

# Lietuviškos kirčiuotos raidės: kodavimas ir įvedimas iš klaviatūros

## Vladas Tumasonis

Vilniaus universiteto Matematikos ir informatikos fakulteto  
Informatikos katedros docentas, daktaras  
Vilnius University, Dept. of Computer Science,  
Assoc. Prof., PhD  
Naugarduko g. 24, LT-03225 Vilnius  
Tel. (+370-5) 219 3074  
El. paštas: vladas.tumasonis@maf.vu.lt

*Straipsnyje pateikiama būtina ir pakankama informacija apie lietuviškų kirčiuotų raidžių realų vartojimą kompiuteriniame tekste – jų kodavimas ir įvedimas standartine lietuviška klaviatūra. Pateikiamas neturinčių kodų kirčiuotų raidžių kodavimas kompozicinėmis sekomis. Aprašomas vadinamasis tęsties ženklas (dead character) metodas, kuriuo galima praplėsti įvedamų ženklų rinkinį. Pasinaudodamas šiuo metodu (ir jį įgyvendinančia tvarkykle), kiekvienas kompiuterio naudotojas (kalbininkas, mokytojas, leidėjas, duomenų bazių kūrėjas ir t. t.) galės įvesti į savo kuriamus dokumentus lietuviškas kirčiuotas raides. Ir netgi prancūziškas, vokiškas ir kitų kalbų raides.*

Kaip žinoma, lietuvių kalba, be pagrindinės abėcėlės, turi kirčiuotas raides. Jos gali būti pavartotos žodynuose, enciklopedijose, vadovėliuose, tekstyvuose, terminologinėse duomenų bazėse ir pan. Kirčiuotų raidžių nėra kompiuterio klaviatūroje, todėl kyla klausimas, kaip jas įvesti. Be to, dalis kirčiuotų raidžių neturi kodų, o tai irgi kelia nemažą sumaištį. Čia pateikiami būdai minėtoms problemoms spręsti. Kiekvienas kompiuterio naudotojas galės įvesti į savo kuriamus dokumentus lietuviškas kirčiuotas raides. Ir netgi prancūziškas, vokiškas, kitų kalbų raides. Svarbu pažymėti, kad pačiai klaviatūrai nereikia jokio papildomo graviravimo. Keičiasi tik klaviatūros išdėstymą įgyvendinanti programa – vadinamoji klaviatūros tvarkyklė.

## Kirčiuotų raidžių kodavimas

Bendrinė lietuvių kalba turi 68 kirčiuotas raides: 34 didžiąsias ir 34 mažąsias. Iš

jų 33 turi standartinius dviejų baitų kodus (Tumasonis, 1999; LST ISO/IEC 10646, 2005) (kartais sakoma *Unikodo* kodus), o 35 – neturi. Neturinčias kodų raides (o kirčiuotos raidės yra sudėtinės raidės, sudarytos iš pamatinės raidės ir diakritinio ženklo) galima išreikšti vadinamąja *kompozicine seka*, kurią sudaro raidė (ji atstovauja pamatinei raidei) ir kombinacinis ženklas (jis atstovauja diakritiniam ženklui – kirčiui). Kombinaciniai ženklai (tiksliau – jų vaizdai) neturi pločio (jie piešiami virš raidės arba po raide), todėl kartais vadinami *nulinio pločio ženklais*. Beje, sekomis galima išreikšti ir turinčias kodus raides.

Kombinaciniai ženklai šio straipsnio tekste vaizduojami iš taškų sudarytu apskritimu, virš kurio yra reikiamas diakritinis ženklas. Kirčiuotoms raidėms koduoti vartojami kombinaciniai ženklai pateikiami 1 lentelėje.

1 lentelė. **Kombinaciniai ženklai**

Ženk- las	Pavadinimas	Pavadinimas pagal LST ISO/ IEC 10646	Kodas pagal LST ISO/IEC 10646
ò	Kombinacinis kairinis kirtis	COMBINING GRAVE ACCENT	0300
ó	Kombinacinis dešininis kirtis	COMBINING ACUTE ACCENT	0301
õ	Kombinacinis riestinis kirtis	COMBINING TILDE	0303
ô	Kombinacinis taškas viršuje	COMBINING DOT ABOVE	0307

Visos kirčiuotos raidės ir jų kodai pateikiami 2 lentelėje.

2 lentelė. **Kirčiuotos raidės ir jų kodai**

Ženklas	Kodas pagal LST ISO/IEC 10646	Kompozicinė žen- klų seka	Kompozicinė kodų seka	Ženklas	Kodas pagal LST ISO/IEC 10646	Kompozicinė ženklų seka	Kompozicinė kodų seka
À	00C0	A ò	0041 0300	Ý	00DD	Y ó	0059 0301
à	00E0	a ò	0061 0300	ý	00FD	y ó	0079 0301
Á	00C1	A ó	0041 0301	ÿ	1EF8	Y õ	0059 0303
á	00E1	a ó	0061 0301	ÿ	1EF9	y õ	0079 0303
Ā	00C3	A õ	0041 0303	Ĳ		J õ	004A 0303
ā	00E3	a õ	0061 0303	ĳ		j õ õ	006A 0307 0303
Ą		Ą ó	0104 0301	Ĺ		L õ	004C 0303
ą		ą ó	0105 0301	ĺ		l õ	006C 0303
Ą		Ą õ	0104 0303	Ļ		M õ	004D 0303
ą		a õ	0105 0303	ļ		m õ	006D 0303
Ē	00C8	E ò	0045 0300	Ñ	00D1	N õ	004E 0303
è	00E8	e ò	0065 0300	ñ	00F1	n õ	007E 0303
É	00C9	E ó	0045 0301	Ò	00D2	O ò	004F 0300
é	00E9	e ó	0065 0301	ò	00F2	o ò	006F 0300
Ĕ	1EBC	E õ	0045 0303	Ó	00D3	O ó	004F 0301
ĕ	1EBD	e õ	0065 0303	ó	00F3	o ó	006F 0301
Ě		Ě ó	0118 0301	Õ	00D5	O õ	004F 0303
ě		ě ó	0119 0301	õ	00F5	o õ	006F 0303
Ě		Ě õ	0118 0303	Ř		R õ	0052 0303
ě		ě õ	0119 0303	ř		r õ	0072 0303
Ě		Ě ó	0116 0301	Ů	00D9	U ò	0055 0300
ě		ě ó	0117 0301	ů	00F9	u ò	0075 0300
Ě		Ě õ	0116 0303	Ú	00DA	U ó	0055 0301
ě		ě õ	0117 0303	ú	00FA	u ó	0075 0301
Ĭ	00CC	I ò	0049 0300	Û	0168	U õ	0055 0303
ì		i ò ò	0069 0307 0300	ü	0169	u õ	0075 0303
Í	00CD	I ó	0049 0301	Ů		U ó	0172 0301
í		i ó ó	0069 0307 0301	ů		u ó	0173 0301
Ī	0128	I õ	0049 0303	Û		U õ	0172 0303
ī		i ó õ	0069 0307 0303	ü		u õ	0173 0303
Ĳ		Ĳ ó	012E 0301	Ů		U ó	016A 0301
ĳ		j ó ó	012F 0307 0301	ů		u ó	016B 0301
Ĺ		Ĺ õ	012E 0303	Ů		U õ	016A 0303
ĺ		j ó õ	012F 0307 0303	ů		u õ	016B 0303

### Kirčiuotų raidžių su taškais ypatumai.

Lietuviškos kirčiuotos mažosios raidės <i>, <į> bei <j> turi vieną ypatumą, susijusį su jų grafiniu vaizdu, t. y. simboliu. Vakarų Europos kalbose raidė <i> su kairiniu kirčiu ir kitais kirčiais vaizduojama be taško, o dabartinės lietuvių kalbos taisyklės reikalauja tašką išlaikyti, plg.:

<î> Vakarų Europos kalbose;

<ï> lietuvių kalboje.

Komponuojant pagal įprastas taisykles (*mažoji raidė i ir kairinis kirtis*), gaunama raidė *i su kairiniu kirčiu* be taško. Norint šios raidės su tašku, jį reikia uždėti specialiai. Toks ženklas Unikodo standarte egzistuoja – tai jau minėtas kombinacinis taškas viršuje, jo kodas 0307. Todėl mūsų tikslams *mažoji raidė i su kairiniu kirčiu* išreiškiama taip: *mažoji raidė i, kombinacinis taškas viršuje* ir *kombinacinis kairinis kirtis*, t. y. trijų ženklų seka <i 0 0>. Kodų seka atitinkamai tokia: 0069 0307 0300.

### Kirčiuotų raidžių įvedimas iš klaviatūros

**Lietuviška kompiuterio klaviatūra (pagal Lietuvos standartą LST 1582).** Klaviatūra yra gana svarbi priemonė ženklams įvesti į kompiuterį. Paprastai klaviatūra galima įvesti tam tikrą ženklų rinkinį, kuriame yra visos nacionalinės raidės, ASCII ženklai ir dar keletas rašte vartojamų ženklų. Pavyzdžiui, lietuviška standartine klaviatūra (LST 1582, 2000), kuri pateikta paveiksle, galima surinkti:

- visus ASCII ženklus;
- visas lietuviškos abėcėlės raides;

- kitus lietuviško rašto ženklus (kabutes, brūkšnį, paragrafo ženklą);
- dešininį kirčio ženklą;
- jungiamąjį tarpą (gaunamas kartu paspaudus trečiojo lygio klavišą *Lyg3* ir tarpo klavišą);
- euro ženklą.

Tai visavertė lietuviška klaviatūra, kurioje yra visi reikalingi ženklai, todėl klaviatūros nereikia perjunginėti.

Klaviatūra yra trijų lygių. Ant klavišo gali būti 2 arba 3 ženklai. Pavyzdžiui, ant klavišo

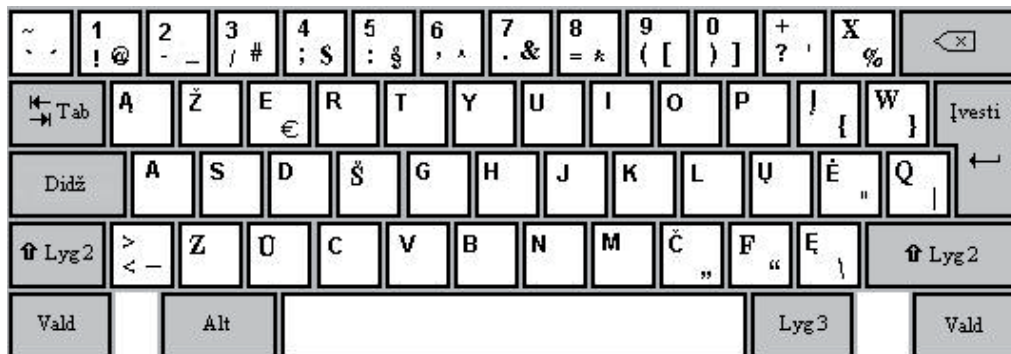


yra trys ženklai. Jie renkami šitaip:

paspaudus šį klavišą, gaunamas kabliataškis <,>, t. y. kairysis apatinis ženklas; nuspaudus *Lyg2* ir paspaudus šį klavišą, gaunamas <4>, t. y. kairysis viršutinis ženklas; nuspaudus *Lyg3* ir paspaudus šį klavišą, gaunamas <\$>, t. y. dešinysis ženklas.

Taigi, klavišų graviravimas atitinka įvedamus ženklus. Įgyvendintas principas „Ką matau, tą gaunu“ (angl. WYSIWYG – *What You See Is What You Get*).

Ši lietuviška klaviatūra sukurta laikantis tarptautinio standarto (ISO/IEC 9995, 1994), reglamentuojančio bendriausius klaviatūros ženklų išdėstymo principus, nuostatų. Tai trys lygiai, skaitmenų išdėstymas viršutinėje eilėje pirmu arba antru lygiu, nacionalinių raidžių nediskriminavimas, įvedamų ženklų rinkinio išsamumas, atitinkamas graviravimas. Nė vieno iš šių principų netenkina Lietuvoje dar naudojama vadinamoji skaitmeninė klaviatūra.



Pav. Lietuviška standartinė klaviatūra

Įprastomis klaviatūromis, kaip buvo minėta, įvedamos nacionalinės raidės ir reikalingi ASCII ženklai. Kirčiuotos raidės į šį rinkinį nepatenka. Todėl reikia kokių nors metodų ir būdų įvedamų ženklų rinkiniui išplėsti.

Šiame straipsnyje teikiamas būdas, kuris remiasi tęsties ženklo (angl. *dead character*) naudojimu. Tęsties ženklas – tai klaviatūros ženklas, kurį surinkus nieko negaunama: jo veikimas pasireiškia surinkus kitą ženklą. Tai pakankamai žinomas tiesioginis klaviatūros galios išplėtimas, pirmą kartą sistemingai pritaikytas lietuviškoms kirčiuotoms raidėms įvesti. Šis būdas nereikalauja nei papildomų klaviatūros fizinių klavišų, nei papildomo graviravimo, nei komplikuoto įsiminimo. Visiškai tinka ta pati fizinė lietuviška klaviatūra. Būdas pakankamai universalus, veikia visose terpėse.

#### Turinčių kodą kirčiuotų raidžių įvedimas.

Iš viso yra trys tęsties ženklai. Tai kairinis, dešininis ir riestinis kirčio ženklai. Šie ženklai turi savo vietas lietuviškoje standartinėje klaviatūroje (t. y. renkami įprasta tvarka). Lengvai įsimenami, nes visi yra viename kairiajame viršutiniame klaviše, t. y. išdėstyti pirmu, antru ir trečiu lygiu atitinkamai. Kirčiuota raidė (o ji yra sudėtinė raidė) gaunama surinkus atitinkamą tęsties ženklą ir po to – reikiamą pamatinę raidę. Pavyzdžiui, raidė <à> (t. y. a-kairinis) gaunama surinkus kairinį kirtį ( ` ) ir raidę a, raidė <â> (a-riestinis) – surinkus riestinį kirtį ( ~ ) ir raidę a, o raidė <á> (a-dešininis) – surinkus dešininį kirtį ( ´ ) ir raidę a. Tai galima užrašyti tokiu formalizmu:

` + a → à

~ + a → â

´ + a → á

Savaime suprantama, kad didžiosios raidės renkamos su didžiąja pamatine raide. Pvz.:

` + A → À

Šitaip (kirtis + raidė) renkamos visos lietuviškos kirčiuotos raidės, turinčios kodus (žr. 2 lentelę). Priminsime, kad atmintyje šios raidės bus vaizduojamos vienu savo kodu. Pvz., raidė à kodu 00E0.

Tęsties ženklai šiek tiek „kainuoja“, nes patys yra užblokuojami (negali būti įvedami

tradiciniu būdu, t. y. vienu klavišo paspaudimu). Todėl jie, kaip savarankiški ženklai, gaunami surinkus po jų tarpą, t. y.:

` + tarpas → ` (0060)

~ + tarpas → ~ (007E)

´ + tarpas → ´ (00B4)

**Neturinčių kodų kirčiuotų raidžių įvedimas.** Neturinčios kodų kirčiuotos raidės išreiškiamos seka, kuri ir turi būti įvesta. Priminsime, kad seką sudaro pamatinė raidė ir kombinacinis ženklas. Kombinacinių ženklų klaviatūroje nėra, todėl ir juos (kaip savarankiškus ženklus) siūloma įvesti pagal tą patį tęsties ženklo metodą:

` + ` → ò

~ + ` → õ

´ + ` → ó

` + . → ô

Tai pakankamai asociatyvus pasirinkimas, leidžiantis lengvai įsiminti reikiamus veiksmus. Tai yra kombinaciniai kirčių ženklai įvedami, surinkus tokį patį atitinkamą paprastą kirtį (kaip tęsties ženklą) ir paskui kairinį kirtį. Tik kombinaciniam taškui taikoma išimtis: deklaravus tašką tęsties ženklu, jis būtų užblokuotas. O taškas, skirtingai nuo kirčio ženklų, yra dažnesnis rašto ženklas.

Iš principo asociatyvia prasme neblogas būtų ir toks variantas, kai kombinaciniai ženklai „užkabinami“ ant vieno ir to paties tęsties ženklo – kairinio kirčio:

` + ` → ò

` + ~ → õ

` + ´ → ó

Bet kai kurios programos (pvz., *Word*) „nemėgsta“, kad antruoju ženklu būtų trečiojo lygio ženklas (o dešininis kirtis toks ir yra). Jis tokiu atveju tiesiog ignoruojamas. Todėl buvo pasirinktas pirmasis variantas.

Taigi, pavyzdžiui, kodo neturinti raidė <ą> yra išreiškiamą kompozicine seka <ą õ>, kuri renkama taip:

ą + ~ + ` → ą õ

Gauta seka vizualizuojama kaip ą. Atmintyje bus abiejų sekos dėmenų kodai 0105 0303.

O <î>, kurią išreiškia trijų ženklų seka <î ò õ>, renkama taip:

i + ` + . + ` + ` → i ô ò

Gauta seka vizualizuojama kaip ì. Atmintyje bus tokie kodai: 0069 0307 0300.

Analogiškai renkamos visos raidės, neturinčios savo kodų.

Reikėtų pasakyti, kad naudotojui nėra reikalo žinoti kodų: juk įvedant paprastas raides, kodai nedomina. Reikia tik aiškiai suvokti (ar įsivaizduoti), iš kokių komponentų sudaroma sudėtinė raidė, ir juos įvesti.

**Prancūziškų, vokiškų ir kitų kalbų raidžių įvedimas.** Be minėtų trijų tęsties ženklų, šių kalbų raidėms įvesti vartojami dar du tęsties ženklai: stogelio ženklas ( ^ ), kuris asocijuojasi su stogelio diakritiniu ženklu, ir paprastos kabutės ( “ ), kurios asocijuojasi su umliauto ženklu. Sudėtinės raidės renkamos analogiškai. Pavyzdžiui:

^ + a → â

“ + a → ä

Toliau pateikiamas visų įvedamų turinčių kodą raidžių rinkinys (įskaitant lietuviškas kirčiuotas).

à è ò ù	ñ ý ù	ç
À È Ì Ò Ù	Ñ Ý ù	Ç
á é ý ó ú	ć ġ k l m n p r s w z	
Á É Í Ý Ó Ú	Ć Ğ K L M N P R S W Z	
ã ë ÿ ñ õ ü	ÿ	
Ã Ë Ĩ Ÿ Ñ Õ Ü	ÿ	
â ĉ ê ĝ ĥ î j ô š ū ŵ ŷ ž	ä ë ĥ i ö ü ŵ x ŷ	
Â Ć Ê Ğ Ĥ Î J Ô Š Ū Ŵ Ŷ Ž	Ä È Ħ İ Ö Ü Ŵ X Ŷ	

*Pastaba.* Raidės ç ir Ç gaunamos su kairinio kirčio tęsties ženklu:

` + c → ç

` + C → Ç

Patys umliauto ir stogelio kombinaciniai ženklai (jų gali prirėkti kad ir dialektologui, norint gauti kokią nors sudėtinę raidę su šiais diakritikais, neturinčią *Unikodo* kodo) įvedami pagal tą pačią schemą:

^ + ` → ô (0302)

“ + ` → ö (0308)

## Šriftai ir *OpenType* keitiniai

Naudojant kombinacinius ženklus, iš esmės galima konstruoti bet kokius sudėtinius ženklus. Savaime suprantama, kad naudojamame šrifte turi būti reikiami kombinaciniai ženklai (jų mums reikia keturių). Ne visi šriftai juos turi. Pavyzdžiui, populiarūs šriftai *Times New Roman* ir *Arial* neturi kombinacinio taško viršuje (0307).

Ir dar viena – ženklų vizualizavimo, t. y. vaizdavimo ekrane ar popieriuje, – problema. Mažosios ir didžiosios raidės yra skirtingo aukščio. Todėl kirtis turėtų atitinkamai keisti savo padėtį. Pavyzdžiui, raidžių ř ir Ŗ riestiniai kirčiai turėtų būti skirtingo aukščio. Tačiau kombinacinio kirčio vaizdas ir padėtis yra fiksuoti šrifte (kaip ir kiekvieno kito ženklo). Todėl vaizdavimo rezultatas bus toks: ř ir Ŗ. Visgi egzistuoja specialus būdas (vadinamieji *OpenType* technologijos keitiniai), leidžiantis gražiai pavaizduoti šias raides. Raidės su skirtingo aukščio kirčiais (mūsų pavyzdyje ř ir Ŗ) iš anksto nupiešiamos ir įtraukiamos į šriftą.

Vėliau, vaizduojant ženklą ekrane arba spausdinant, „negrąžios“ sekos ř ir Ŗ keičiamos į tuos anksčiau nupieštus vaizdus ř ir Ŗ.

Pasakytina, kad dar nedaug terpių įgyvendina *OpenType* keitinius. Kol kas galima pažymėti tik *Word* 2003 ir 2007, *Internet Explorer* ir maketavimo sistemą *InDesign*. Be to, ir pačiuose šriftuose gali nebūti keitinių. Neseniai sukurtas liuanistinis šriftas *Palemonas* (Aleknavičienė ir kt., 2005)

turi visus reikiamus kombinacinius ženklus ir keitinius. Tai, atrodo, kol kas vienintelis šriftas Lietuvoje, tinkamas veiksams su kirčiuotomis lietuviškomis raidėmis, prancūziškomis, vokiškomis ir kitų kalbų sudėtinėmis raidėmis.

## Tvarkyklės

Klaviatūros tvarkyklės, sukurtas autoriaus ir įgyvendinančias straipsnyje išdėstytas idėjas minėtoms sudėtinėms raidėms koduoti ir įvesti, galima rasti Valstybinės lietuvių kalbos

komisijos tinklalapyje <http://www.vlkk.lt/palemonas.html>. Į paketą įeina tvarkyklės aprašas, diegimo ir naudojimo instrukcija. Ten pateikta ir tvarkyklė tam atvejui, kai neturinčios kodų

lietuviškos kirčiutos raidės koduojamos vadinamojoje *Privačiojoje srityje*. Iš čia galima taip pat atsisiųsti ir *Palemono* šriftą.

## LITERATŪRA

ALEKNEVIČIENĖ, O.; GRUMADIENĖ, L.; GURSKAS, A.; SKIRMANTAS, P.; STROCKIS, M.; TUMASONIS, V. (2005) Lituanistinis šriftas *Palemonas*. Vilnius: Tautinių bendrijų namai. 302 p.

ISO/IEC 9995 (1994) *Information technology – Keyboard layouts for text and office systems*.

LST 1582 (2000) *Informacijos technologija. Lietuviška kompiuterio klaviatūra. Ženklių išdėstymas*.

LST ISO/IEC 10646 (2005) *Informacijos technologija. Universalus kelias baitais koduotų ženklų rinkinys (tapatus ISO/IEC 10646:2003)*.

TUMASONIS V. (1999) Encoding of Lithuanian Accented Letters. Iš *Multilinguale Corpora: Encoding, Structuring, Analysis*. Proceedings of the Conference, Frankfurt a/Main, p. 15–20.

## LITHUANIAN ACCENTED LETTERS: ENCODING AND INPUT FROM THE KEYBOARD

Vladas Tumasonis

### Summary

In standard modern Lithuanian there are three accent (stress) signs: grave ( ` ), acute ( ´ ), and circumflex ( ~ ). All in all there are 68 (34 lower case and 34 upper case) Lithuanian accented letters. Only 33 letters have Unicode codes. Encoding of Lithuanian accented letters and its input from the keyboard are presented in this article. Codeless letters are expressed by composite sequences: base letter and combining accent character (grave, acute, tilde and sometimes dot

above). Usual keyboards have no keys for accented letters. The input extension method (dead character method) for Lithuanian standard computer keyboard is proposed as well as the corresponding driver. So computer user (a linguist, a teacher, a publisher, a data base creator etc.) with Lithuanian standard keyboard will be able in very suitable manner to enter Lithuanian accented letters as well as French, German and other languages letters.