pl

Dwudziesta pierwsza Międzynarodowa Olimpiada Lingwistyczna

Brasília (Brazylia), 23–31 lipca 2024

Zadanie turnieju drużynowego

Leksykostatystyka to zbiór metod mających za zadanie umożliwienie oszacowania stopnia pokrewieństwa języków w oparciu o ich słownictwo. W typowym przypadku metody te stosuje się do długich list wyrazów ręcznie anotowanych przez ekspertów, którzy wskazują pary słów uważanych za wywodzące się ze wspólnego źródła. Czasami jednak językoznawcy stosują metody leksykostatystyczne do list anotowanych przy pomocy automatyzowanych procedur. Jedna z procedur oparta jest na koncepcji *klas spółgłoskowych*, zaproponowanej przez radziecko-izraelskiego językoznawcę Arona Dołgopolskiego w 1964 roku.

P.	p b ɓ ɸ β f v	K.	k g x ɣ q ɢ χ ɯ	Y.	j ç (na początku rdzenia)	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 (na początku rdzenia)	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, samogłoski oraz j ç w ʍ (z wyłączeniem początku rdzenia)						

Klasy spółgłoskowe Dołgopolskiego

Poniżej znajdziecie anotowane fragmenty list wyrazów z kilku rodzin językowych z całego świata. Anotacje podano za pomocą cyfr w indeksie dolnym. W oparciu o podane listy za pomocą tzw. algorytmu *StarlingNʃ* skonstruowano drzewa genealogiczne rodzin językowych. Każdemu wyrazowi przypisano *indeks stabilności*. Drzewa i indeksy stabilności u góry bazują na listach anotowanych ręcznie, zaś dolne — na listach z automatyzowaną anotacją. Przy tym każdej liście odpowiadają dwa drzewa, oparte odpowiednio na algorytmach A i B. Zwróćcie uwagę, że w niektórych przypadkach jednej liście wyrazów może odpowiadać kilka możliwych drzew; w takich przypadkach wybrano losowo jedną z istniejących możliwości. Każdemu węzłowi na każdym drzewie przyporządkowana jest odległość leksykostatystyczna. Im wyższa jest jej wartość, tym bliżej spokrewnione są języki. Trafniej byłoby zatem mówić nie o „odległości leksykostatystycznej”, a o „odwrotnej odległości leksykostatystycznej”. Dla uproszczenia posługujemy się w tym zadaniu pojęciem „odległość leksykostatystyczna”.

Podobnie jak indeksy stabilności, odległości leksykostatystyczne zaokrąglone zostały do dwóch miejsc po przecinku. Jeżeli trzecia cyfra po przecinku jest mniejsza niż 5, zaokrąglamy w dół, w przeciwnym razie — w górę. Przykładowo, 2,836 zaokrągla się do 2,84, 0,705 do 0,71, a 0,703 do 0,70. Zaokrąglanie stosuje się wyłącznie do wartości przedstawianych odczytującemu je człowiekowi. Oznacza to, że komputer wykonujący algorytm „widzi” wartości niezaokrąglone.

Zwróćcie uwagę, że niektóre wyrazy stanowią znane lub prawdopodobne zapożyczenia z innych języków. Przykład: wyraz **jok:i** ‘sól’ w języku kadiwéu jest zapożyczeniem guarańskiego **jukɨ**, a **ʔa:nʲ** ‘rok’ w ipai (Mesa Grande) pochodzi od hiszpańskiego **ˈaɲo**.

W niektórych przypadkach do jednego znaczenia przypisano kilka synonimów, oddzielonych przecinkiem. Przykładem może tu być ‘stopa’ w języku vejoz.

W poniższym zbiorze danych wszystkie przedrostki oddzielono znakiem „=”, a przyrostki — znakiem „-”. Niektóre wyrazy występują zawsze z przedrostkami. Poprzedzone są one znakiem „=”.

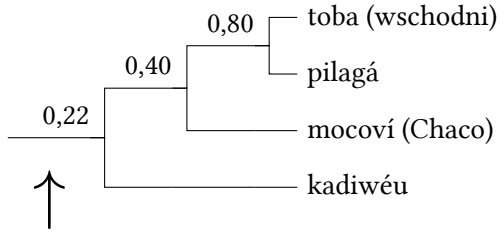
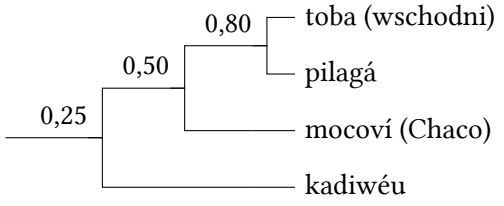
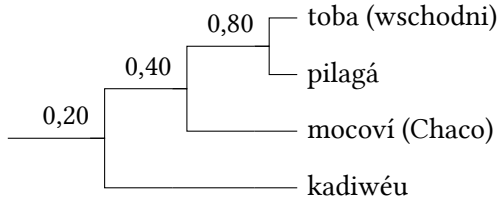
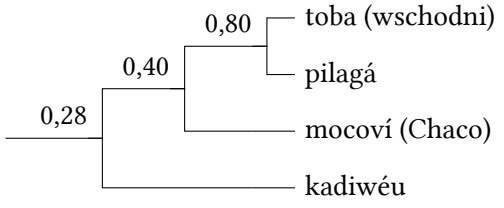
Dane zapisano w transkrypcji za pomocą Międzynarodowego Alfabetu Fonetycznego. ˈ = akcent główny, ˌ = akcent poboczny (słabszy niż akcent główny), ○ː = długa głoska, ○̆ = bardzo krótka głoska, X͡Y = X i Y wymawiane są razem, tj. jako pojedyncza głoska, ○́ = ton wysoki, ○̀ = ton niski, ○̂ = ton opadający, ˀ○ = głoska preglotalizowana (poprzedzona krótką blokadą przepływu powietrza w krtani), ○’ =

spółgłoska ejektywna (wymawiana poprzez gwałtowne zwarcie krtani), ◌̥ = spółgłoska bezdźwięczna, ◌̃ = głoska nosowa (wymawiana przez nos), ◌̰ = wymowa skrzypiąca (niski, skrzeczący głos), ⁿ◌ oznacza przepływ powietrza przez nos poprzedzający spółgłoskę, ◌ʰ = spółgłoska przydechowa (wymawiana z gwałtownym wydechem powietrza), ◌ʷ = spółgłoska labializowana (tj. wymawiana z zaokrągleniem warg), ◌ʲ = spółgłoska palatalizowana (tj. miękka). **ɑ**, **æ**, **ɛ**, **ɪ**, **ɨ**, **ɔ**, **ʊ**, **ʉ**, **ə**, **ʌ**, **ɒ**, **ɘ**, **ʏ**, **ɵ**, **ø** są samogłoskami. Pozostałe znaki specjalne oznaczają spółgłoski.

⚠ Znajomość żadnego z języków pojawiających się w zadaniu nie daje przewagi przy jego rozwiązywaniu.

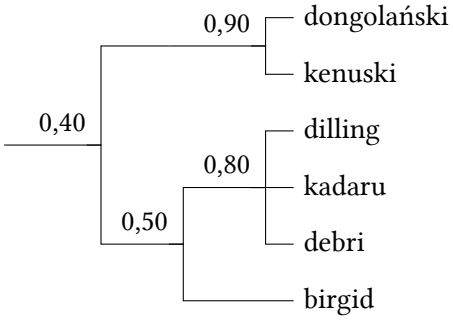
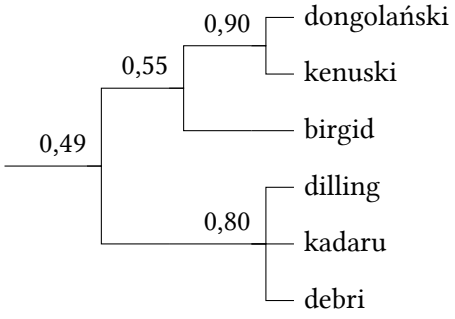
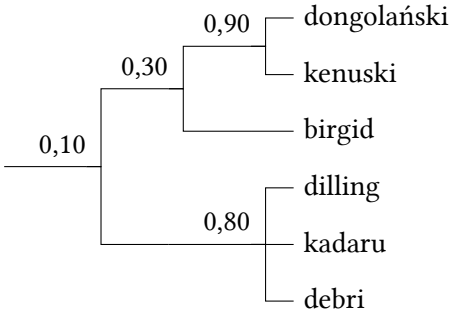
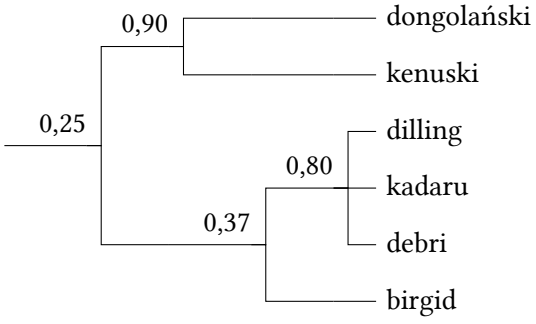
Część I. Rodzina guaicuru (Argentyna, Brazylia, Paragwaj)

	toba (wschodni)	pilagá	mocoví (Chaco)	kadiwéu
chmura	l=ʔok$_1$	ˈlo=ʔok$_1$	naweɣelek$_2$	lol:adi$_3$
ogień	nodek$_1$	ˈd=oleʔ$_2$	noɾek$_1$	n=ol:edi$_2$
ryba	njaq$_1$	ˈnijaq$_1$	naʌin$_2$	nij:oɢo-d͡ʒegi$_3$
głowa	=qajk$_1$	=ˈqajk$_1$	=qaik$_1$	=ak:ilo$_2$
zabić	=alawat$_1$	=aˈla:t$_1$	=alawat$_1$	=el:owadi$_1$
księżyc	ʔawoʁojk$_1$	ʔaˈwoʃojk$_1$	ʃiɾajɣo$_2$	ep:enaj$_3$
nos	=mik$_1$	=ˈmik$_1$	=mik$_1$	=m:iq:o$_1$
sól	towe$_1$	olˈɣek$_2$	ʔwe$_1$	jok:i$_{-1}$
kamień	qaʔ$_1$	ˈqaʔ$_1$	qaʔ$_1$	wet:iɢa$_2$
język	=at͡ʃ-aʁat$_1$	=aˈt͡ʃ-aʃat$_1$	=oʔleɣ-aʁan-aʁat$_2$	=ok:el:i$_3$

	algorytm A	algorytm B	
ręczne			Indeksy stabilności: chmura 0,50 ogień 0,50 ryba 0,50 głowa 0,75 zabić 1,00 księżyc 0,50 nos 1,00 sól 0,67 kamień 0,75 język 0,50
automatyzowane			Indeksy stabilności: chmura 0,50 ogień 0,50 ryba 0,75 głowa 0,75 zabić 1,00 księżyc 0,50 nos 1,00 sól 0,25 kamień 0,75 język 0,50

Część II. Rodzina nubijska (Egipt, Sudan)

	dongolański	kenuski	dilling	kadaru	debri	birgid
zabić	ˈbɛː$_1$	beː$_1$	hur$_2$	wur-i$_2$	wur-i$_2$	fila:l-e$_1$
księżyc	uˈn-at-t$_1$	an-at-ti$_1$	nɔn-ti$_1$	nɔn-tu$_1$	nɔn-to$_1$	ma:l$_2$
woda	ˈɛss$_1$	essi$_1$	ɔti$_1$	ɔto$_1$	ɔtu$_1$	eɟi$_1$
dać	ˈtɪr$_1$	tir$_1$	ti$_1$	ti$_1$	ti$_1$	teː-n$_1$
dobry	ˈsɛrɛː$_1$	sereː$_1$	ken$_2$	kɛn$_2$	kɛŋ$_2$	azze-n$_3$
wiatr	ˈturug$_1$	turug$_1$	irʃ-i$_2$	irʃ-o$_2$	irʃ-o$_2$	kurr-i$_3$
włosy	ˈdɪl-tɪ$_1$	siːr$_2$	tel-ti$_1$	til-tu$_1$	til-tu$_1$	ur=dill-e$_1$
brzuch	ˈtuː$_1$	tuː$_1$	te-te$_2$	to$_1$	to$_1$	tuː$_1$
spać	ˈnɛːr$_1$	neːr$_1$	ɟer$_1$	dwallɛli$_2$	ɟer-i$_1$	neːr-i$_1$
słońce	ˈmasɪl$_1$	masɪl$_1$	ɛɟ$_2$	aju$_2$	eŋgal-to$_3$	ʔiːzi$_2$

	algorytm A	algorytm B	
ręczne			Indeksy stabilności: zabić 0,50 księżyc 0,83 woda 1,00 dać 1,00 dobry 0,50 wiatr 0,50 włosy 0,83 brzuch 0,83 spać 0,83 słońce 0,50
automatyzowane			Indeksy stabilności: zabić 0,33 księżyc 0,50 woda 0,50 dać 0,67 dobry 0,50 wiatr 0,50 włosy 0,83 brzuch 1,00 spać 0,50 słońce 0,50

(A) **(2 punkty)** Spółgłoska **ʁ** wymawiana jest jak francuskie *r*, za pomocą tylnej części języka. Do której klasy Dołgopolskiego należy i w jaki sposób to ustaliliście?

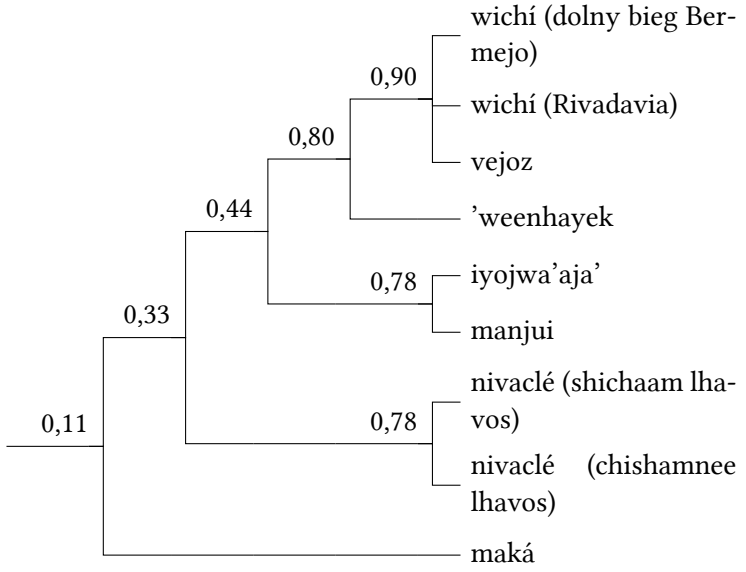
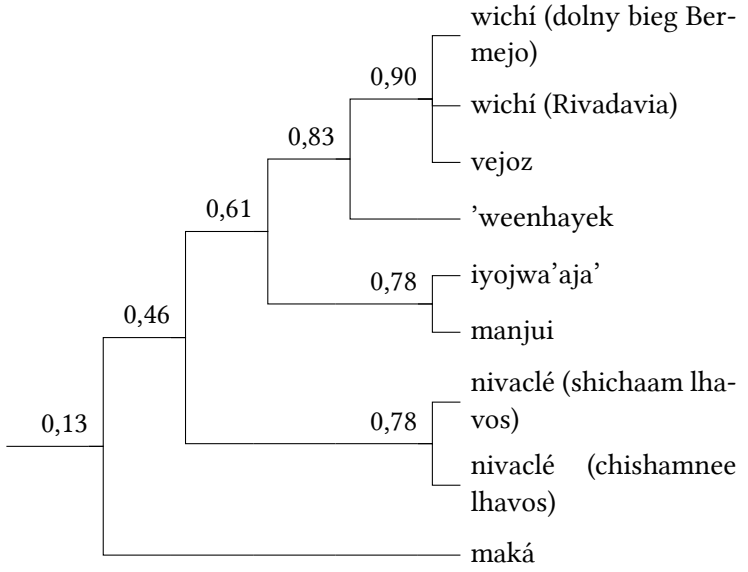
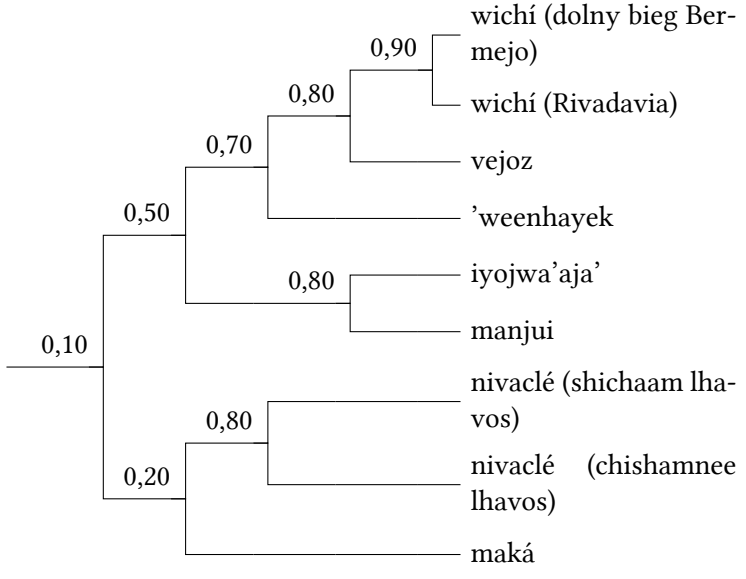
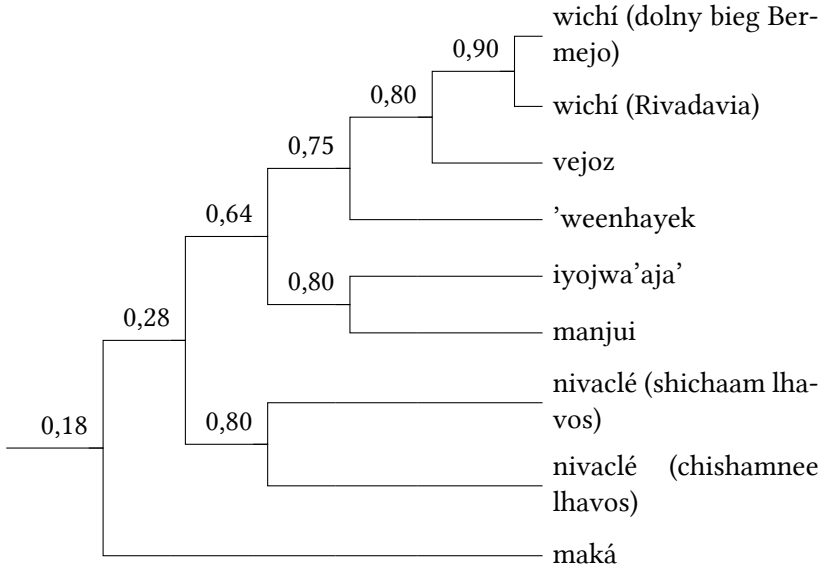
(B) **(2 punkty)** Nubijskie drzewo w lewym górnym rogu jest jednym z dwóch możliwych dla danej kombinacji algorytmu i typu anotacji. Narysujcie drugie możliwe drzewo.

(C) **(2 punkty)** Nubijskie drzewo w lewym dolnym rogu jest jednym z dwóch możliwych dla danej kombinacji algorytmu i typu anotacji. Narysujcie drugie możliwe drzewo.

(D) **(2 punkty)** Odległość leksykostatystyczna 0,49, przypisana do korzenia drzewa nubijskiego w prawym górnym rogu, została zaokrąglona do dwóch miejsc po przecinku, podobnie jak niektóre inne odległości w niniejszym zadaniu. Ile wynosi dokładna odległość?

Część III. Rodzina mataguayańska (Argentyna, Boliwia, Paragwaj)

	wichí (dolny bieg Bermejo)	wichí (Rivadavia)	vejoz	’weenhayek	iyojwa’aja’	manjui	nivaclé (shichaam lhavos)	nivaclé (chisham-nee lhavos)	maká
ogień	ʔitoχ$_{1}$	ʔitɔx$_{1}$	ʔitɑh$_{1}$	ʔi:tɑχ$_{1}$	ˈh͡wat$_{2}$	ˈʔe͡itʲe$_{1}$	ʔitɑx$_{1}$	ʔitɑx$_{1}$	feʔt$_{2}$
ryba	ʔwahat$_{1}$	wahat$_{1}$	wahat$_{1}$	ʔwa:hat$_{1}$	siˈʔjus$_{-1}$	ʃiˈʔjus$_{-1}$	saxet͡ʃ$_{-1}$	saxet͡ʃ$_{-1}$	seheț͡s$_{-1}$
stopa	=pat͡ʃu$_{1}$	=qɔlɔ$_{2}$	=pat͡ʃo$_{1}$, =kɑla$_{2}$	=pa:kʲʼoʔ$_{1}$	=ˈsat$_{3}$	=kaˈʔlaʔ$_{2}$	=ɸoʔ$_{4}$	=ɸoʔ$_{4}$	=fʼiʔ$_{5}$
woda	ʔinot$_{1}$	ʔinɔt$_{1}$	wah$_{2}$	ʔinɑ:t$_{1}$	ʔiˈnʲat$_{1}$	ʔaˈʔnat$_{1}$	jinɑʔt$_{1}$	jinɑʔt$_{1}$	iweliʔ$_{3}$
dać	=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$_{5}$
dobry	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ˈʔes$_{1}$	ˈʔe͡is$_{1}$	ʔis$_{1}$	ʔis$_{1}$	t=ejkʼun-ej$_{2}$
wiatr	ʔinwokʷ$_{1}$	ʔinwɔk$_{1}$	ʔihwokʷ$_{1}$	=ja:ɬ$_{2}$, =xʷoxʷ$_{3}$	ˈh͡lah͡wuʔ$_{4}$	ˈh͡lah͡wuuʔ$_{4}$	ɬaβiʔm$_{5}$	ɬaβiʔm$_{5}$	tʼunikʼi$_{6}$
drzewo	haʔlo$_{1}$	halɔ$_{1}$	haʔlɑ$_{1}$	haʔlɑʔ$_{1}$	ʔaˈʔlaʔ$_{1}$	ʔaˈʔla-k$_{1}$	ʔaʔkxi-juk$_{2}$	jiʔk͡lɑʔ$_{1}$	naxka-k$_{3}$
włosy	=ʔwule-j$_{1}$	=wule-j$_{1}$	=ʔwole-j$_{1}$	=ʔwo:le-ç$_{1}$, hi:lenaχ$_{2}$	=ˈʔwole$_{1}$	=ˈʔwole-j$_{1}$	=ʃateʔt͡ʃ$_{3}$	=jeʔs$_{4}$	=ʔewkux-it͡s$_{5}$
zabić	=lon$_{1}$	=lɔn$_{1}$	=lɑn$_{1}$	=lɑ:n̥$_{1}$	=ˈlaʔan$_{1}$	=ˈlan$_{1}$	=k͡lɑn$_{1}$	=k͡lɑn$_{1}$	=lan$_{1}$

<table>
<tr><th></th><th>algorytm A</th><th>algorytm B</th><th></th></tr>
<tr><td>ręczne</td><td> 
</td><td> 
</td><td>Indeksy stabilności:
ogień 0,78
ryba 1,00
stopa 0,33
woda 0,78
dać 0,44
dobry 0,89
wiatr 0,33
drzewo 0,78
włosy 0,67
zabić 1,00</td></tr>
<tr><td>automatyzowane</td><td> 
</td><td> 
</td><td>Indeksy stabilności:
ogień 0,78
ryba 0,44
stopa 0,33
woda 0,56
dać 0,67
dobry 0,89
wiatr 0,22
drzewo 0,67
włosy 0,67
zabić 1,00</td></tr>
</table>

Część IV. Rodzina mongolska (Chińska Republika Ludowa, Mongolia, Federacja Rosyjska)

(E) (10 punktów) Zapoznajcie się z poniższą listą wyrazów. Obliczcie indeksy stabilności odpowiadające anotacji ręcznej i automatyzowanej.

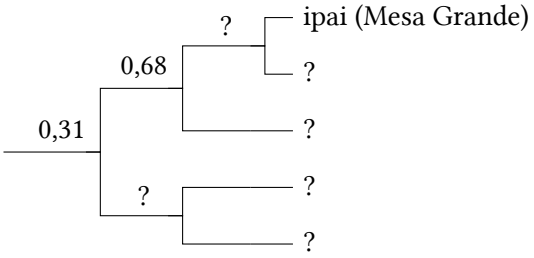
Aby ułatwić Wam zadanie, obliczyliśmy już obydwa indeksy stabilności dla wyrazu ‘wszystkie’. W przypadkowym porządku wynoszą one 0,36 oraz 0,40.

	dagurski (Hailar)	chamnigański (dial. mandżurski)	buriacki (dial. choryński)	nowobargucki	olocki	choszucki	kałmucki	chałchaski	ordoski	szera-jögurski	baoański
wszystkie	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$_{-2}$
kora	hai̯l͡s$_{1}$	qalihʊn$_{1}$	χɔltəhɔn$_{2}$	xal^{j}huː$_{1}$	xolts$_{2}$	xalis$_{1}$	dursn$_{3}$	xɔɮt^{h}ɔ̆s$_{2}$	turusu$_{3}$	χalsən$_{1}$	arasun$_{4}$
brzuch	keːli$_{1}$	ɡətəhʉn$_{2}$	gedehen$_{2}$	gedyː$_{2}$	geːs$_{2}$	gets$_{2}$	gesn$_{2}$	gitĭs$_{2}$, xiwɮĭj$_{-1}$	ketysy$_{2}$	ketesən$_{2}$	kele$_{1}$
ptak	dəgi$_{-1}$	ɕiwan$_{1}$	ʃubuːn$_{1}$	ʃuwuː$_{1}$	ʃuvuː$_{1}$	ʃuwuː$_{1}$	ʃowun$_{1}$	ʃuwu$_{1}$	ʃuβuː$_{1}$	ʃuːn$_{1}$, pelt͡ʃɔr$_{2}$	bend͡zer$_{2}$
ogień	ɡali$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	qal$_{1}$	qal$_{1}$	χal$_{1}$
droga	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$_{4}$
sól	hataː$_{1}$	dawhʊn$_{2}$	dabhan$_{2}$	dawuhuː$_{2}$	daws$_{2}$	daws$_{2}$	dawsn$_{2}$	tawsă$_{2}$	taβusu$_{2}$	taːpsən$_{2}$	dabsuŋ$_{2}$
pływać	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-$_{1}$
woda	ɔsɔ$_{1}$	ʊxʊn$_{1}$	uhan$_{1}$	uːha$_{1}$	usn$_{1}$	us$_{1}$	usn$_{1}$	ʊsŏ$_{1}$	usun$_{1}$	q^{h}usun$_{1}$	sə$_{1}$
wiatr	kei̯n$_{1}$	halkin$_{2}$	halxin$_{2}$	halxi$_{2}$	sal^{j}xin$_{2}$	salk^{j}i$_{2}$	sal^{j}kn$_{2}$	saɮxĭ$_{2}$	k^{h}iː$_{1}$	k^{h}iː$_{1}$	ki$_{1}$

Część V. Rodzina jumańska (Meksyk, Stany Zjednoczone)

(F) (8 punktów) Zapoznajcie się z poniższą listą wyrazów. Poniżej znajduje się drzewo stworzone w oparciu o tę samą listę. Część danych (nazwy języków i odległości leksykostatystyczne) została pominięta. Uzupełnijcie luki. Określcie, czy drzewo sporządzono metodą ręczną czy automatyzowaną oraz czy wygenerowano je przy wykorzystaniu algorytmu A czy B.

	mohave	kokopa	yavapai	tipai (Jamul)	ipai (Mesa Grande)
krótki	wena=wen-a_1	ˈxʌ=ʔut$_2$	ˈt͡ʃkr=ot-i_2	lə=ˈʔuɲ$_1$	mə=put-k_3
ptak	ʔit͡ʃ=i=jer$_1$	ˈʃa$_2$	ˈʔ=t͡ʃ=sa$_2$	aʔ=ˈʃa$_2$	ʔaː=ʂaː$_2$
kość	ɲ=a=s=ak$_1$	ˈɲ=j=aːk$_1$	ˈt͡ʃ=j=aːk-a_1	ˈak$_1$	aq$_1$
suchy	i=roː-v-k$_1$	ˈʂ=ʔar$_2$	ˈru-β-i$_1$	ˈs=ʔaːj$_3$	saːj$_3$
mięso	kʷiːkʷay$_1$	ʔi=ˈmaːt͡ʃ$_2$	ˌkʷeː=ˈθo-β-a_3	ˈkʷak$_4$	kukʷaːj-p_1
szyja	maʌaqe$_1$	ˈm=puk$_2$	ˈmlqíˀ$_1$	iː=ˈpuk$_2$	iː=puk$_2$
widzieć	i=juː-k$_1$	ˈwiː$_2$	ˈʔuː$_1$	ˈwiːw$_2$	ə=wuːw$_2$
ogon	iː=ʔar$_1$	ˈʃ=juʌ̊$_2$	ˈβ=hé$_3$	ʃə=ˈjuɬ$_2$	xə=juɬ$_2$
dwa	havik-k_1	ˈx=wak$_1$	ˈh͡wâk-i_1	xə=ˈwak$_1$	xə=wak$_1$
rok	huːðe$_1$	ˌmat-ˈkaːm$_2$	ˈʔ=t͡ʃʰur-a_3	mat-ˈwam$_2$	ʔaːnʲ$_{-1}$



(G) (20 punktów) Dla języków jumańskich wygenerowano kilka dalszych drzew, z następującymi odległościami leksykostatystycznymi w korzeniach drzew (tj. na lewym końcu każdego drzewa):

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

Narysujcie każde z tych drzew. Dla każdego z drzew określcie, czy zostało sporządzone metodą ręczną czy automatyzowaną oraz czy wygenerowano je przy wykorzystaniu algorytmu A czy B.

(H) (3 punkty) Dwie odległości wymienione w poleceniu **(G)** zostały zaokrąglone do dwóch miejsc po przecinku: 0,23 uzyskano poprzez zaokrąglenie 0,225. Jaką jeszcze odległość zaokrąglono i ile wynosi jej dokładna wartość?

(I) (4 punkty) Opiszcie, jak obliczane są indeksy stabilności.

(J) (5 punktów) Opiszcie, jak obliczane są odległości leksykostatystyczne.

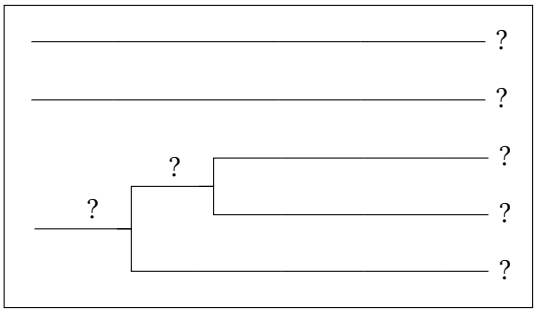
(K) (4 punkty) Wytłumaczcie różnicę między algorytmem A i algorytmem B.

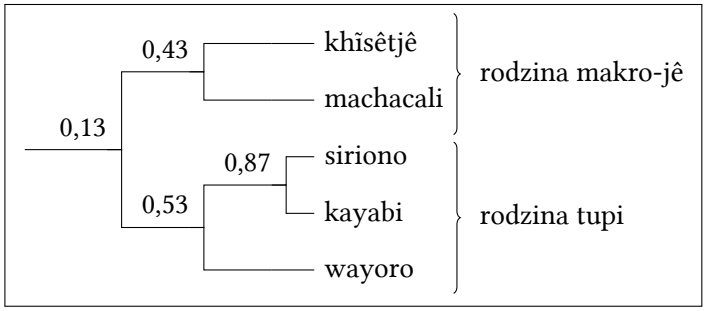
Część VI. Rodzina makro-jê i rodzina tupi (Brazylia, Boliwia)

(L) (28 punktów) Języki makro-jê i języki tupi stanowią dwie duże rodziny językowe w Ameryce Południowej. Niektórzy językoznawcy uważają je za odlegle spokrewnione. Zapoznajcie się z poniższymi listami wyrazów.

	A	B	Γ	Δ	E
kora	e='e-ke	hʷĩ='kʰʌ	kʉp='pɛ	mĩβ͡m='t͡ɕaj	='pɛ
brzuch	'e=ɾje	='ʈʰigi	=ã'ʉ̃n	='tæj	=ɾɛ'wɛk
krew	e='ɾuki	=ka'ⁿbɽo	=d͡ʒ=a'ʉ	='hɛβ͡p	=ɾu'ɨ
palić	='ɾai̯	=ɾɔ='kʰɹə̃	=po'kʷa	mũ=...='haβ͡p	=ɾa'pɨ
tłuszcz	e='kiɾa	='ʈʰwɜmɨ	='d͡ʒ=ap	='tuβ͡p	='kap
stopa	'e=i	='hʷaji	='βi	=pɒ'ta	='pɨ
ręka	'e=o	=ɲĩ'kʰɹa	='βo	='ɲĩβ͡m	='pɔ
ciężki	e='usi	=wɨ'ʈʰĩ	=po'ti	=β͡p'təj	=pɔ'ɨj
wątroba	'e=ja	='ⁿba	=pi'a	=t͡ɕɨβ͡pkɨ̃'nãj	=pɨ'ʔa
nowy	e='jasu	='ⁿdɨwɨ	=pa'gop	='tɨβ͡p	=pɨa'u
korzeń	e='ɾao	=ja'ɽe	kʉp=kʉjo'pɛ	mĩβ͡m=ɲĩβ͡m=t͡ɕa'ti̯ə	=ɾa'pɔ
skóra	'e=ɨ	='kʰʌ	='pɛ	='t͡ɕaj	='pit
ogon	e='ɾoko̯ɨ	='ⁿbɨ	=d͡ʒ=o'kʷaj	=nã:='kɨβ͡p	='ɾaj
biały	'e=ʃɨ	=ja'kʰa	=d͡ʒi'ɾa	=β͡p'douɻ	='sĩŋ
skrzydło	e='heo	=ja'ɽa	=pɛ'o	=ɲĩ'mãuɻ	=pɛ'pɔ, =jɨ'wa

Poniżej znajdują się dwa drzewa stworzone w oparciu o te same listy. Część danych (nazwy języków i odległości leksykostatystyczne) została pominięta. Uzupełnijcie luki. Dla każdego z drzew określcie, czy zostało sporządzone metodą ręczną czy automatyzowaną oraz czy wygenerowano je przy wykorzystaniu algorytmu A czy B.





A	B	Γ	Δ	E
?	?	?	?	?

⚠ Ręczne anotacje i indeksy stabilności zostały w tym poleceniu celowo pominięte.

(M) (10 punktów) Automatyzowane procedury oparte na klasach Dołgopolskiego mogą dawać nieprawidłowe wyniki. W tym przykładzie automatyzowana procedura wykrywa więcej podobieństw pomiędzy siriono i pewnym językiem makro-jê (khĩsêtjê) niż pomiędzy siriono a innymi językami tupi. Zaproponujcie i *krótko* opiszcie taką wersję automatyzowanej procedury, która dałaby prawidłową klasyfikację przy zastosowaniu do list wyrazów z rodzin makro-jê i tupi.

⚠ Odpowiedź na to polecenie będzie sprawdzona tylko w przypadku remisu drużyn z najwyższym wynikiem.

Autorzy dziękują Alejandrze Vidal, Marii Konoszenko, Ilji Gruntowowi i Jamthô Suyá za odpowiedzi na pytania dotyczące poszczególnych języków. —*Andriej Nikulin, Milena Wenewa*

Redakcja: Iwan Derżanski (red. techn.), Hugh Dobbs, Stanisław Gurewicz, Boris Iomdin, Liam McKnight, Andriej Nikulin (red. odp.), Aleksejs Peguševs, Jan Petr, Aleksandr Piperski, Maria Rubinsztejn, Elysia Warner, Milena Wenewa.

Tekst polski: Przemysław Podleśny.

Powodzenia!