

Hiperteksto rašyklių palyginimas lokalizavimo galimybių požiūriu

Rimgaudas LAUCIUS (MII)

el. paštas: rimga@ktl.mii.lt

Ižanga

Globalizacijos ir kultūrų asimiliacijos kontekste mažoms valstybėms itin aktuali nacionalinės kalbos ir kultūros puoselėjimo problema. Lietuvos informacinės visuomenės plėtros strateginiame plane vienas iš keturių prioritetų yra „lietuvių kalbos ir kultūros paveldo puoselėjimas panaudojant informacines technologijas ir telekomunikacijas“ (IVPK, 2001). Šiam tikslui įgyvendinti vienas iš svarbiausių šiandienos darbų yra programinės įrangos (PĮ) lokalizavimas.

Tai, kaip šis tikslas bus įgyvendintas, priklauso ir nuo mūsų pačių iniciatyvos. Džiugu, kad šiuo metu internete galima rasti nemažai sulietuvintos PĮ. Didelė jos dalis yra sulietuvinta be jokio atlygio, o tik iš idėjos. Deja, tačiau nemaža jos dalis yra sulietuvinta neišbaigtai arba nekorektiškai. Kaip vieną to priežasčių, galima paminėti lokalizavimo žinių bei metodinės literatūros šia tema trūkumą. Todėl šiame straipsnyje pristatomas vienas iš būdų kaip įvertinti, ar pasirinkta PĮ gali būti pilnai lokalizuota ir jį taikant iširta 21 nemokamai platinama hiperteksto rašyklė.

PĮ lokalizavimo darbai

Tam, kad geriau suvoktume PĮ lokalizavimo procesą jį bendrais bruožais apibūdinsime, išskirdami ir paaiškindami jo etapus.

Bendru atveju išskiriami trys PĮ lokalizavimo darbai: adaptavimas lokalės normoms, sąsajos teksto vertimas ir žinytų vertimas (1 pav.). Tik pilnai atlikus šiuos tris darbus galime teigti, kad PĮ yra pilnai lokalizuota [1].



1 pav. PĮ lokalizavimo darbai.

PĮ lokalizavimo darbų pagrindą sudaro adaptavimas lokalės normoms. Lokalė – tai „naudotojo aplinkos poaibio, priklausančio nuo kalbos ir kultūros normų, apibrėžimas“ (LST ISO/IEC 15897). Ji apima ženklų kodavimą, klasifikaciją ir rikiavimą, skaičių, datų ir piniginių vienetų formatus, bei daugelį kitų nuo kultūrinės terpės priklausomų nuostatų. Tačiau šis terminas gali būti naudojamas ir šiek tiek platesne prasme – pavyzdžiui, pažymint vietovę, kurioje šios nuostatos galioja.

Bendru atveju tik adaptavus PĮ lokalės normoms galima atlikti sąsajos teksto bei žinytų vertimą. Mat lokalės normoms neadaptuota PĮ gali nepalaikyti lokalės ženklų ir neteisingai rodyti išverstą tekstą.

Kad šiuos darbus būtų galima sklandžiai ir išbaigtai atlikti, tam PĮ kūrėjo turi būti sudarytos galimybės – PĮ pilnai internacionalizuota. Nuo gamintojo pasirinkto internacionalizacijos modelio taip pat priklauso koku iš dviejų būdų ji gali būti lokalizuojama: pradinio kodo lygyje arba dvejetainiame lygyje. Kadangi pirmasis internacionalizacijos modelis laikomas pasenusiu ir nebuvo sutiktas tiriant hiperteksto rašykles panagrinėsime tik antrąjį.

PĮ lokalizavimo etapai

PĮ lokalizavimo darbus bendru atveju galima suskirstyti į tris etapus:

1. Lokalizuojamų išteklių išskyrimas.
2. Lokalizuojamų išteklių adaptavimas lokalės normoms ir vertimas.
3. Lokalizuojamų išteklių įkėlimas.

Lokalizuojami ištekliai tai PĮ naudojami paveikslai, piktogramos, teksto eilutės, dialogų šablonai, dokumentų šablonai, informacija apie lokalę ir kitokie ištekliai, kuriuos reikia adaptuoti lokalės normoms arba išvesti. Kuriant PĮ šie ištekliai yra įkompiliuojami į vykdomą programą arba prie jos jungiamas dinamines bibliotekas. Tam, kad juos vėliau būtų galima išskirti, jie turi būti įkompiliuoti į šiose bylose specialiai tam skirtas išteklių sekcijas.

Lokalizuojant išteklius reikia dekompiliuoti arba kitaip sakant išskirti. Tam gali būti naudojamos įvairios specializuotos priemonės – pavyzdžiui: „RLTools“, „Resource Hacker“, „Resource Workshop“ ir kt.

Kuriant PĮ dauguma programavimo sistemų automatiškai sukuria išteklių sekcijas ir perkelia į jas dalį lokalizuojamų išteklių. Tačiau, kad būtų perkelti visi lokalizuojami ištekliai turi pasirūpinti programuotojas – pavyzdžiui, programoje jis turi nurodyti kad visos lokalizuojamo teksto eilutės kompiliuojant būtų padėtos į išteklių sekcijas.

Išskirti ištekliai yra adaptuojami lokalės normoms ir verčiami. Adaptuojant PĮ lokalės normoms pirmiausia pakeičiama informacija apie lokalę. Pagrindinė šios informacijos dalis yra lokalės kodas, identifikuojantis konkrečią lokalę. Paprastai tai yra skaičius, atitinkantis numatytą lokalę identifikuojančią konstantą – pavyzdžiui, „Windows“ programose lietuvių lokalę identifikuoja skaičius 427₍₁₆₎ [5]. Tą nepadarius PĮ gali neteisingai rodyti išverstą sąsajos tekstą.

Adaptuojant PĮ lokalės normoms taip pat pakeičiami lokalės nuostatų neatitinkantys paveikslai, piktogramos. Pavyzdžiui, tai gali būti paveikslai su rankų gestais arba kitokiu

turiniu, kurie konkrečioje lokalėje turi kitokią prasmę nei siekiama perteikti. Priklausomai nuo PĮ taip pat gali tekti pakeisti lokalės normų neatitinkančius dokumentų, datų, skaičių, piniginių sumų ir kitokių elementų formavimui naudojamus šablonus.

Verčiant sąsajos tekstą, dalis išverstų eilučių gali netilpti į joms paskirtą vietą. Šią problemą galima išspręsti dviem būdais: pataisant ištekliuose esančius dialogų šablonus arba kaip sutrumpinant teksto eilutę (pvz., išsireiškiant lakoniškiau).

Lokalizuoti ištekliai įkeliami atgal. Tam paprastai gali būti naudojamos tos pačios priemonės kaip ir jų išskyrimui, arba išteklių kompiliatoriai pateikiami PĮ kūrimo priemonių paketuose (pvz., „Delphi“, „Visual C++“).

Bendras hiperteksto rašyklių internacionalizacijos lygio įvertinimas

Atlikto tyrimo objektu pasirinktos nemokamai platinamos hiperteksto rašyklės. Norėdami įvertinti jų internacionalizacijos lygį pirmiausia apžvelgsime jų lokalizavimo galimybes bendrais bruožais, siekdami atrinkti programas, kurios apskritai gali būti lokalizuojamos. Vėliau atrinktas programas išnagrinėsime detaliau bandydami rasti jų internacionalizacijos trūkumų.

Apžvelgdami programas bendrais bruožais nustatysime, ar galima išskirti lokalizuojamus išteklius: paveikslus ir piktogramas, sąsajos teksto eilutes, dialogų šablonus ir kaip jie užkoduoti. Rezultatai pateikti 1 lentelėje. 4 pirmosiose lentelės eilutėse nurodytos programos atitiko iškeltus reikalavimus.

1 lentelė. Hiperteksto rašyklių bendros lokalizavimo galimybės

Programos pavadinimas	Paveikslai, piktogramos	Sąsajos teksto eilutės	Dialogų šablonai
„AceHTML 5“, „HTML-Kit 5.0“	Vykdomoji byla	Vykd. ir tekstinė byla	DFM
„Amaya 8.0“	Vykdomoji byla	Tekstinė byla	WINAPI
„Arachnophilia 4.0“	Vykdomoji byla	Vykdomoji byla	WINAPI
„WebCoder 3.00“	Vykdomoji byla	Vykdomoji byla	DFM
„BPlainPro 2.2.6“, „EasyHTML 2.2“, „EzPad 3.0“, „HomeSite 1.2“, „HTML Beauty 1.5“, „HTML Buddy 1.2“, „HTML Builder 5.6“, „Webniac 1.4“, „Web Notepad 1.02“, „WebZard HTML 1.0“	Vykdomoji byla	Nepilnai	DFM
„Web Dwarf 2.00“	Vykdomoji byla	Nepilnai	WINAPI
„Abacus 1.0.3“	Vykdomoji byla	Negalima	Negalima
„JmkHtml 4.01“, „Web Crafter 4.0“, „Web Edit 2002“, „Web-Studio 3.0“	Negalima	Negalima	Negalima

„Vykdomoji byla“ – ištekliai įkompiluoti į vykdomosios bylos išteklių sekcija.

„Tekstinė byla“ – ištekliai iškelti į išorinę tekstinę bylą.

„Nepilnai“ – į išteklių sekcijas įkompilautos ne visos išteklių eilutės.

„WINAPI“ – „Windows“ API stiliaus šablonai.

„DFM“ – „Delphi form“ stiliaus šablonai.

„Negalima“ – išteklių išskirti negalima.

Šis įvertinimas buvo atliekamas tokiais etapais:

1. Nustatyta koku būdu pateikiami lokalizuojami ištekliai.
2. Išskirti lokalizuojami ištekliai.
3. Nustatyta, ar pavyko išskirti visus lokalizuojamus išteklius.
4. Nustatyta, koku būdu koduoti dialogo langų šablonai.

Dauguma PĮ kūrimo priemonių automatiškai į išteklių sekciją padeda paveikslus, piktogramas, dialogų šablonus, todėl problemų paprastai būna tik su teksto eilutėmis. Paprasčiausias būdas įsitikinti jog išskirtos ne visos eilutės, patikrinti ar ištekliuose yra eilutė kuria klausiama ar išsaugoti pakeitimus uždarant programą.

Kaip matyti iš pateiktos lentelės tik 5 programos iš 21 tenkina pagrindinius internacionalizacijos reikalavimus. Viena to priežasčių PĮ internacionalizavimo žinių ir patirties trūkumas. Šią hipotezę patvirtina gilesnis hiperteksto rašyklių tyrimas. Kaip pavyzdį pagnagrinėjime programą „EasyHTML 2.2“. Joje iš dalies įdiegtos dvi kalbos: suomių ir anglų, tai rodo jog programuotojas siekė sukurti PĮ, palaikančią daugiau nei vieną lokale. Programa parašyta naudojant „Delphi“ sistemą, kuri turi nesudėtingas priemones, bei galimybes skirtas pilnam PĮ internacionalizavimui. Tačiau jomis nebuvo pasinaudota, o pasirinktas sudėtingesnis ir didelių trūkumų turintis būdas: abiejų kalbų sąsajos teksto eilutės įkompiliuotos į programą kaip konstantos. Pagrindiniai šio būdo trūkumai:

1. PĮ negalima papildyti naujomis lokalėmis.
2. Bet kokie sąsajos teksto pakeitimai reikalauja pilno PĮ perkompiliavimo.

PĮ internacionalizavimo žinių trūkumą taip pat rodo tai, jog visose programose lokalizuojami ištekliai įkompiliuoti į vykdomąją bylą, o ne dinamines bibliotekas, atskiras kiekvienai lokalei. Pagrindiniai šio būdo trūkumai:

1. PĮ apribota viena lokale, kuri negali būti pakeista neįdiegiant kitai lokalei skirto PĮ paketo.
2. Sudėtinga papildyti PĮ naujomis lokalėmis, jas atnaujinti.

Pažymėtina, kad nepakankamos PĮ internacionalizacijos žinių trūkumo priežastimi dažnai tampa, išankstinis nusistatymas, kad PĮ internacionalizavimas yra sudėtingas arba, kad tai kažkieno kito – pavyzdžiui, lokalizuotojo problema.

Detalus hiperteksto rašyklių internacionalizacijos lygio įvertinimas

Bendrus internacionalizacijos reikalavimus tenkinančias hiperteksto rašykles išnagrinėsimė detaliau.

Tam buvo nustatyta, ar nagrinėjamos hiperteksto rašyklės:

- 1) teisingai koduoja bylų vardus, turinčius lietuviškų raidžių,
- 2) teisingai rodo ir koduoja lietuvišką tekstą,
- 3) teisingai vykdo lietuviškų raidžių turinčių žodžių paiešką,
- 4) teisingai įterpia datą ir laiką,
- 5) leidžia įvesti visus klaviatūros išdėstyme numatytus ženklus dirbant su lietuvišką standartą LST 1582:2000 atitinkančia klaviatūra,
- 6) turi kitokių nuo lokalės priklausomų savybių ir ar jas galima adaptuoti konkrečiai lokalei,

7) ar galima pašalinti pirmiau paminėtus internacionalizacijos trūkumus, jei PI jų turi. Paminėsime kiekvienos hiperteksto rašyklės pranašumus ir trūkumus lokalizavimo galimybių atžvilgiu atskirai.

„**AceHTML 5**“. Ši rašyklė vienintelė palaiko unikodą ir leidžia pasirinkti, kuria kodavimo lentele koduotą tekstą rodyti. Įvedus ženklą, kurio nurodytoje kodų lentelėje nėra jį pakeičia į skaitmeninį kodą, todėl yra išsaugomi ir naršyklėje teisingai rodomi visi įvesti ženklai. Programa palaiko Windows 1257 kodavimo lentele, tačiau pasigesta ISO 8859-13 kodų lentelės.

„**Amaya 8.0**“. Neteisingai atlieka žodžių, turinčių lietuviškų raidžių paiešką – nekonvertuoja didžiųjų ir mažųjų lietuviškų raidžių.

„**Arachnophilia 4.0**“. Programa neteisingai įterpia datą ir laiką. Jie įterpiami neteisingu formatu ir su angliškais pavadinimais, kurių išversti negalima.

„**HTML-Kit 5.0**“. Į programą įeina papildoma priemonė, kuri leidžia rašyti unikodu koduotą tekstą, tačiau pati rašyklė jį rodo neteisingai.

Neteisingai atlieka ir žodžių turinčių lietuviškų raidžių paiešką – nekonvertuoja didžiųjų ir mažųjų lietuviškų raidžių.

Data ir laikas įterpiami teisingai, programa turi plačią gamą jų formatų pasirinkimo. Šie formatai imami iš operacinės sistemos lokalės.

Lietuviška standartine klaviatūra negalima įvesti atidaranciu kabučių. Vietoj to yra įterpiama numatyta hiperteksto gairė. Kaip šį trūkumą pašalinti sprendimo rasti nepavyko.

„**WebCoder 3.00**“. Su lokale susijusių trūkumų šioje programoje nepastebėta. Bazineje jos versijoje neteisingai įterpiama data ir laikas, tačiau jie formuojami pagal šablonus, kuriuos galima pakeisti.

Išvados

1. Tyrimas parodė, kad daugumoje nagrinėtų hiperteksto rašyklių nėra numatytos lokalizavimo galimybės, arba jos nepakankamos.
2. Tyrimo metu programų internacionalizacijai iškeltus reikalavimus geriausiai atitiko „AceHTML 5“ ir „WebCoder 3.00“ programos.

Literatūra

- [1] G. Grigas, Interneto programų paketo lietuvinimo patirtis, in: *Informacinės technologijos*, Konferencijos pranešimų medžiaga, KTU, Kaunas (2002).
- [2] IVPK, 2002–2004 metų strateginis veiklos planas (2001).
http://www.ivpk.lt/teises/strateginis_planas02_04.pdf
- [3] LST 1582:2000 Informacijos technologija, Lietuviška kompiuterio klaviatūra, Ženklų išdėstymas.
- [4] LST ISO/IEC 15897 Informacijos technologija, Kultūros elementų registravimo procedūros.
- [5] Win32 SDK (2002). <http://msdn.microsoft.com>

The comparison of localization features of hypertext editors

R. Laucius

There are results of investigation accomplished in comparison of freely distributed hypertext editors localization problems presented in this paper. The aims of this investigation are: to evaluate freely distributed hypertext editors internationalization level, find out their localization problems and their causes.