Vigesimoprimera Olimpiada Internacional de Lingüística

Brasilia (Brasil), 23–31 de julio de 2024

Problema del torneo de grupos

La lexicoestadística es un grupo de métodos creados para evaluar qué tan distante es la relación entre determinadas lenguas con base en su vocabulario. Normalmente esos métodos se aplican a largas listas de palabras anotadas por expertos, quienes señalan si, según su juicio, un determinado par de palabras proviene de la misma fuente. Sin embargo, a veces los lingüistas aplican métodos lexicoestadísticos a listas de palabras anotadas mediante procedimientos automatizados. Uno de esos procedimientos se basa en el concepto de *clases consonánticas*, introducido por Aharon Dolgopolsky, un lingüista soviético e israelí, en 1964.

P.	p b ɓ ɸ β f v	K.	k g x ɣ q ɢ χ ɯ	Y.	j ç (en el inicio de la raíz)	M.	m ɱ
T.	t d ɗ θ ð ʈ ɖ	R.	r ɾ ɽ ɹ l ɬ ɮ ɭ ʎ ɬ	W.	w ʍ (en el inicio de la raíz)	N.	n ɳ ɲ ŋ
S.	s z ʃ ʒ ʂ ʐ ɕ ʑ c ɟ					Q.	t͡ɬ d͡ɮ
H.	ħ ʕ ʜ ʢ ʔ h ɦ ʔ, vocales y j ç w ʍ (excepto en el comienzo de la raíz)						

Clases consonánticas de Dolgopolsky

A continuación, encontrarán fragmentos de listas de palabras de distintas familias lingüísticas del mundo. Las anotaciones se representan mediante índices numéricos. Con base en esas listas, se construyeron árboles de las respectivas familias lingüísticas usando dos versiones simplificadas del llamado algoritmo *StarlingNJ*, y se atribuyó un *índice de estabilidad* a cada palabra. Los árboles y los índices de estabilidad en la parte superior se basan en listas anotadas manualmente, mientras que los de la parte inferior se basan en listas anotadas automáticamente. En cada caso, hay dos árboles para cada lista, construidas con versiones distintas del algoritmo: el algoritmo A y el algoritmo B. Nótese que en algunos casos hay múltiples árboles posibles que corresponden a una lista de palabras; en esos casos se eligió un árbol aleatoriamente. A cada nudo en cada árbol se atribuye una distancia lexicoestadística. Mientras mayor es la distancia, más cercana la relación entre las lenguas. De este modo, el término más exacto sería «distancia lexicoestadística invertida» y no «distancia lexicoestadística». Por simplicidad, en este problema utilizamos el termino «distancia lexicoestadística».

Tanto los índices de estabilidad como las distancias lexicoestadísticas fueron redondeadas a dos cifras decimales. Si la tercera cifra decimal es inferior a 5, el valor se redondea por defecto; en el caso contrario, se redondea por exceso. Por ejemplo, 2,836 se redondea a 2,84, 0,705 a 0,71, 0,703 a 0,70. El redondeamiento se aplica únicamente a los valores mostrados a lectores humanos. Es decir, la máquina «ve» los valores no redondeados mientras aplica los algoritmos.

Nótese que algunas palabras son préstamos conocidos o probables de otras lenguas. Por ejemplo, la palabra **jokːi** ‘sal’ de la lengua kadiweo es un préstamo del guaraní (**jukɨ**), mientras que la palabra **ʔaːnʲ** ‘año’ del ’iipai (Mesa Grande) es un préstamo del castellano (**ˈaɲo**).

A veces se listan múltiples sinónimos, separados por comas, que corresponden a un solo concepto. Un ejemplo es el concepto ‘pie’ en la lengua vejoz.

En los datos que siguen, todos los prefijos se separan mediante el símbolo «=», mientras que los sufijos se separan mediante el símbolo «-». Algunas palabras jamás ocurren sin prefijos. Éstas comienzan con el símbolo «=».

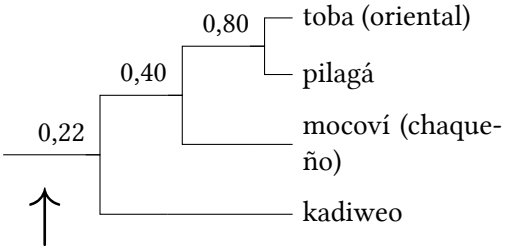
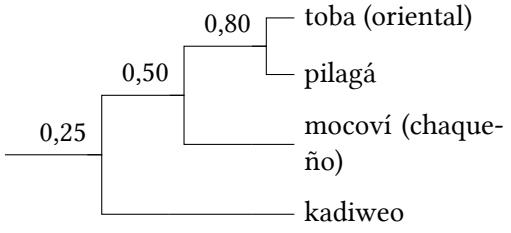
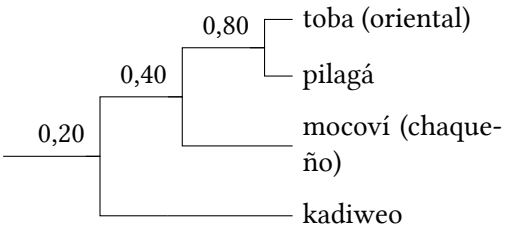
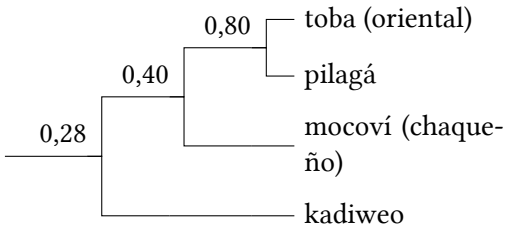
Los datos aparecen transcritos usando el Alfabeto Fonético Internacional. ˈ = acento primario, ˌ = acento secundario (más débil que el primario), ◌ː = sonido largo, ◌̆ = sonido muy corto, $\widehat{\text{XY}}$ = X y Y se

pronuncian como un único sonido, ◌́ = tono alto, ◌̀ = tono bajo, ◌̂ = tono descendente, ˀ◌ = consonante preglotalizada (precedida por un corto bloqueo del flujo de aire en la glotis), ◌’ = consonante ejectiva (acompañada de un corto bloqueo del flujo de aire en la glotis), ◌̥ = sonido sordo, ◌̃ = sonido nasalizado (pronunciado por la nariz), ◌̰ = laringización (sonido bajo, crujiente), ⁿ◌ señala un flujo de aire a través de la nariz antes de la consonante, ◌ʰ = consonante aspirada (pronunciada con un soplo de aire), ◌ʷ = consonante labializada (pronunciada con los labios redondeados), ◌ʲ = consonante palatalizada (una parte de la lengua se acerca al palato duro). **ɑ**, **æ**, **ɛ**, **ɪ**, **ɨ**, **ɔ**, **ʊ**, **ʉ**, **ə**, **ʌ**, **ɒ**, **ɘ**, **ɣ**, **ɵ**, **ø** son vocales. Otros símbolos especiales denotan consonantes.

⚠ El conocimiento de las lenguas mencionadas en este problema no da ninguna ventaja para su solución.

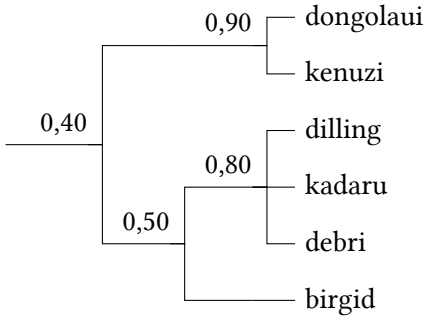
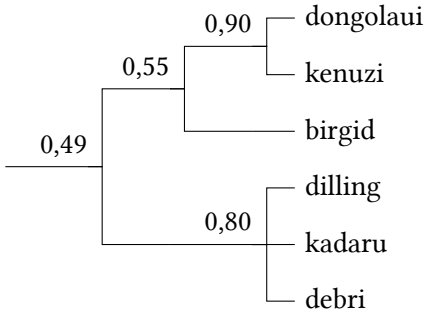
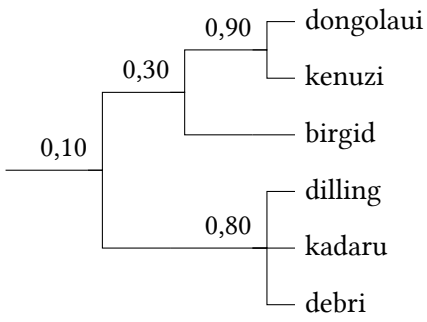
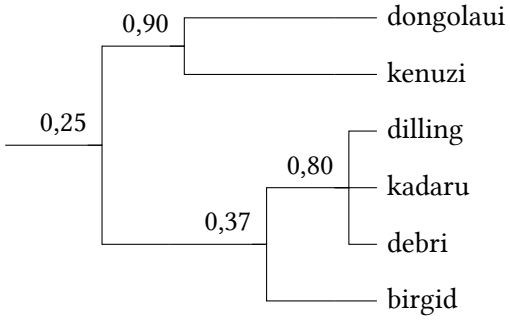
Parte I. La familia guaycurú (Argentina, Brasil, Paraguay)

	toba (oriental)	pilagá	mocoví (chaqueño)	kadiweo
nube	l=ʔok$_1$	ˈlo=ʔok$_1$	naweɣelek$_2$	lolːadi$_3$
fuego	nodek$_1$	ˈd=oleʔ$_2$	noɾek$_1$	n=olːedi$_2$
pez	ɲjaq$_1$	ˈnijaq$_1$	naʌin$_2$	nijːoɢo-d͡ʒegi$_3$
cabeza	=qajk$_1$	=ˈqajk$_1$	=qaik$_1$	=akːilo$_2$
matar	=alawat$_1$	=aˈlaːt$_1$	=alawat$_1$	=elːowadi$_1$
luna	ʔawoʁojk$_1$	ʔaˈwoʃojk$_1$	ʃiɾajɣo$_2$	epːenaj$_3$
nariz	=mik$_1$	=ˈmik$_1$	=mik$_1$	=mːiqːo$_1$
sal	towe$_1$	olˈɣek$_2$	ʔwe$_1$	jokːi$_{-1}$
piedra	qaʔ$_1$	ˈqaʔ$_1$	qaʔ$_1$	wetːiɢa$_2$
lengua	=at͡ʃ-aʁat$_1$	=aˈt͡ʃ-aʃat$_1$	=oʔleɣ-aʁan-aʁat$_2$	=okːelːi$_3$

<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 estabilidad:
nube 0,50
fuego 0,50
pez 0,50
cabeza 0,75
matar 1,00
luna 0,50
nariz 1,00
sal 0,67
piedra 0,75
lengua 0,50</td></tr>
<tr><td>automática</td><td> 
</td><td> 
</td><td>Índices de estabilidad:
nube 0,50
fuego 0,50
pez 0,75
cabeza 0,75
matar 1,00
luna 0,50
nariz 1,00
sal 0,25
piedra 0,75
lengua 0,50</td></tr>
</table>

Parte II. La familia nubia (Egipto, Sudán)

	dongolaui	kenuzi	dilling	kadaru	debri	birgid
matar	ˈbɛː$_1$	beː$_1$	hur$_2$	wur-i$_2$	wur-i$_2$	fila:l-e$_1$
luna	uˈn-at-t$_1$	an-at-ti$_1$	nɔn-ti$_1$	nɔn-tu$_1$	nɔn-to$_1$	maːl$_2$
agua	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$_1$
bueno	ˈsɛrɛː$_1$	sereː$_1$	ken$_2$	kɛn$_2$	kɛŋ$_2$	azze-n$_3$
viento	ˈturug$_1$	turug$_1$	irʃ-i$_2$	irʃ-o$_2$	irʃ-o$_2$	kurr-i$_3$
pelo	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$_2$	aɟu$_2$	eŋgal-to$_3$	ʔiːzi$_2$

	algoritmo A	algoritmo B	
manual			Índices de estabilidad: matar 0,50 luna 0,83 agua 1,00 dar 1,00 bueno 0,50 viento 0,50 pelo 0,83 barriga 0,83 dormir 0,83 sol 0,50
automática			Índices de estabilidad: matar 0,33 luna 0,50 agua 0,50 dar 0,67 bueno 0,50 viento 0,50 pelo 0,83 barriga 1,00 dormir 0,50 sol 0,50

(A) **(2 puntos)** La consonante **ʁ** se pronuncia como la *r* del francés, con la parte trasera de la lengua. ¿A qué clase de Dolgopolsky pertenece y cómo lo descubrieron?

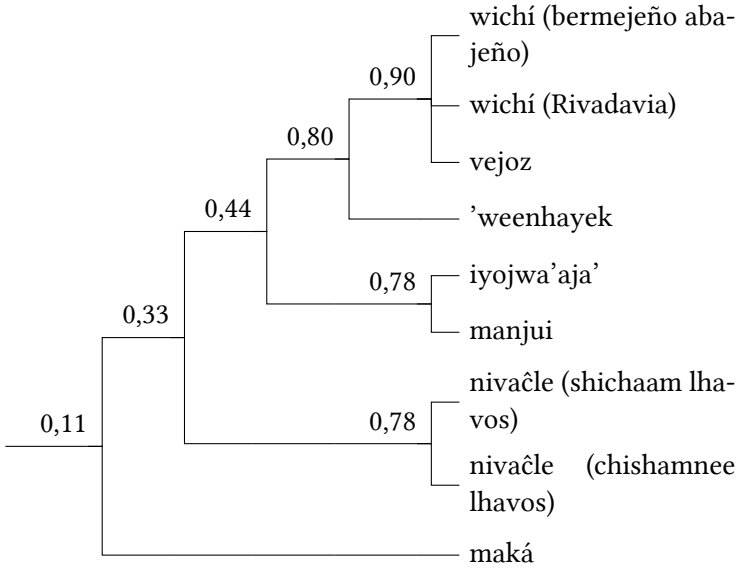
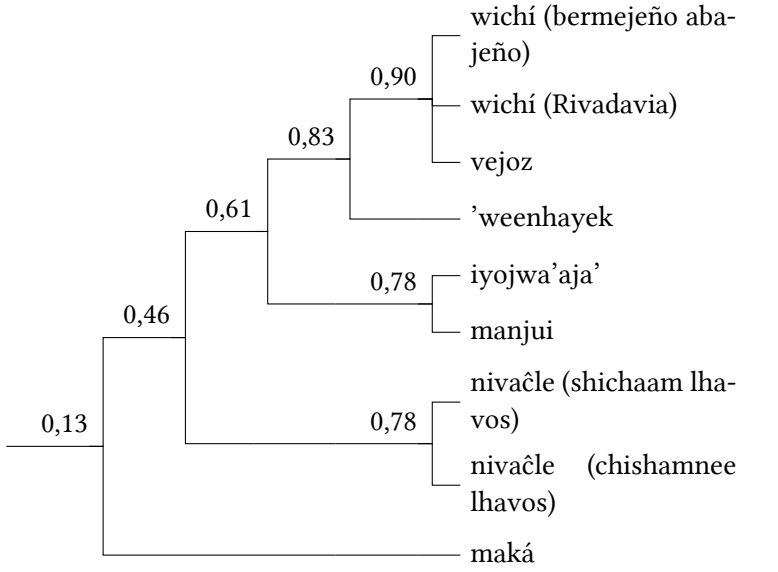
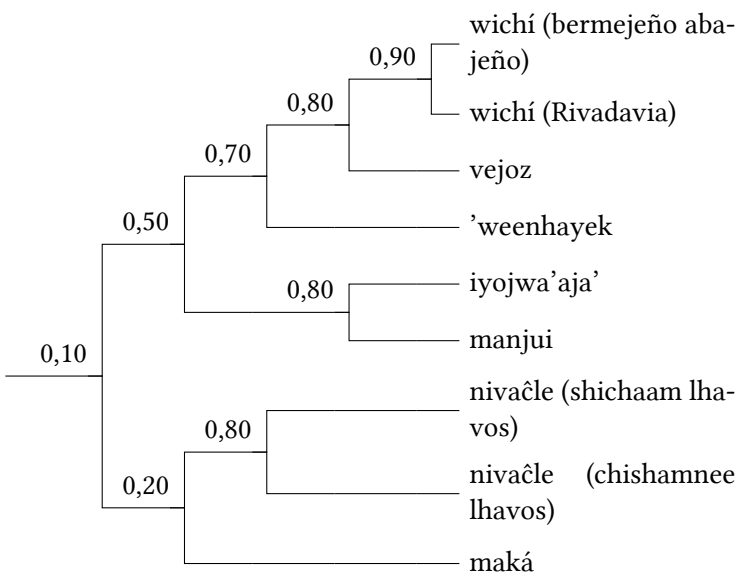
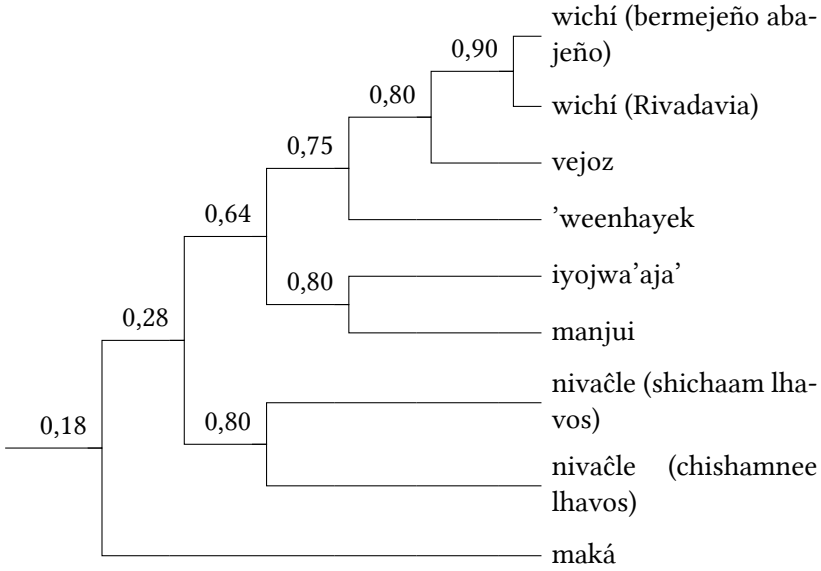
(B) **(2 puntos)** El árbol nubio en la parte superior izquierda es apenas uno de dos árboles posibles para esta combinación del algoritmo y del tipo de las anotaciones. Dibujen el otro árbol posible.

(C) **(2 puntos)** El árbol nubio en la parte inferior izquierda es apenas uno de dos árboles posibles para esta combinación del algoritmo y del tipo de las anotaciones. Dibujen el otro árbol posible.

(D) **(2 puntos)** La distancia lexicoestadística 0,49, atribuida a la raíz del árbol nubio en la parte superior derecha, fue redondeada a dos cifras decimales, como otras distancias en este problema. ¿Cuál es la distancia exacta?

Parte III. La familia mataguaya (Argentina, Bolivia, Paraguay)

	wichí (bermejeño abajeño)	wichí (Rivadavia)	vejoz	’weenhayek	iyojwa’aja’	manjui	nivaĉle (shichaam lhavos)	nivaĉle (chisham-nee lhavos)	maká
fuego	ʔitoχ$_{1}$	ʔitɔx$_{1}$	ʔitɑh$_{1}$	ʔiːtɑχ$_{1}$	ˈh͡wat$_{2}$	ˈʔei̯t̪e$_{1}$	ʔitɑx$_{1}$	ʔitɑx$_{1}$	feʔt$_{2}$
pez	ʔwahat$_{1}$	wahat$_{1}$	wahat$_{1}$	ʔwaːhat$_{1}$	siˈʔjus$_{-1}$	ʃiˈʔjus$_{-1}$	saxet͡ʃ$_{-1}$	saxet͡ʃ$_{-1}$	sehet͡s$_{-1}$
pie	=pat͡ʃu$_{1}$	=qɔlɔ$_{2}$	=pat͡ʃo$_{1}$, =kɑla$_{2}$	=paːkʲ’oʔ$_{1}$	=ˈsat$_{3}$	=kaˈʔlaʔ$_{2}$	=ɸoʔ$_{4}$	=ɸoʔ$_{4}$	=f’iʔ$_{5}$
agua	ʔinot$_{1}$	ʔinɔt$_{1}$	wah$_{2}$	ʔinɑːt$_{1}$	ʔiˈn̪at$_{1}$	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$_{5}$
bueno	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ˈʔes$_{1}$	ˈʔei̯s$_{1}$	ʔis$_{1}$	ʔis$_{1}$	t=ejk’un-ej$_{2}$
viento	ʔinwokʷ$_{1}$	ʔinwɔk$_{1}$	ʔihwokʷ$_{1}$	=jaːɬ$_{2}$, =xʷoxʷ$_{3}$	ˈh͡lah͡wuʔ$_{4}$	ˈh͡lah͡wuuʔ$_{4}$	ɬaβiʔm$_{5}$	ɬaβiʔm$_{5}$	t’unik’i$_{6}$
árbol	haʔlo$_{1}$	halɔ$_{1}$	haʔlɑ$_{1}$	haʔlɑʔ$_{1}$	ʔaˈʔlaʔ$_{1}$	ʔaˈʔla-k$_{1}$	ʔaʔkxi-juk$_{2}$	jiʔk͡lɑʔ$_{1}$	naxka-k$_{3}$
pelo	=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 estabilidad:
fuego 0,78
pez 1,00
pie 0,33
agua 0,78
dar 0,44
bueno 0,89
viento 0,33
árbol 0,78
pelo 0,67
matar 1,00</td></tr>
<tr><td>automática</td><td> 
</td><td> 
</td><td>Índices de estabilidad:
fuego 0,78
pez 0,44
pie 0,33
agua 0,56
dar 0,67
bueno 0,89
viento 0,22
árbol 0,67
pelo 0,67
matar 1,00</td></tr>
</table>

Parte IV. La familia mongólica (República Popular China, Mongolia, Rusia)

(E) (10 puntos) Examinen la lista siguiente. Calculen los índices de estabilidad que corresponden a las anotaciones manuales y automáticas.

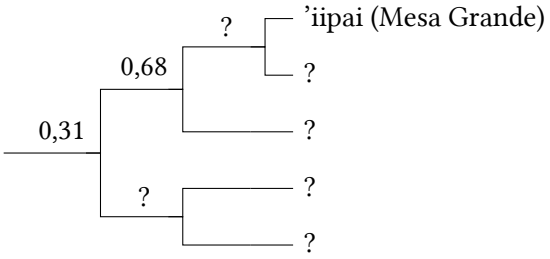
Para hacer su tarea más fácil, nosotros ya calculamos ambos índices para el concepto 'todos'. En orden aleatorio, son: 0,36 y 0,40.

	daur (jailar)	jamnigan (manchú)	buriato (jori)	bargut nuevo	eleuto	joshut	calmuco	jalja	ordós	shira yugur	bonán
todos	hɔː$_{1}$	bʊlt$_{2}$	bʊxɨː$_{3}$	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$_{-2}$
corteza	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$_{1}$
pájaro	dəgi$_{-1}$	ɕiwan$_{1}$	ʃubuːn$_{1}$	ʃuwuː$_{1}$	ʃuvuː$_{1}$	ʃuwuː$_{1}$	ʃowun$_{1}$	ʃuwu$_{1}$	ʃuβuː$_{1}$	ʃuːn$_{1}$, pelt͡ʃɔr$_{2}$	bend͡ʑer$_{2}$
fuego	gali$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	qal$_{1}$	qal$_{1}$	χal$_{1}$
camino	terg-uːl$_{1}$	qarɢʊi̯$_{2}$	χarɢɨ$_{2}$, 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-$_{1}$
agua	ɔsɔ$_{1}$	ʊxʊn$_{1}$	uhan$_{1}$	uːha$_{1}$	usn$_{1}$	us$_{1}$	usn$_{1}$	ʊsŏ$_{1}$	usun$_{1}$	q^{h}usun$_{1}$	sə$_{1}$
viento	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. La familia yumana (México, EE.UU.)

(F) **(8 puntos)** Examinen la lista siguiente. Abajo ustedes pueden observar un árbol construido con base en la misma lista. Algunos datos (nombres de lenguas y distancias lexicoestadísticas) fueron omitidos. Llenen las lagunas. Indiquen si el árbol se basa en anotaciones manuales o automáticas, así como si fue generado mediante el algoritmo A o B.

	mojave	cucapá	yavapai	tiipai (Jamul)	’iipai (Mesa Grande)
corto	wena=wen-a_1	ˈxʌ=ʔut$_2$	ˈt͡ʃkr=ot-i$_2$	lə=ˈʔuɲ$_1$	mə=put-k$_3$
pájaro	ʔit͡ʃ=i=jer$_1$	ˈʃa$_2$	ˈʔ=t͡ʃ=sa$_2$	aʔ=ˈʃa$_2$	ʔaː=ʂaː$_2$
hueso	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$_1$
cuello	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$_2$
cola	iː=ʔar$_1$	ˈʃ=juʌ$_2$	ˈβ=hé$_3$	ʃə=ˈjuɬ$_2$	xə=juɬ$_2$
dos	havik-k$_1$	ˈx=wak$_1$	ˈh͡wâk-i$_1$	xə=ˈwak$_1$	xə=wak$_1$
año	huːðe$_1$	ˌmat-ˈkaːm$_2$	ˈʔ=t͡ʃʰur-a$_3$	mat-ˈwam$_2$	ʔaːnʲ$_{-1}$



(G) **(20 puntos)** Se generaron algunos otros árboles para la familia yumana, con las siguientes distancias lexicoestadísticas en la raíz (las cuales se encuentran en el margen izquierdo de cada árbol):

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

Dibujen cada uno de estos árboles. Para cada árbol indiquen si se basa en anotaciones manuales o automáticas, así como si fue generado mediante el algoritmo A o B.

(H) **(3 puntos)** En la tarea **(G)**, dos distancias fueron redondeadas a dos cifras decimales: el valor 0,23 se obtuvo mediante redondeamiento de 0,225. ¿Cuál otra distancia fue redondeada y cuál es su valor exacto?

(I) **(4 puntos)** Expliquen cómo se calculan los índices de estabilidad.

(J) **(5 puntos)** Expliquen cómo se calculan las distancias lexicoestadísticas.

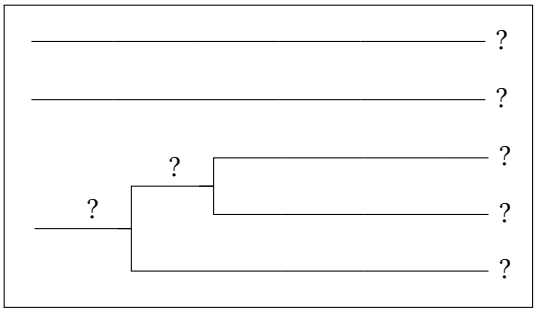
(K) **(4 puntos)** Expliquen la diferencia entre los algoritmos A y B.

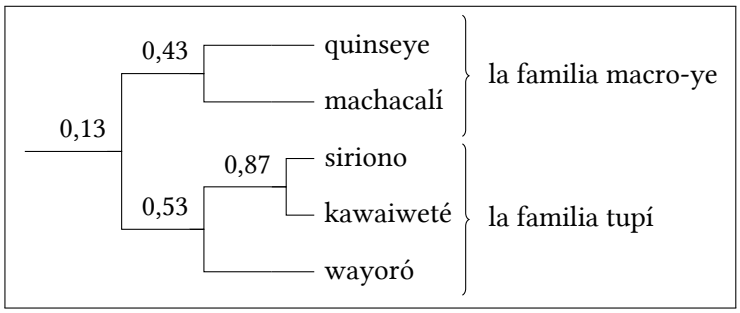
Parte VI. La familia macro-ye y la familia tupí (Brasil, Bolivia)

(L) (28 puntos) Las familias macro-ye y tupí son dos familias lingüísticas importantes de Sudamérica. Algunos lingüistas creen que están lejanamente emparentadas. Examinen las listas siguientes.

	A	B	Γ	Δ	E
corteza	e=ˈe-ke	hʷĩ=ˈkʰʌ	kʉp=ˈpɛ	mĩβ͡m=ˈt͡ɕaj	=ˈpɛ
barriga	ˈe=ɾje	=ˈʈʰigi	=ãˈʉ̃n	=ˈtæj	=ɾɛˈwɛk
sangre	e=ˈɾuki	=kaˈⁿbɭo	=d͡ʒ=aˈʉ	=ˈhɛβ͡p	=ɾuˈɨ
quemar	=ˈɾai͡	=ɾɔ=ˈkʰɹə̃	=poˈkʷa	mũ=...=ˈhaβ͡p	=ɾaˈpɨ
grasa	e=ˈkiɾa	=ˈʈʰwəmɨ	=ˈd͡ʒ=ap	=ˈtuβ͡p	=ˈkap
pie	ˈe=i	=ˈhʷaji	=ˈβi	=pɒˈta	=ˈpɨ
mano	ˈe=o	=ɲɪ̃ˈkʰɹa	=ˈβo	=ˈɲĩβ͡m	=ˈpɔ
pesado	e=ˈusi	=wɨˈʈʰĩ	=poˈti	=β͡pˈtəj	=pɔˈɨj
hígado	ˈe=ja	=ˈⁿba	=piˈa	=t͡ɕɨβ͡pkɪ̃ˈnãj	=pɨˈʔa
nuevo	e=ˈjasu	=ˈⁿdɨwɨ	=paˈgop	=ˈtɨβ͡p	=pɨaˈu
raíz	e=ˈɾao	=jaˈɽe	kʉp=kʉjoˈpɛ	mĩβ͡m=ɲĩβ͡m=t͡ɕaˈti͡ə	=ɾaˈpɔ
piel	ˈe=ɨ	=ˈkʰʌ	=ˈpɛ	=ˈt͡ɕaj	=ˈpit
cola	e=ˈɾoko͡i	=ˈⁿbɨ	=d͡ʒ=oˈkʷaj	=nãː=ˈkɨβ͡p	=ˈɾaj
blanco	ˈe=ʃɪ̃	=jaˈkʰa	=d͡ʒiˈɾa	=β͡pˈdouɻ	=ˈsĩŋ
ala	e=ˈheo	=jaˈɽa	=pɛˈo	=ɲĩˈmãuɻ	=pɛˈpɔ, =jɨˈwa

Abajo ustedes pueden observar dos árboles construidos con base en las mismas listas. Algunos datos (nombres de lenguas y distancias lexicoestadísticas) fueron omitidos. Llenen las lagunas. Para cada árbol indiquen si se basa en anotaciones manuales o automáticas, así como si fue generado mediante el algoritmo A o B.





A	B	Γ	Δ	E
?	?	?	?	?

⚠ Las anotaciones manuales y los índices de estabilidad en esta tarea han sido omitidos de manera intencional.

(M) (10 puntos) Los procedimientos automáticos basados en las clases de Dolgopolsky pueden arrojar resultados incorrectos. En este ejemplo, el procedimiento automático detecta más semejanzas entre el siriono y una determinada lengua macro-ye (quinseye) que entre el siriono y las demás lenguas tupíes. Propongan y describan *brevemente* un procedimiento modificado que generaría una clasificación correcta de las lenguas macro-ye y tupíes a partir de las listas dadas.

⚠ Esta tarea se evaluará únicamente en caso de empate entre equipos con mayor puntuación.

Los autores del problema agradecen a Alejandra Vidal, María Konoshenko, Ilyá Gruntov y Jamthô Suyá por haber respondido preguntas sobre lenguas específicas. —*Andrey Nikulin, Milena Véneva*

Editores: Iván Derzhanski (editor técnico), Hugh Dobbs, Stanislav Gurévich, Borís Iomdin, Eimear McKnight, Andrey Nikulin (editor jefe), Aleksejs Peguševs, Jan Petr, Alexánder Piperski, María Rubinstein, Milena Véneva, Elysia Warner.

Texto en castellano: Andrey Nikulin, Jenifer Vega Rodríguez.

¡Suerte!