

## 제 21 회 국제 언어학 올림피아드

브라질 브라질리아, 2024년 7월 23~31일

## 단체전 문제

어휘 통계학은 어휘의 유사성을 기반으로 언어 사이의 관련성을 추정하기 위해 만들어진 여러 방법들을 일컫는다. 일반적으로 이러한 방법들은 전문가가 수동으로 주석을 단 긴 어휘 목록에 적용된다. 전문가는 한 쌍의 단어마다 동일한 어휘에서 유래되었다고 생각하는지를 표시한다. 그러나 때때로 언어학자들은 자동화된 과정을 통해 주석이 달린 어휘 목록에 어휘통계적 방법을 적용한다. 그러한 절차 중 하나는 소련 출신 이스라엘 언어학자 아하론 돌고폴스키가 1964년에 도입한 자음 분류라는 개념에 기초한다.

P. p b b φ β f v	K. k g x y q g x u	Y. j ç (어근의 첫 소리일 경우)	M. m ɱ
T. t d d θ ð t d	R. r r r l l k k l l	W. w m (어근의 첫 소리일 경우)	N. n ɲ ɳ ɳ
S. s z ʃ ʒ s z ɛ z c j			Q. t̪ d̪
H. h ʃ h ʃ ʔ h h ʔ, 모음, 그리고 j ç w m (어근의 첫 소리가 아닐 경우)			

## 돌고폴스키의 자음 분류

아래에는 세계 여러 어족의 어휘 목록 중 일부에 주석이 달려 제시되어 있다. 주석은 아래 첨자 숫자의 형태로 주어졌다. 이 목록들을 기반으로 이른바 *StarlingNJ* 알고리즘의 두 가지 단순화된 버전을 사용하여 어족 수형도가 작성되었으며, 각 단어에 안정 지수가 할당되었다. 상단에 있는 수형도와 안정 지수는 수동으로 주석을 단 어휘 목록에 근거하며, 하단에 있는 수형도와 안정 지수는 자동으로 주석을 단 어휘 목록에 근거한다. 어휘 목록마다 A 알고리즘과 B 알고리즘을 이용하여 각각 두 개의 수형도가 생성되었다. 경우에 따라 하나의 어휘 목록에 여러 개의 수형도가 대응할 수 있다는 점에 유의하여야. 그런 경우에는 하나의 수형도만 무작위로 선택되었다. 각 수형도의 각 분기점마다 어휘통계적 거리가 할당되었다. 거리의 값이 클수록 언어 사이의 관계가 가깝다는 뜻이다. 따라서 더욱 정확한 용어는 “어휘통계적 거리”가 아니라 “역 어휘통계적 거리”라고 할 수 있다. 편의상 이 문제에서는 “어휘통계적 거리”라는 용어가 사용된다.

안정 지수와 어휘통계적 거리는 둘 다 소수점 셋째 자리에서 반올림한다. 소수점 셋째 자리의 숫자가 5 미만일 경우 버림하고 초과할 경우 올림한다. 예를 들어, 2.836은 2.84로, 0.705는 0.71로, 0.703은 0.70으로 반올림한다. 반올림은 사용자에게 보이는 값에만 적용된다. 즉, 알고리즘을 실행하는 컴퓨터는 반올림하지 않은 값을 그대로 “본다”.

일부 어휘는 다른 언어로부터 빌려온 것으로 알려져 있거나 의심된다는 점에 유의하여야. 예를 들어, 카디웨우어어 *joki* ‘소금’이란 단어는 과라니어의 *juki*에서 차용되었으며, 이파이어 (메사그란테 방언) 어 *ʔamj* ‘년(시간)’은 에스파냐어의 *ajno*에서 차용되었다.

하나의 의미에 대한 여러 유의어가 어휘 목록에 제시되는 경우, 심표로 구분되어 나열된다. 하나의 예시로는, 베호스어 ‘밭(신체)’이 여기에 해당한다.

아래 데이터에서 모든 접두사는 기호 “=”로 구분되며, 모든 접미사는 기호 “-”로 구분된다. 일부 단어는 접두사 없이 사용되지 않는다. 그러한 경우, 단어 맨 앞에 기호 “=”가 붙는다.

모든 자료는 국제 음성 기호를 이용하여 전사되어 있다. ' = 제1 강세, , = (제1 강세보다 약한) 제2 강세, ː = 긴 소리, ˘ = 아주 짧은 소리,  $\widehat{XY}$  = X와 Y는 하나의 소리로 발음됨, ˆ = 높은 성조, ˜ = 낮은 성조, ˆ = 하강하는 성조, ˆ˘ = 전성문음화된 소리 (목구멍에서 공기의 흐름을 잠시 차단한 다음 내는 소리), ˘˘ = 방출음 (목구멍에서 공기의 흐름을 잠시 차단하면서 내는 소리), ˘˘ = 무성음, ˘˘ = 비음화된 소리 (코를 통해 발음하는 소리), ˘˘ = 뺨격이는 소리(낮고 굽는 듯한 소리), ˘˘은 자음 앞에 공기의 일부를 코로 내 보내며 발음하는 것을 나타낸다, ˘˘ = 유기음(거센 공기를 터뜨리며 발음하는 소리), ˘˘ = 순음화된 자음(입술을 둥글게 오므리고 발음하는 자음), ˘˘ = 구개음화된 소리(혀바닥을 입천장 단단한 부분 가까이 대고 발음하는 소리). **a, æ, ɛ, i, ɨ, ɔ, u, ʌ, ɒ, ɔ, y, ɐ, ø**는 모음이다. 그 외의 특수문자들은 자음이다.

⚠ 본 문제에 등장하는 언어에 대한 지식은 문제를 해결하는 데에 있어 유리하게 작용하지 않는다.

제 1 부. 과이쿠루아어족 (아르헨티나, 브라질, 파라과이)

	토바어 (동부)	필라가어	모코비어 (차코 방언)	카디웨우어
구름	l=ʔok <sub>1</sub>	'lo=ʔok <sub>1</sub>	naweyelek <sub>2</sub>	lol:adi <sub>3</sub>
불	nodek <sub>1</sub>	'd=oleʔ <sub>2</sub>	norek <sub>1</sub>	n=ol:edi <sub>2</sub>
물고기	njaq <sub>1</sub>	'nijaq <sub>1</sub>	naʎin <sub>2</sub>	nij:ogo-dʒegi <sub>3</sub>
머리	=qajk <sub>1</sub>	=ʔajk <sub>1</sub>	=qaik <sub>1</sub>	=ak:ilo <sub>2</sub>
죽이다	=alawat <sub>1</sub>	=a'la:t <sub>1</sub>	=alawat <sub>1</sub>	=el:owadi <sub>1</sub>
달(천체)	ʔawoʝok <sub>1</sub>	ʔa'woʝok <sub>1</sub>	ʃirajʝo <sub>2</sub>	ep:enaj <sub>3</sub>
코	=mik <sub>1</sub>	=ʔmik <sub>1</sub>	=mik <sub>1</sub>	=m:iq:o <sub>1</sub>
소금	towe <sub>1</sub>	ol'ʝek <sub>2</sub>	ʔwe <sub>1</sub>	jok:i <sub>-1</sub>
돌	qaʔ <sub>1</sub>	'qaʔ <sub>1</sub>	qaʔ <sub>1</sub>	wet:iga <sub>2</sub>
혀	=atʃ-awat <sub>1</sub>	=a'ʃ-awat <sub>1</sub>	=oʔley-awan-awat <sub>2</sub>	=ok:eli:3

	A 알고리즘	B 알고리즘	
변수	<p>어휘통계적 거리</p>		지정 지수: 구름 0.50 불 0.50 물고기 0.50 머리 0.75 죽이다 1.00 달(천체) 0.50 코 1.00 소금 0.67 돌 0.75 혀 0.50
지동			지정 지수: 구름 0.50 불 0.50 물고기 0.75 머리 0.75 죽이다 1.00 달(천체) 0.50 코 1.00 소금 0.25 돌 0.75 혀 0.50

제 2부. 누비아어족 (이집트, 수단)

	동골라위어	케누지어	딜링어	카다루어	데브리어	비르기드어
죽이다	'bɛ:₁	be:₁	hur₂	wur-i₂	wur-i₂	fila:l-e₁
달(천체)	u'n-at-t₁	an-at-ti₁	nɔn-ti₁	nɔn-tu₁	nɔn-to₁	ma:l₂
물	'ɛss₁	essi₁	ɔti₁	ɔto₁	ɔtu₁	eji₁
주다	'tir₁	tir₁	ti₁	ti₁	ti₁	te:-n₁
좋은	'sɛrɛ:₁	sere:₁	ken₂	kɛn₂	kɛŋ₂	azze-n₃
바람	'turug₁	turug₁	irf-i₂	irf-o₂	irf-o₂	kurr-i₃
머리카락	'dil-ti₁	si:r₂	tel-ti₁	til-tu₁	til-tu₁	ur=dill-e₁
배(신체)	'tu:₁	tu:₁	te-te₂	to₁	to₁	tu:₁
자다	'nɛ:r₁	ne:r₁	jer₁	dwallɛli₂	jer-i₁	ne:r-i₁
해(천체)	'masil₁	masil₁	ɛj₂	aju₂	ɛŋgal-to₃	ʔi:zi₂

	A 알고리즘	B 알고리즘	
상수			<p>상수</p> <p>안정 지수:</p> <ul style="list-style-type: none"> <li>죽이다 0.50</li> <li>달(천체) 0.83</li> <li>물 1.00</li> <li>주다 1.00</li> <li>좋은 0.50</li> <li>바람 0.50</li> <li>머리카락 0.83</li> <li>배(신체) 0.83</li> <li>자다 0.83</li> <li>해(천체) 0.50</li> </ul>
자동			<p>자동</p> <p>안정 지수:</p> <ul style="list-style-type: none"> <li>죽이다 0.33</li> <li>달(천체) 0.50</li> <li>물 0.50</li> <li>주다 0.67</li> <li>좋은 0.50</li> <li>바람 0.50</li> <li>머리카락 0.83</li> <li>배(신체) 1.00</li> <li>자다 0.50</li> <li>해(천체) 0.50</li> </ul>

- (A) (2 점) 자음 **ʁ**은 마치 프랑스어의 *r*처럼, 혀의 뒷부분에서 발음된다. 이 자음은 돌고폴스키의 어느 분류에 속하며, 어떻게 그 결론을 도출하였는가?
- (B) (2 점) 왼쪽 상단에 있는 누비아어군의 수형도는 지정된 알고리즘과 주석 유형을 조합했을 경우 가능한 두 개의 수형도 중 하나일 뿐이다. 다른 가능한 수형도를 그려라.
- (C) (2 점) 왼쪽 하단에 있는 누비아어군의 수형도는 지정된 알고리즘과 주석 유형을 조합했을 경우 가능한 두 개의 수형도 중 하나일 뿐이다. 다른 가능한 수형도를 그려라.
- (D) (2 점) (오른쪽 상단의 누비아어군 수형도의 뿌리에 할당된) 어휘통계적 거리 0.49는 이 문제의 다른 거리와 마찬가지로 소수점 셋째 자리에서 반올림되었다. 정확한 거리는 얼마인가?

제 3부. 마타과야어족 (아르헨티나, 볼리비아, 파라과이)

	위치어 (베르메호 강의 하류 방언)	위치어 (리바다비아 방언)	베호스어	웬나예크어	이요화아하어	만후이어	니바클레어 (하류 방언)	니바클레어 (상류 방언)	마카어
불	ʔitox <sub>1</sub>	ʔitox <sub>1</sub>	ʔitah <sub>1</sub>	ʔi:tax <sub>1</sub>	ʰwat <sub>2</sub>	ʔeith <sub>1</sub>	ʔitax <sub>1</sub>	ʔitax <sub>1</sub>	feʔt <sub>2</sub>
물고기	ʔwahat <sub>1</sub>	wahat <sub>1</sub>	wahat <sub>1</sub>	ʔwa:hat <sub>1</sub>	siʔjus <sub>-1</sub>	ʃiʔjus <sub>-1</sub>	saxetʃ <sub>-1</sub>	saxetʃ <sub>-1</sub>	sehets <sub>-1</sub>
밭(신체)	=patʃu <sub>1</sub>	=qolb <sub>2</sub>	=patʃo <sub>1</sub> , =kala <sub>2</sub>	=pa:kʰoʔ <sub>1</sub>	=ʰsat <sub>3</sub>	=kaʔlaʔ <sub>2</sub>	=φoʔ <sub>4</sub>	=φoʔ <sub>4</sub>	=fʰiʔ <sub>5</sub>
물	ʔinot <sub>1</sub>	ʔinot <sub>1</sub>	wah <sub>2</sub>	ʔina:t <sub>1</sub>	ʔiʰnat <sub>1</sub>	ʔaʰnat <sub>1</sub>	jinaʔt <sub>1</sub>	jinaʔt <sub>1</sub>	iweliʔ <sub>3</sub>
주다	=ʔwen̩-u <sub>1</sub>	=wen̩-u <sub>1</sub>	=ʔwen̩-o <sub>1</sub>	=ʔwen̩-oʔ <sub>1</sub>	=ʰwɛhn̩-aʔm <sub>2</sub>	=ʰhajʔ <sub>3</sub> , =ʰwɛn̩ <sub>2</sub>	=xut <sub>4</sub>	=xut-ej <sub>4</sub>	tis-ix <sub>5</sub>
좋은	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=ejkʰun-ej <sub>2</sub>
바람	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>	ʰhlahwuʔ <sub>4</sub>	ʰhlahwu <sup>u</sup> ʔ <sub>4</sub>	ʔaβiʔm <sub>5</sub>	ʔaβiʔm <sub>5</sub>	tʰunikʰi <sub>6</sub>
나무	haʔlo <sub>1</sub>	halb <sub>1</sub>	haʔla <sub>1</sub>	haʔlaʔ <sub>1</sub>	ʔaʔlaʔ <sub>1</sub>	ʔaʔla-k <sub>1</sub>	ʔaʔkxi-juk <sub>2</sub>	ʔiʔklaʔ <sub>1</sub>	naxka-k <sub>3</sub>
머리카락	=ʔwule-j <sub>1</sub>	=wule-j <sub>1</sub>	=ʔwole-j <sub>1</sub>	=ʔwo:le-ç <sub>1</sub> , hi:lenax <sub>2</sub>	=ʔwole <sub>1</sub>	=ʔwole-j <sub>1</sub>	=ʔateʔtʃ <sub>3</sub>	=jeʔs <sub>4</sub>	=ʔewkux-its <sub>5</sub>
죽이다	=lon <sub>1</sub>	=lɔn <sub>1</sub>	=lan <sub>1</sub>	=la:n̩ <sub>1</sub>	=ʰlaʔan <sub>1</sub>	=ʰlan <sub>1</sub>	=klan <sub>1</sub>	=klan <sub>1</sub>	=lan <sub>1</sub>

	A 알고리즘	B 알고리즘	
수동	<p>0.11</p> <p>0.33</p> <p>0.44</p> <p>0.78</p> <p>0.80</p> <p>0.90</p> <p>0.78</p> <p>0.78</p>	<p>0.13</p> <p>0.46</p> <p>0.61</p> <p>0.78</p> <p>0.83</p> <p>0.90</p> <p>0.78</p> <p>0.78</p>	<p>안정 지수:</p> <p>불 0.78</p> <p>물고기 1.00</p> <p>밭(신체) 0.33</p> <p>물 0.78</p> <p>주다 0.44</p> <p>좋은 0.89</p> <p>바람 0.33</p> <p>나무 0.78</p> <p>머리카락 0.67</p> <p>죽이다 1.00</p>
자동	<p>0.10</p> <p>0.50</p> <p>0.70</p> <p>0.80</p> <p>0.90</p> <p>0.80</p> <p>0.80</p> <p>0.20</p> <p>0.80</p>	<p>0.18</p> <p>0.28</p> <p>0.64</p> <p>0.80</p> <p>0.75</p> <p>0.80</p> <p>0.90</p> <p>0.80</p> <p>0.80</p>	<p>안정 지수:</p> <p>불 0.78</p> <p>물고기 0.44</p> <p>밭(신체) 0.33</p> <p>물 0.56</p> <p>주다 0.67</p> <p>좋은 0.89</p> <p>바람 0.22</p> <p>나무 0.67</p> <p>머리카락 0.67</p> <p>죽이다 1.00</p>

제 4부. 몽골어족 (중국, 몽골, 러시아)

(E) (10 점) 다음 어휘 목록을 살펴보아라. 수동 및 자동 주석에 대응하는 안정 지수를 계산하여라.

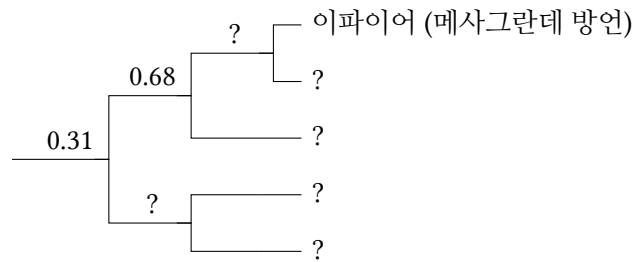
문제 해결을 쉽게 하기 위하여, 단어 ‘모든’에 대한 해당 두 가지의 안정 지수를 미리 계산하여 제공한다. 이는 0.36, 0.40 이다 (순서는 상관 없음).

	다 구 르 어 (하 일 라 르 방 언)	함 니 간 어 (만 주 방 언)	부 라 트 어 (호 리 방 언)	신 바 르 구 트 어	울 드 어	호 슈 트 어	칼 미 크 어	할 하 어	오 르 도 스 어	시 라 유 구 르 어	바 오 안 어
모든	hɔ:₁	bolt₂	buxi:₃	bygd₄	tsug₅	lug₅	tsuk₅, xamak₋₁	pux₃, pugt₄, xamāg₋₁	pyyite₄, xamukᵃ₋₁	tʃᵃuq₅	hanə₂
나무껍질	hails₁	qalihon₁	χoltōhōn₂	xalʃhu:₁	xolts₂	xalis₁	dursn₃	xɔɮtᵃᵃs₂	turusu₃	χalsən₁	arasun₄
배 (신체)	ke:li₁	getəhən₂	gedehen₂	gedy:₂	ge:s₂	gets₂	gesn₂	gitis₂, xiwɮij₋₁	ketysy₂	ketesən₂	kele₁
새 (동물)	dəgi₋₁	eiwan₁	ʃubun₁	ʃuwu:₁	ʃuvu:₁	ʃuwu:₁	ʃowun₁	ʃuwu₁	ʃuβu:₁	ʃu:n₁, peltʃər₂	bendzer₂
불	gali₁	gal₁	gal₁	gal₁	gal₁	gal₁	gal₁	gal₁	qal₁	qal₁	χal₁
길	terg-u:l₁	qargūi₂	χargi₂, zam₋₁	zam₋₁	dzam₋₁	dzam₋₁	xa:-lɔə₃	tsam₋₁	tʃam₋₁	mør₄	mor₄
소금	hata:₁	dawhōn₂	dabhan₂	dawuhu:₂	daws₂	daws₂	dawsn₂	tawsă₂	taβusu₂	ta:psən₂	dabsuŋ₂
헤엄치다	unpa-du₁	umba₋₁	tᵃamar₋₂	umb₋₁	sele₋₃	umba₋₁	us-təi₋₄, ø:m₋₅	siɮi₋₃	usu-tʃᵃi-la₋₄	umpa₋₁	mba₋₁
물	ɔsɔ₁	oxon₁	uhan₁	u:ha₁	usn₁	us₁	usn₁	ʊšö₁	usun₁	qᵃusun₁	sə₁
바람	kein₁	halkin₂	halxin₂	halxi₂	salʃxin₂	salkji₂	salʃkn₂	saɮxi₂	kᵃi:₁	kᵃi:₁	ki₁

**제 5부. 유마어족 (멕시코, 미국)**

(F) (8 점) 다음 어휘 목록을 살펴보아라. 아래에는 같은 목록을 기반으로 작성된 수행도가 있다. 일부 데이터(언어 이름과 어휘통계적 거리)는 숨겨져 있다. 빈칸을 채워라. 수행도가 수동으로 생성되었는지, 자동으로 생성되었는지를 명시하고, A 알고리즘을 이용하여 생성되었는지, B 알고리즘을 이용하여 생성되었는지를 명시하여라.

	모하비어	코코파어	야바파이어	티이파이어 (하물 방언)	이파이어 (메사그란데 방언)
짧은	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>
새(동물)	ʔitʃ=i=jer <sub>1</sub>	'ʃa <sub>2</sub>	'ʔ=ʃsa <sub>2</sub>	aʔ='ʃa <sub>2</sub>	ʔa:=ʃa:₂
뼈	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>
건조한	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>
고기	kwi:kway <sub>1</sub>	'ʔi='ma:tʃ <sub>2</sub>	'kʷe:='θo-β-a <sub>3</sub>	'kʷak <sub>4</sub>	kukʷa:j-p <sub>1</sub>
목	maʎaqe <sub>1</sub>	'm=puk <sub>2</sub>	'mlq <sub>1</sub>	i:='puk <sub>2</sub>	i:=puk <sub>2</sub>
보다	i=ju:-k <sub>1</sub>	'wi:₂	'ʔu:₁	'wi:w <sub>2</sub>	ə=wu:w <sub>2</sub>
꼬리	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>
둘	havik-k <sub>1</sub>	'x=wak <sub>1</sub>	'h̄wâk-i <sub>1</sub>	xə='wak <sub>1</sub>	xə=wak <sub>1</sub>
년(시간)	hu:ðe <sub>1</sub>	'mat-'ka:m <sub>2</sub>	'ʔ=tʃ̄ur-a <sub>3</sub>	mat-'wam <sub>2</sub>	ʔa:m̄ <sub>1</sub>



(G) (20 점) 이 외에 유마어족 수행도가 몇 개 더 생성되었으며, 각 수행도의 뿌리에 있는 어휘통계적 거리(즉, 각 수행도의 가장 왼쪽에 있는 어휘통계적 거리)는 다음과 같다:

1. 0.20
2. 0.23
3. 0.24

각각의 수행도를 그려라. 각 수행도마다 수동으로 생성되었는지, 자동으로 생성되었는지를 명시하고, A 알고리즘을 이용하여 생성되었는지, B 알고리즘을 이용하여 생성되었는지를 명시하여라.

(H) (3 점) 문항 (G)에 나열된 거리 중 두 개는 소수점 셋째 자리에서 반올림되었다: 0.23은 0.225에서 반올림되었다. 또 하나의 반올림된 거리는 무엇이며, 그 정확한 값은 무엇인가?

(I) (4 점) 안정 지수를 계산하는 방법을 설명하여라.

(J) (5 점) 어휘통계적 거리를 계산하는 방법을 설명하여라.

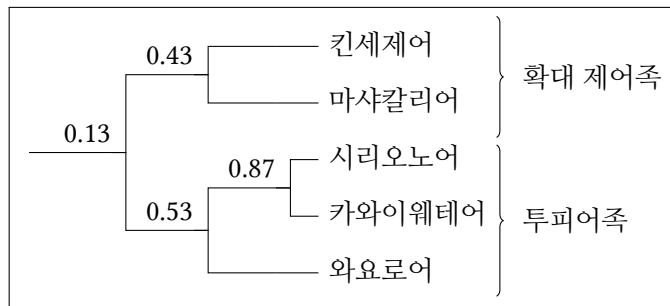
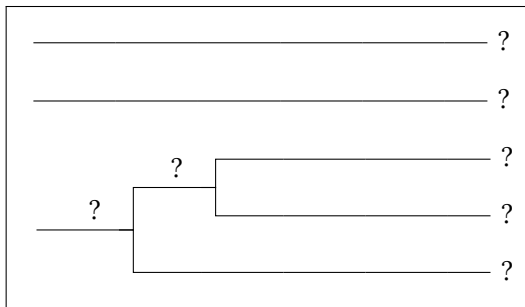
(K) (4 점) A 알고리즘과 B 알고리즘 사이의 차이를 설명하여라.

제 6 부. 확대 제어족, 투피어족 (브라질, 볼리비아)

(L) (28 점) 확대 제어족과 투피어족은 남아메리카의 주요 어족이다. 어떤 언어학자들은 그들이 먼 친족 관계라고 주장한다. 다음 어휘 목록들을 살펴보아라.

	A	B	Γ	Δ	E
나무껍질	e='e-ke	h <sup>w</sup> i='k <sup>h</sup> Λ	kup='pε	mīβm='tεaj	'pε
배 (신체)	'e=rje	=t <sup>h</sup> igi	=ã'ün	'tæj	=rε'wek
피	e='ruki	=ka'nbrɔ	=d̄z='a'ɯ	'hεβ̄p	=ru'i
(불에) 태우다	'rai	=rɔ='k <sup>h</sup> ɔs̄	=po'k <sup>w</sup> a	mũ=...='haβ̄p	=ra'pi
지방	e='kira	=t <sup>h</sup> wəmi	'd̄z=ap	'tuβ̄p	'kap
발 (신체)	'e=i	=h <sup>w</sup> aji	'βi	=pɔ'ta	'pi
손	'e=o	=ɲi'k <sup>h</sup> ɔa	'βo	'ɲiβ̄m	'pɔ
무거운	e='usi	=wi't <sup>h</sup> i	=po'ti	=β̄p'təj	=pɔ'ij
간 (신체)	'e=ja	=n <sup>h</sup> ba	=pi'a	=tεiβ̄pkɪ'nāj	=pi'ʔa
새로운	e='jasu	=n <sup>h</sup> diwi	=pa'gop	'tiβ̄p	=pia'u
뿌리	e='rao	=ja'ɾe	kup=kujɔ'pε	mīβm=ɲiβ̄m=tεa'tiə	=ra'pɔ
피부	'e=i	=k <sup>h</sup> Λ	'pε	'tεaj	'pit
꼬리	e='rokoī	=n <sup>h</sup> bi	=d̄z='o'k <sup>w</sup> aj	=nā:='kiβ̄p	'raj
하얀	'e=f̄i	=ja'k <sup>h</sup> a	=d̄zi'ra	=β̄p'douɥ	's̄iŋ
날개	e='heo	=ja'ɾa	=pε'o	=ɲi'māuɥ	=pε'pɔ, =ji'wa

아래에는 같은 목록들을 기반으로 작성된 수형도가 두 개 있다. 일부 데이터(언어 이름과 어휘통계적 거리)는 숨겨져 있다. 빈칸을 채워라. 각 수형도마다 수동으로 생성되었는지, 자동으로 생성되었는지를 명시하고, A 알고리즘을 이용하여 생성되었는지, B 알고리즘을 이용하여 생성되었는지를 명시하여라.



A	B	Γ	Δ	E
?	?	?	?	?

⚠ 이 문항에서 수동 주석과 안정 지수는 의도적으로 생략되었다.

(M) (10 점) 돌고폴스키의 분류를 기반으로 하여 자동화된 절차는 잘못된 결과를 낼 가능성이 있다. 이 예시에서 자동화된 절차는 시리오노어와 같은 투피어족의 다른 언어들 사이보다 시리오노어와 확대 제어족의 한 언어(킨세제어) 사이에 더 많은 유사성을 감지한다. 위의 확대 제어족과 투피어족의 어휘 목록에 적용할 경우 올바른 결과가 나오는 자동화된 절차를 제안하고 이를 간단히 설명하라.

⚠ 이 문항은 최고 점수를 받은 팀들 간에 동점이 발생한 경우에만 채점됨.

저자들은 특정 언어에 대한 질문에 답변해 준 알레한드라 비달, 마리아 코노센코, 일리야 그룬토프, 얀 토 수야에게 감사를 표합니다. —안드레이 니콜린, 밀레나 베네바

---

편집자: 스타니슬라프 구레비치, 안드레이 니콜린 (편집장), 이반 데르잔스키 (기술편집자), 휴 도브스, 마리아 루빈시테인, 이머 맥나이트, 밀레나 베네바, 엘리시아 워너, 보리스 이웬딘, 알렉세이 페구셰프, 얀 페트르, 알렉산드르 피페르스키.

한국어 텍스트: 이태훈, 김민규.

행운을 빕니다!