d ɗ θ ð ʈ ɖ	R.	r ɾ ʀ ɹ l ɬ ɮ ɭ ʎ ɬ	W.	w ʍ (no início da raiz)	N.	n ɳ ɲ ŋ
S.	s z ʃ ʒ ʂ ʐ ɕ ʑ c ɟ					Q.	t͡ɬ d͡ɮ
H.	ħ ʕ ʜ ʢ ʡ h ɦ ʔ, vogais e j ç w ʍ (exceto no início da raiz)						

Classes consonantais de Dolgopolsky

Abaixo vocês encontrarão fragmentos de listas de palavras de várias famílias linguísticas do mundo. As anotações são representadas por dígitos subscritos. Com base nessas listas, foram construídas árvores das respectivas famílias linguísticas por meio de duas versões simplificadas do chamado algoritmo *StarlingNJ*, e foi atribuído um *índice de estabilidade* a cada palavra. As árvores e os índices de estabilidade na parte de cima baseiam-se em listas anotadas manualmente, enquanto as árvores e os índices de estabilidade na parte de baixo se baseiam em listas anotadas automaticamente. Em cada caso, há duas árvores para cada lista, construídas com versões diferentes do algoritmo: o algoritmo A e o algoritmo B. Notem que em alguns casos há múltiplas árvores possíveis que correspondem a uma lista de palavras; em tais casos, uma árvore foi escolhida aleatoriamente. Uma distância lexicoestatística é atribuída a cada nó em cada árvore. Quanto maior a distância, mais próxima a relação entre as línguas. Deste modo, o termo mais exato seria "distância lexicoestatística invertida" e não "distância lexicoestatística". Por uma questão de simplicidade, neste problema fazemos uso do termo "distância lexicoestatística".

Tanto os índices de estabilidade como as distâncias lexicoestatísticas foram arredondadas para duas casas decimais. Caso a terceira casa seja menor que 5, o valor é arredondado para baixo; caso contrário, é arredondado para cima. Por exemplo, 2,836 arredonda-se para 2,84, 0,705 para 0,71, 0,703 para 0,70. O arredondamento é aplicado unicamente aos valores mostrados a leitores humanos. Em outras palavras, o computador que roda os algoritmos "vê" os valores não arredondados.

Notem que algumas palavras são empréstimos conhecidos ou prováveis de outras línguas. Por exemplo, a palavra **jokːi** 'sal' da língua kadiwéu é um empréstimo do guarani (**jukɨ**), ao passo que a palavra **ʔaːnʲ** 'ano' do 'iipay (Mesa Grande) é um empréstimo do espanhol (**ˈaɲo**).

Às vezes são listados múltiplos sinônimos que correspondem a um único conceito, separados por uma vírgula. Um exemplo é o conceito 'pé' na língua vejoz.

Nos dados abaixo, todos os prefixos são separados pelo símbolo "=", ao passo que todos os prefixos são separados pelo símbolo "-". Algumas palavras jamais ocorrem sem prefixos. Estas vêm precedidas pelo símbolo "=".

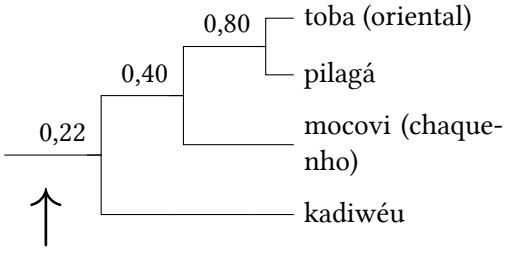
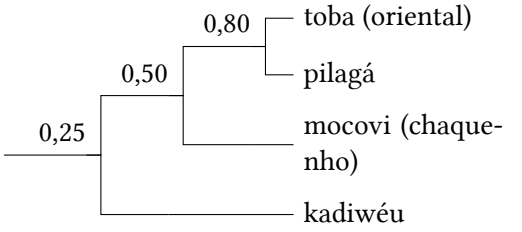
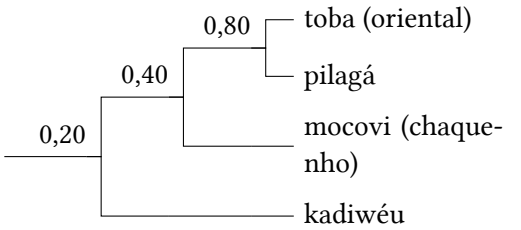
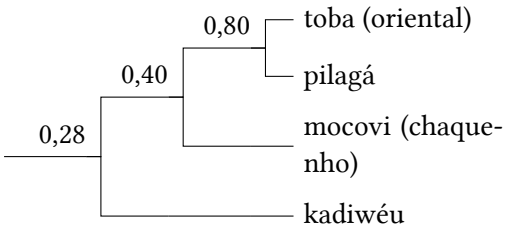
A transcrição dos dados segue o Alfabeto Fonético Internacional. ˈ = acento primário, ˌ = acento secundário (mais fraco que o primário), ◌ː = som longo, ◌̆ = som muito curto, X͡Y = X e Y são pronunciados

como um único som, ◌́ = tom alto, ◌̀ = tom baixo, ◌̂ = tom descendente, ◌ˀ = consoante preglotalizada (precedida por um curto bloqueio do fluxo de ar na glote), ◌’ = consoante ejetiva (pronunciada com um curto bloqueio do fluxo de ar na glote), ◌̥ = som desvozeado, ◌̃ = som nasalizado (pronunciado pelo nariz), ◌̰ = laringalização (som rouco, crepitante), ◌ⁿ assinala um fluxo de ar através do nariz que precede ao som da consoante, ◌ʰ = consoante aspirada (pronunciada com um sopro), ◌ʷ = consoante labializada (pronunciada com os lábios arredondados), ◌ʲ = consoante palatalizada (pronunciada com uma parte da língua aproximando-se ao palato duro). **ɑ**, **æ**, **ɛ**, **ɪ**, **ɨ**, **ɔ**, **ʊ**, **ʉ**, **ə**, **ʌ**, **ɒ**, **ɘ**, **ʏ**, **ɵ**, **ø** são vogais. Outros caracteres especiais denotam consoantes.

⚠ O conhecimento de qualquer uma das línguas mencionadas neste problema não dá vantagem para sua solução.

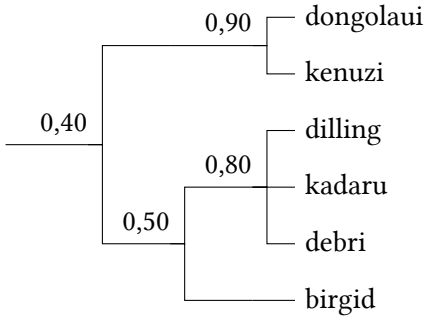
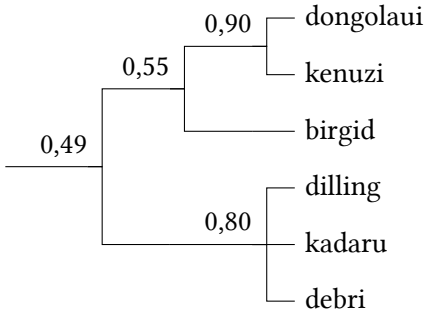
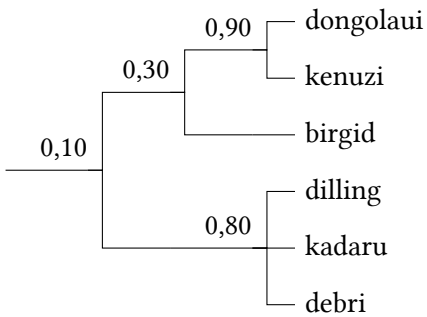
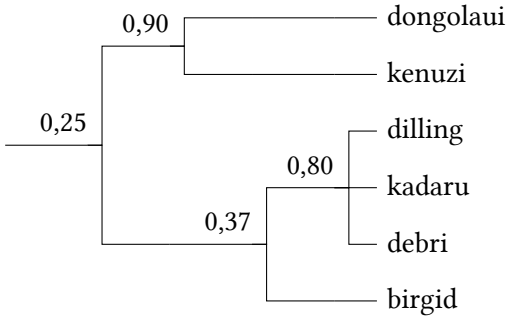
Parte I. Família guaicuru (Argentina, Brasil, Paraguai)

	toba (oriental)	pilagá	mocovi (chaquenho)	kadiwéu
nuvem	l=ʔok$_1$	ˈlo=ʔok$_1$	naweɣelek$_2$	lol:adi$_3$
fogo	nodek$_1$	ˈd=oleʔ$_2$	noɾek$_1$	n=ol:edi$_2$
peixe	njaq$_1$	ˈnijaq$_1$	naʌin$_2$	nij:oɢo-d͡ʒegi$_3$
cabeça	=qajk$_1$	=ˈqajk$_1$	=qaik$_1$	=ak:ilo$_2$
matar	=alawat$_1$	=aˈla:t$_1$	=alawat$_1$	=el:owadi$_1$
lua	ʔawoʁojk$_1$	ʔaˈwoʃojk$_1$	ʃɪrajɣo$_2$	ep:enaj$_3$
nariz	=mik$_1$	=ˈmik$_1$	=mik$_1$	=m:iq:o$_1$
sal	towe$_1$	olˈɣek$_2$	ʔwe$_1$	jok:i$_{-1}$
pedra	qaʔ$_1$	ˈqaʔ$_1$	qaʔ$_1$	wet:iɢa$_2$
língua	=at͡ʃ-aʁat$_1$	=aˈt͡ʃ-aʃat$_1$	=oʔleɣ-aʁan-aʁat$_2$	=ok:el:i$_3$

<table>
<tr><th></th><th>algoritmo A</th><th>algoritmo B</th><th></th></tr>
<tr><td>manual</td><td> 
</td><td> 
</td><td>Índices de estabilidade:
nuvem 0,50
fogo 0,50
peixe 0,50
cabeça 0,75
matar 1,00
lua 0,50
nariz 1,00
sal 0,67
pedra 0,75
língua 0,50</td></tr>
<tr><td>automatizada</td><td> 
</td><td> 
</td><td>Índices de estabilidade:
nuvem 0,50
fogo 0,50
peixe 0,75
cabeça 0,75
matar 1,00
lua 0,50
nariz 1,00
sal 0,25
pedra 0,75
língua 0,50</td></tr>
</table>

Parte II. Família núbia (Egito, Sudão)

	dongolaui	kenuzi	dilling	kadaru	debri	birgid
matar	ˈbɛː$_1$	beː$_1$	hur$_2$	wur-i$_2$	wur-i$_2$	filaːl-e$_1$
lua	uˈn-at-t$_1$	an-at-ti$_1$	nɔn-ti$_1$	nɔn-tu$_1$	nɔn-to$_1$	maːl$_2$
água	ˈɛss$_1$	essi$_1$	ɔti$_1$	ɔto$_1$	ɔtu$_1$	eɟi$_1$
dar	ˈtɪr$_1$	tir$_1$	ti$_1$	ti$_1$	ti$_1$	teː-n$_1$
bom	ˈsɛrɛː$_1$	sereː$_1$	ken$_2$	kɛn$_2$	kɛŋ$_2$	azze-n$_3$
vento	ˈturug$_1$	turug$_1$	irʃ-i$_2$	irʃ-o$_2$	irʃ-o$_2$	kurr-i$_3$
cabelo	ˈdɪl-tɪ$_1$	siːr$_2$	tel-ti$_1$	til-tu$_1$	til-tu$_1$	ur=dill-e$_1$
barriga	ˈtuː$_1$	tuː$_1$	te-te$_2$	to$_1$	to$_1$	tuː$_1$
dormir	ˈnɛːr$_1$	neːr$_1$	ɟer$_1$	dwallɛli$_2$	ɟer-i$_1$	neːr-i$_1$
sol	ˈmasɪl$_1$	masɪl$_1$	ɛɟ$_2$	aju$_2$	eŋgal-to$_3$	ʔiːzi$_2$

	algoritmo A	algoritmo B	
manual			Índices de estabilidade: matar 0,50 lua 0,83 água 1,00 dar 1,00 bom 0,50 vento 0,50 cabelo 0,83 barriga 0,83 dormir 0,83 sol 0,50
automatizada			Índices de estabilidade: matar 0,33 lua 0,50 água 0,50 dar 0,67 bom 0,50 vento 0,50 cabelo 0,83 barriga 1,00 dormir 0,50 sol 0,50

(A) **(2 pontos)** A consoante **ʁ** é pronunciada como o *r* do francês ou de algumas variedades do português de Portugal, usando a parte de trás da língua. A qual classe de Dolgopolsky ela pertence e como vocês descobriram isso?

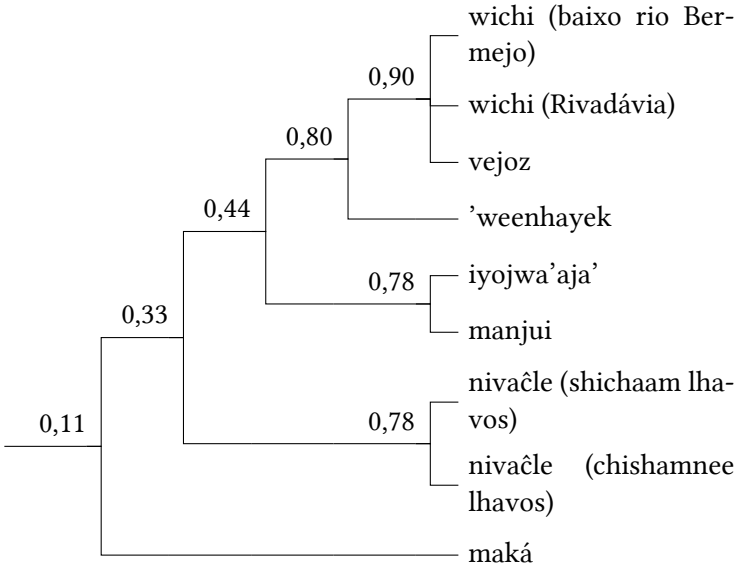
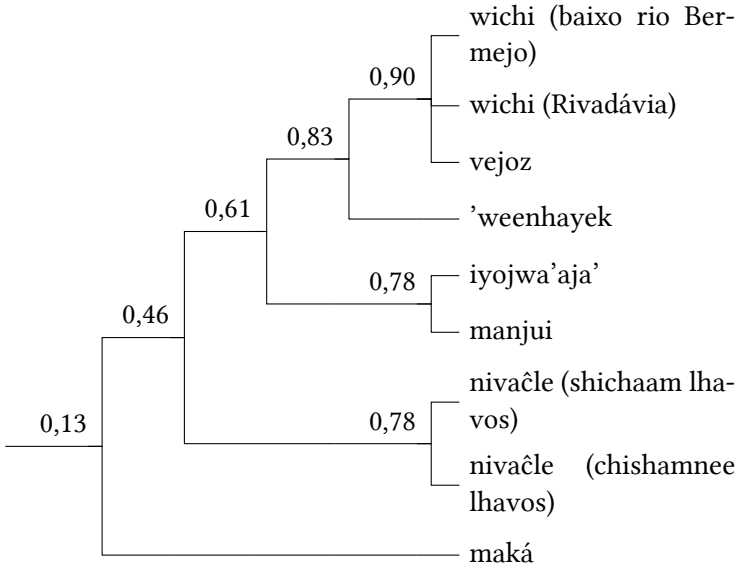
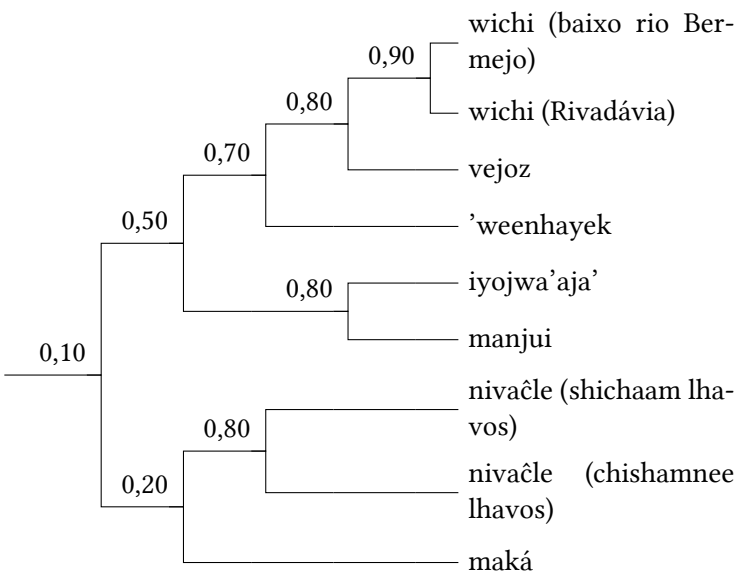
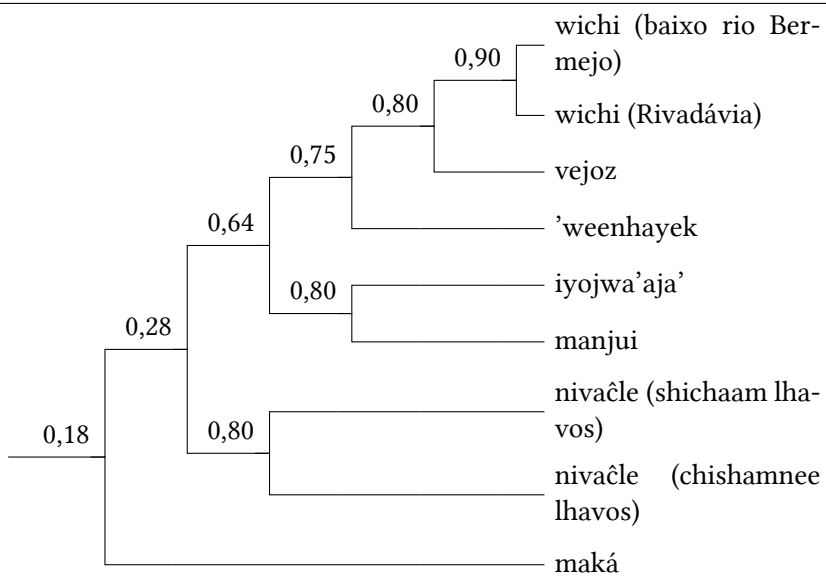
(B) **(2 pontos)** A árvore núbia no canto superior esquerdo é apenas uma de duas árvores possíveis para esta combinação do algoritmo e do tipo de anotações. Desenhem a outra árvore possível.

(C) **(2 pontos)** A árvore núbia no canto inferior esquerdo é apenas uma de duas árvores possíveis para esta combinação do algoritmo e do tipo de anotações. Desenhem a outra árvore possível.

(D) **(2 pontos)** A distância lexicoestatística 0,49, atribuída à raiz da árvore núbia no canto superior direito, foi arredondada para duas casas decimais, assim como outras distâncias neste problema. Qual é a distância exata?

Parte III. Família mataguaia (Argentina, Bolívia, Paraguai)

	wichi (baixo rio Bermejo)	wichi (Ri-vadávia)	vejoz	’weenhayek	iyojwa’aja’	manjui	nivaĉle (shichaam lhavos)	nivaĉle (chisham-nee lhavos)	maká
fogo	ʔitoχ$_{1}$	ʔitɔx$_{1}$	ʔitɑh$_{1}$	ʔiːtɑχ$_{1}$	ˈh͡wat$_{2}$	ˈʔe͡itʲe$_{1}$	ʔitɑx$_{1}$	ʔitɑx$_{1}$	feʔt$_{2}$
peixe	ʔwahat$_{1}$	wahat$_{1}$	wahat$_{1}$	ʔwaːhat$_{1}$	siˈʔjus$_{-1}$	ʃiˈʔjus$_{-1}$	saxet͡ʃ$_{-1}$	saxet͡ʃ$_{-1}$	seheț͡s$_{-1}$
pé	=pat͡ʃu$_{1}$	=qɔlɔ$_{2}$	=pat͡ʃo$_{1}$, =kɑla$_{2}$	=paːkʲʼoʔ$_{1}$	=ˈsat$_{3}$	=kaˈʔlaʔ$_{2}$	=ɸoʔ$_{4}$	=ɸoʔ$_{4}$	=fʼiʔ$_{5}$
água	ʔinot$_{1}$	ʔinɔt$_{1}$	wah$_{2}$	ʔinɑːt$_{1}$	ʔiˈnʲat$_{1}$	ʔaˈʔnat$_{1}$	jinɑʔt$_{1}$	jinɑʔt$_{1}$	iweliʔ$_{3}$
dar	=ʔweⁿ̥-u$_{1}$	=weⁿ̥-u$_{1}$	=ʔweⁿ̥-o$_{1}$	=ʔweⁿ̥-oʔ$_{1}$	=ˈwɛh͡n-aʔm$_{2}$	=ˈhajʔ$_{3}$, =ˈwɛn$_{2}$	=xut$_{4}$	=xut-ej$_{4}$	tis-ix$_{5}$
bom	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ˈʔes$_{1}$	ˈʔe͡is$_{1}$	ʔis$_{1}$	ʔis$_{1}$	t=ejkʼun-ej$_{2}$
vento	ʔinwokʷ$_{1}$	ʔinwɔk$_{1}$	ʔihwokʷ$_{1}$	=jaːɬ$_{2}$, =xʷoxʷ$_{3}$	ˈh͡lah͡wuʔ$_{4}$	ˈh͡lah͡wuuʔ$_{4}$	ɬaβiʔm$_{5}$	ɬaβiʔm$_{5}$	tʼunikʼi$_{6}$
árvore	haʔlo$_{1}$	halɔ$_{1}$	haʔlɑ$_{1}$	haʔlɑʔ$_{1}$	ʔaˈʔlaʔ$_{1}$	ʔaˈʔla-k$_{1}$	ʔaʔkxi-juk$_{2}$	jiʔk͡lɑʔ$_{1}$	naxka-k$_{3}$
cabelo	=ʔwule-j$_{1}$	=wule-j$_{1}$	=ʔwole-j$_{1}$	=ʔwoːle-ç$_{1}$, hiːlenaχ$_{2}$	=ˈʔwole$_{1}$	=ˈʔwole-j$_{1}$	=ʃateʔt͡ʃ$_{3}$	=jeʔs$_{4}$	=ʔewkux-it͡s$_{5}$
matar	=lon$_{1}$	=lɔn$_{1}$	=lɑn$_{1}$	=lɑːn̥$_{1}$	=ˈlaʔan$_{1}$	=ˈlan$_{1}$	=k͡lɑn$_{1}$	=k͡lɑn$_{1}$	=lan$_{1}$

<table>
<tr><th></th><th>algoritmo A</th><th>algoritmo B</th><th></th></tr>
<tr><td>manual</td><td> 
</td><td> 
</td><td>Índices de estabilidade:
fogo 0,78
peixe 1,00
pé 0,33
água 0,78
dar 0,44
bom 0,89
vento 0,33
árvore 0,78
cabelo 0,67
matar 1,00</td></tr>
<tr><td>automatizada</td><td> 
</td><td> 
</td><td>Índices de estabilidade:
fogo 0,78
peixe 0,44
pé 0,33
água 0,56
dar 0,67
bom 0,89
vento 0,22
árvore 0,67
cabelo 0,67
matar 1,00</td></tr>
</table>

Parte IV. Família mongólica (República Popular da China, Mongólia, Rússia)

(E) (10 pontos) Examinem a seguinte lista. Calculem os índices de estabilidade que correspondem às anotações manuais e automáticas.

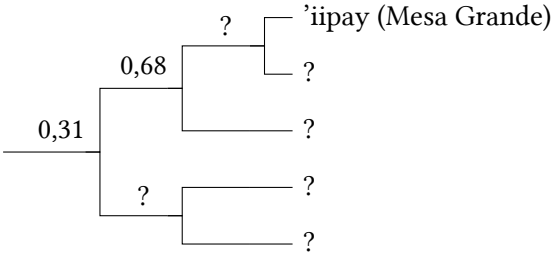
Para facilitar a tarefa de vocês, nós já calculamos os dois índices para o conceito 'todos'. Em ordem aleatória, são esses: 0,36 e 0,40.

	daur (hailar)	khamnigan (manchu)	buriato (khori)	bargut novo	ööld	khochut	calmuco	calca	ordos	chira yugur	bonã
todos	hɔː$_{1}$	bʊlt$_{2}$	bʊxiː$_{3}$	bygd$_{4}$	t͡sug$_{5}$	lug$_{5}$	t͡suk$_{5}$, xamak$_{-1}$	pux$_{3}$, pugt$_{4}$, xamăɢ$_{-1}$	pyɣyte$_{4}$, xamuk^{h}$_{-1}$	t͡ʃ^{h}uq$_{5}$	hanə$_{-2}$
casca	hai̯l͡s$_{1}$	qalihʊn$_{1}$	χɔltəhɔn$_{2}$	xaʲhuː$_{1}$	xolts$_{2}$	xalis$_{1}$	dursn$_{3}$	xɔɮt^{h}ɔ̆s$_{2}$	turusu$_{3}$	χalsən$_{1}$	arasun$_{4}$
barriga	keːli$_{1}$	gʊtəhʉn$_{2}$	gedehen$_{2}$	gedyː$_{2}$	geːs$_{2}$	gets$_{2}$	gesn$_{2}$	gitĭs$_{2}$, xiwɮĭj$_{-1}$	ketysy$_{2}$	ketesən$_{2}$	kele$_{1}$
pássaro	dəgi$_{-1}$	ɕiwan$_{1}$	ʃubuːn$_{1}$	ʃuwuː$_{1}$	ʃuvuː$_{1}$	ʃuwuː$_{1}$	ʃowun$_{1}$	ʃuwu$_{1}$	ʃuβuː$_{1}$	ʃuːn$_{1}$, pelt͡ʃɔr$_{2}$	bend͡ʑer$_{2}$
fogo	ɢali$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	qal$_{1}$	qal$_{1}$	χal$_{1}$
caminho	terg-uːl$_{1}$	qarɢʊ̯i̯$_{2}$	χarɢi̯$_{2}$, zam$_{-1}$	zam$_{-1}$	d͡zam$_{-1}$	d͡zam$_{-1}$	xaː-lʁə$_{3}$	t͡sam$_{-1}$	t͡ʃam$_{-1}$	mør$_{4}$	mor$_{4}$
sal	hataː$_{1}$	dawhʊn$_{2}$	dabhan$_{2}$	dawuhuː$_{2}$	daws$_{2}$	daws$_{2}$	dawsn$_{2}$	tawsă$_{2}$	taβusu$_{2}$	taːpsən$_{2}$	dabsuŋ$_{2}$
nadar	unpa-du$_{1}$	ʊmba-$_{1}$	t^{h}amar-$_{2}$	umb-$_{1}$	sele-$_{3}$	umba-$_{1}$	us-t͡ɕi-$_{4}$, øːm-$_{5}$	siɮĭ-$_{3}$	usu-t͡ʃ^{h}i-la-$_{4}$	umpa-$_{1}$	mba-$_{1}$
água	ɔsɔ$_{1}$	ʊxʊn$_{1}$	uhan$_{1}$	uːha$_{1}$	usn$_{1}$	us$_{1}$	usn$_{1}$	ʊsŏ$_{1}$	usun$_{1}$	q^{h}usun$_{1}$	sə$_{1}$
vento	kei̯n$_{1}$	halkin$_{2}$	halxin$_{2}$	halxi$_{2}$	salʲxin$_{2}$	salkʲi$_{2}$	salʲkn$_{2}$	saɮxĭ$_{2}$	k^{h}iː$_{1}$	k^{h}iː$_{1}$	ki$_{1}$

Parte V. Família yumana (México, EUA)

(F) **(8 pontos)** Examinem a seguinte lista. Abaixo vocês podem ver uma árvore construída com base na mesma lista. Alguns dados (nomes de línguas e distâncias lexicoestatísticas) estão faltando. Preencham as lacunas. Indiquem se a árvore é baseada em anotações manuais ou automáticas, bem como se ela foi gerada usando o algoritmo A ou B.

	mojave	cocopa	yavapai	tiipay (Jamul)	’iipay (Mesa Grande)
curto	wena=wen-a_1	ˈxʌ=ʔut$_2$	ˈt͡ʃkr=ot-i_2	lə=ˈʔuɲ$_1$	mə=put-k_3
pássaro	ʔit͡ʃ=i=jer$_1$	ˈʃa$_2$	ˈʔ=t͡ʃ=sa$_2$	aʔ=ˈʃa$_2$	ʔaː=ʂaː$_2$
osso	ɲ=a=s=ak$_1$	ˈɲ=j=aːk$_1$	ˈt͡ʃ=j=aːk-a_1	ˈak$_1$	aq$_1$
seco	i=roː-v-k$_1$	ˈʂ=ʔar$_2$	ˈru-β-i$_1$	ˈs=ʔaːj$_3$	saːj$_3$
carne	kʷiːkʷay$_1$	ʔi=ˈmaːt͡ʃ$_2$	ˌkʷeː=ˈθo-β-a_3	ˈkʷak$_4$	kukʷaːj-p_1
pescoço	maʌaqe$_1$	ˈm=puk$_2$	ˈmlqí$_1$	iː=ˈpuk$_2$	iː=puk$_2$
ver	i=juː-k$_1$	ˈwiː$_2$	ˈʔuː$_1$	ˈwiːw$_2$	ə=wuːw$_2$
rabo	iː=ʔar$_1$	ˈʃ=juʌ$_2$	ˈβ=hé$_3$	ʃə=ˈjuɬ$_2$	xə=juɬ$_2$
dois	havik-k_1	ˈx=wak$_1$	ˈh͡wâk-i_1	xə=ˈwak$_1$	xə=wak$_1$
ano	huːðe$_1$	ˌmat-ˈkaːm$_2$	ˈʔ=t͡ʃʰur-a_3	mat-ˈwam$_2$	ʔaːnʲ$_{-1}$



(G) **(20 pontos)** Foram geradas algumas outras árvores para a família yumana, com as seguintes distâncias lexicoestatísticas na raiz da árvore (as distâncias lexicoestatísticas na margem esquerda de cada árvore):

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

Desenhem cada uma dessas árvores. Para cada uma das árvores, indiquem se ela é baseada em anotações manuais ou automáticas, bem como se ela foi gerada usando o algoritmo A ou B.

(H) **(3 pontos)** Duas distâncias listadas na tarefa **(G)** foram arredondadas para duas casas decimais: o valor 0,23 foi obtido por meio de arredondamento de 0,225. Qual outra distância foi arredondada e qual é seu valor exato?

(I) **(4 pontos)** Expliquem como são calculados os índices de estabilidade.

(J) **(5 pontos)** Expliquem como são calculadas as distâncias lexicostatísticas.

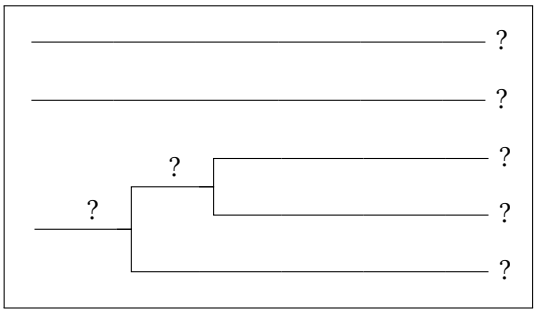
(K) **(4 pontos)** Expliquem a diferença entre os algoritmos A e B.

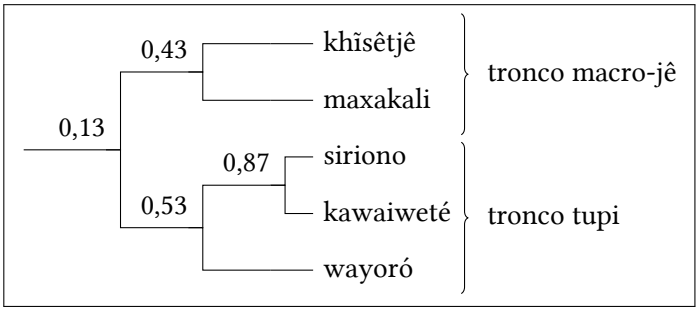
Parte VI. Tronco macro-jê e tronco tupi (Brasil, Bolívia)

(L) **(28 pontos)** Os troncos macro-jê e tupi são dois grandes agrupamentos linguísticos da América do Sul. Alguns linguistas acreditam que há uma relação distante entre eles. Examinem as seguintes listas.

	A	B	Γ	Δ	E
casca	e=ˈe-ke	hʷĩ=ˈkʰʌ	kʉpˈpɛ	mĩβ͡m=ˈt͡ɕɑj	=ˈpɛ
barriga	ˈe=ɾje	=ˈʈʰigi	=ãˈũn	=ˈtæj	=ɾɛˈwɛk
sangue	e=ˈɾuki	=kaˈⁿbɽo	=d͡ʒ=aˈʉ	=ˈhɛβ͡p	=ɾuˈɨ
queimar	=ˈɾai̯	=ɾɔ=ˈkʰɹɜ̃	=poˈkʷa	mũ=...=ˈhaβ͡p	=ɾaˈpɨ
gordura	e=ˈkiɾa	=ˈʈʰwɜmɨ	=ˈd͡ʒ=ap	=ˈtuβ͡p	=ˈkap
pé	ˈe=i	=ˈhʷaji	=ˈβi	=pɒˈta	=ˈpɨ
mão	ˈe=o	=ɲĩˈkʰɹa	=ˈβo	=ˈɲĩβ͡m	=ˈpɔ
pesado	e=ˈusi	=wɨˈʈʰĩ	=poˈti	=β͡pˈtɜj	=pɔˈɨj
fígado	ˈe=ja	=ˈⁿba	=piˈa	=t͡ɕɨβ͡pkɨ̃ˈnãj	=pɨˈʔa
novo	e=ˈjasu	=ˈⁿdɨwɨ	=paˈgop	=ˈtɨβ͡p	=pɨaˈu
raiz	e=ˈɾao	=jaˈɽe	kʉp=kʉjoˈpɛ	mĩβ͡m=ɲĩβ͡m=t͡ɕaˈti̯ə	=ɾaˈpɔ
pele	ˈe=ɨ	=ˈkʰʌ	=ˈpɛ	=ˈt͡ɕɑj	=ˈpit
rabo	e=ˈɾoko͡i̯	=ˈⁿbɨ	=d͡ʒ=oˈkʷaj	=nãː=ˈkɨβ͡p	=ˈɾaj
branco	ˈe=ʃɨ	=jaˈkʰa	=d͡ʒiˈɾa	=β͡pˈdoɯ	=ˈsĩŋ
asa	e=ˈheo	=jaˈɽa	=pɛˈo	=ɲĩˈmãɯ	=pɛˈpɔ, =jɨˈwa

Abaixo vocês podem ver duas árvores construídas com base nas mesmas listas. Alguns dados (nomes de línguas e distâncias lexicoestatísticas) estão faltando. Preencham as lacunas. Para cada uma das árvores, indiquem se ela é baseada em anotações manuais ou automáticas, bem como se ela foi gerada usando o algoritmo A ou B.





A	B	Γ	Δ	E
?	?	?	?	?

⚠ As anotações manuais e os índices de estabilidade nesta tarefa foram omitidos intencionalmente.

(M) (10 pontos) Os procedimentos automatizados baseados nas classes de Dolgopolsky podem dar resultados errados. Neste exemplo, o procedimento automatizado detecta mais semelhanças entre o siriono e uma determinada língua macro-jê (khĩsêtjê) do que entre o siriono e outras línguas do tronco tupi. Proponham e descrevam *brevemente* um procedimento modificado que resultaria em uma classificação correta se aplicado às listas macro-jê e tupi acima.

⚠ Esta tarefa será corrigida somente em caso de empate entre times mais bem classificados.

Os autores da questão agradecem a Alejandra Vidal, Maria Konoshenko, Iliá Gruntov e Jamthô Suyá por terem respondido perguntas sobre línguas específicas. —*Andrey Nikulin, Milena Vêneva*

Boa prova!