

Zapis večjezičnih besedil v e-sperantu

Sanida Omerović, Grega Jakus, Tatjana Filimonova in Sašo Tomažič

Fakulteta za elektrotehniko, Tržaška 25, 1000 Ljubljana, Slovenija

E-pošta: sanida.omerovic@lkn1.fe.uni-lj.si, grega.jakus@guest.arnes.si, tatjana.filimonova@fe.uni-lj.si, saso.tomazic@fe.uni-lj.si

Povzetek. V prispevku je predstavljena zasnova formalnega računalniškega jezika e-speranto, ki služi kot formalen zapis besedil, namenjenih večjezičnim skupinam bralcev, obenem pa je lahko vmesni jezik pri avtomatskem prevajanju besedil med različnimi naravnimi jeziki. Jezik e-speranto je zasnovan tako, da je prijazen računalniku, a hkrati še vedno razumljiv človeku. Zadnje je pomembno predvsem pri razvoju aplikacij, povezanih z e-sperantom. Zato njegova skladnja temelji na razširljivem označevalnem jeziku XML (eXtensible Markup Language) [1]. Temeljni konstrukcijski element e-speranta je stavek. Stavek v e-sperantu je semantična enota, ki ustreza povedi v naravnem jeziku. V nasprotju z naravnimi jeziki bo pomen stavkov v e-sperantu enomen, njegova slovnica pa brez izjem in izražena eksplicitno v obliki XML atributov, pomen in stil sta v največji meri ločena. Stilske posebnosti se izražajo v obliki tako imenovanih prevajalskih namigov. Razvoj e-speranta bo potekal v štirih fazah: 1. faza: razvoj skladnje, slovnice in temeljnega besednjaka; 2. faza: razvoj orodij za pisanje besedil neposredno v e-sperantu; 3. faza: razvoj prevajalnikov za prevajanje iz e-speranta v različne naravne jezike; in 4. faza: razvoj prevajalnikov za prevajanje iz naravnih jezikov v e-speranto, kar bo na koncu omogočilo tudi avtomatsko prevajanje med različnimi naravnimi jeziki prek vmesnega jezika e-speranto.

Ključne besede: e-speranto, formalni jezik, naravni jezik, prevajanje, večjezični dokumenti.

Multilingual documents in e-speranto

Extended abstract. This paper presents the initial design of a formal computer language e-speranto. E-speranto is intended for writing and storing multilingual e-documents and, in its final stage, to serve as an intermediate language in automatic multilingual translation processes. The language is designed to be computer friendly but still understandable to humans; the latter being important mainly for developers of different applications connected to e-speranto. For this reason the syntax of e-speranto is based on eXtensible Markup Language (XML) [1]. The basic building block of e-speranto is the sentence. The sentence in e-speranto is a semantic unit which corresponds to the sentence in natural language, but opposed to natural languages, has unambiguous meaning. The meaning and the style of

the sentence in e-speranto are distinguished as much as possible. The style is expressed in a form of the so called translation hints. The grammar has no exceptions and it is expressed explicitly via XML attributes.

The development of e-speranto consists of the following four phases:

1. development of the syntax, grammar and basic dictionary,
2. development of tools for writing documents in e-speranto,
3. development of interpreters of e-speranto documents in different natural languages, and
4. development of automatic translators from different natural languages to e-speranto.

Because of the syntax based on XML, and therefore compatibility with the format of Web pages, e-speranto could be applied in the domain of World Wide Web. After the completion of the 4th phase, a multilingual translation would be possible. Therefore, e-speranto could introduce multilingual Web browsing.

Keywords: language, multilingual, translation, sentence.

To delo je opravljeno v okviru raziskovalnega programa "Algoritmi in optimizacijski postopki v telekomunikacijah" in projekta "Zasnova sistema za večjezično prevajanje besedil".

Prejet 12. marec, 2007
Odobren 29. marec, 2007

1 Uvod

Množična uporaba svetovnega spleta in globalizacija sta zabrisali meje med različnimi kulturami in narodi. Na svetovnem spletu je mogoče najti skoraj vsako informacijo, vendar pri tem še vedno obstajajo jezikovne ovire. Večina svetovnega prebivalstva govori samo en jezik, povprečno izobražen človek na zahodu pa razume še enega ali dva tuja jezika.

Strani na svetovnem spletu so sicer pogosto dostopne v več jezikih, vendar so to najpogosteje, poleg izvirnega jezika strani, le veliki svetovni jeziki, kot so angleščina, španščina, francoščina in nemščina. Na voljo je tudi kar nekaj prevajalnikov, ki v realnem času prevajajo spletne strani iz enega v drug jezik, vendar so tudi ti ponavadi omejeni na velike svetovne jezike. Orodja za avtomatsko prevajanje, ki so trenutno dostopna na trgu, omogočajo le prevajanje med posameznimi pari jezikov, kot na primer angleščina/francoščina, španščina/francoščina ipd. [2] [3] [4] [5]. Modul za vsak par jezikov je treba kupiti ločeno.

Trenutni prevajalniki delujejo dokaj solidno, kar pomeni, da je večinoma mogoče pravilno razumeti prevedeno besedilo že po prvem prevajanju, vendar so še daleč od popolnosti, saj se pri večkratnem prevajanju tja in nazaj pomen popolnoma spremeni ali pa se izgubi.

Za besedila, ki so namenjena večjezični uporabi, kot so na primer večjezične spletne strani, je zato potreben jezik, ki bi omogočal zapis večjezičnih besedil in njihovo avtomatsko interpretacijo v različnih naravnih jezikih. V primeru spletnih strani bi se morali interpretacija in prevajanje izvajati v realnem času. Podobni vlogi je bil namenjen Esperanto [9], ki naj bi ga poznali vsi in ki bi bil namenjen sporazumevanju med ljudmi z različnimi maternimi jeziki. To vlogo je zdaj bolj ali manj uspešno prevzela angleščina. Ker je angleščina naravni jezik, računalniku ni prijazna in zato ne more najbolje služiti kot jezik za zapis večjezičnih e-dokumentov (oziroma kot vmesni jezik za avtomatsko prevajanje med dvema jezikoma).

Predlagamo, da se uvede nov, standardiziran formalni računalniški jezik. Da bi lahko služil opisanemu namenu, mora tak jezik imeti naslednje značilnosti:

1. Računalniku mora biti razumljiv, kar pomeni, da ga je mogoče avtomatsko razčlenjevati. To je pomembno pri avtomatski obdelavi dokumentov, kar pomeni pri prevajanju v ta jezik in iz tega jezika.
2. Biti mora razumljiv človeku. To je pomembno predvsem pri razvoju različnih

aplikacij, obenem pa omogoča pisanje in preverjanje e-besedil neposredno v tem jeziku.

3. Jezik mora enoumno izražati pomen. Zato mora imeti vsaka beseda v slovarju en sam pomen, skladnja (sintaksa) jezika pa mora biti nedvoumna.
4. Pomen mora biti ločen od stila. V naravnem jeziku je namreč mogoče isti pomen izraziti z različnimi stili. Ločevanje pomena od stila omogoča boljše ohranjanje pomena. Stil je pri tem podan v namigih za prevajanje (sleng, besedna igra, fraza ...), tako da je le od »spretnosti prevajalca« odvisno, kako dober bo prevod stilsko. Pomen se mora v celoti ohraniti, tudi če povsem zanemarimo vse namige za prevajanje.

Na podlagi zgoraj opisanih zahtev smo naredili zasnovo formalnega računalniškega jezika za zapis večjezičnih dokumentov. Jezik smo poimenovali e-speranto. Ime pove, da je osnovna ideja povzeta po Esperantu, obenem pa ločeni »e« pomeni, da je namenjen za zapis elektronskih dokumentov.

Da bi bil e-speranto razumljiv tako računalniku kot človeku, smo kot osnovno skladnjo jezika izbrali razširljiv označevalni jezik XML (eXtensible Markup Language) [1]. Slovnica jezika so povzeta po Esperantu, saj je to jezik, ki ne pozna izjem. Vendar je v nasprotju z Esperantom, v e-sperantu slovnica izražena eksplicitno, v obliki metapodatkov, to je podatkov o podatkih, kar je za računalniško obdelavo primernejše od implicitnega označevanja s predponami, končnicami, ločili in podobno.

V drugem razdelku je najprej podan kratek pregled slovnice Esperanta, ki je bila osnova za slovnico e-speranta. Zasnova slovnice e-speranta je predstavljena v tretjem razdelku, njegova XML skladnja pa v četrtem. V petem razdelku je predstavljeno prvo orodje za pomoč pri izdelavi e-speranto dokumentov, izdelano v okolju Eclipse. Ker je e-speranto šele v začetni fazi razvoja, se bo še precej spreminjal in nadgrajeval. V šestem razdelku so predstavljene predvidene faze njegovega razvoja.

2 Kratek pregled slovnice Esperanta

Osnovna ideja e-speranta izvira iz Esperanta, umetnega jezika, namenjenega sporazumevanju med ljudmi, ki govorijo različne materne jezike. Ustanovitelj jezika je bil L. L. Zamenhof. Njegov glavni cilj je bil ustvariti preprost in fleksibilen jezik, ki bi bil univerzalni drugi jezik.

Kot teoretski koncept ima Esperanto zagotovo izreden potencial, toda v praksi se ni uveljavil, saj je izgubil globalno jezikovno bitko proti angleščini.

Slovnica Esperanta je bila odličen vodnik pri definiciji slovnice formalnega računalniškega jezika e-speranta. Temelj razmeroma preproste slovnice Esperanta sestavlja šestnajst pravil za: člene, samostalnike, pridevnike, števnike, zaimke, glagole, prislove, predloge, izgovarjavo, naglas, sestavljene besede, zanikanje, vezavo, večnamenske predloge, medmete in elizije. Slovnica se v Esperantu izraža implicitno, bodisi z uporabo posameznih besednih pripon (kot pri definiciji glagolskega načina: nedoločnik, povednik, pogojnik ali velelnik) bodisi z uporabo posebnih besednih vrst (tvorjenje vprašalnih stavkov: besedice »cxu« za da-ne vprašanja, in besedice na »ki-« kiu, kio, kia, kie, kiam, kies, kiel, kial, kiom za oziralne vprašalnice) [7]. Za boljše ponazoritev si oglejmo končnice, ki se v Esperantu uporabljajo za razlikovanje besednih vrst in so predstavljene v tabeli 1.

končnica	besedna vrsta
-a	pridevnik
-e	prislov
-o	samostalnik

Tabela 1: Definicije besednih vrst v Esperantu
Table 1. Esperanto type of word definition

Slovnici pravila Esperanta smo skoraj v celoti preslikali v e-speranto, le da so tu eksplicitno izražena v obliki oznak XML in atributov XML.

Med osnovnimi idejami na začetku razvoja e-speranta je bila tudi ideja, da bi poleg osnov slovnice uporabili tudi besednjak Esperanta. To idejo smo opustili predvsem zato, ker se Esperanto aktivno ne uporablja in ima zato precej omejen besednjak, ki se le deloma dopolnjuje s sodobnimi izrazi. Zato besednjak e-speranta ostaja odprto vprašanje, najverjetneje pa bo temeljil na angleščini.

Bistvena ovira pri uvajanju Esperanta je bila, da bi se ga morali uporabniki naučiti, kar pa se ne izplača, dokler ni dosežena neka kritična masa, to je dovolj ljudi, ki bi govorili Esperanto. Ravno v tem je bistvena prednost e-speranta, saj se ga uporabnikom ni treba učiti. Potrebujemo le aplikacijo, ki zna e-speranto tolmačiti (interpretirati) v uporabnikovem jeziku. Da pa bi se ustrezne aplikacije razvile, je seveda potrebno, da je jezik standardiziran.

3 Slovnica e-speranta

Slovnica e-speranta temelji na slovnici Esperanta, z razliko, da so tu slovnici pravila izražena eksplicitno v obliki meta podatkov. Temeljni gradbeni element besedila v e-sperantu je stavek. Stavek v e-sperantu je semantična enota, ki ustreza

povedi v naravnem jeziku. Na splošno so stavki sestavljeni iz enega ali več podstavkov. Podstavek lahko vsebuje elemente kot osebek, predmet, povedek, prilastek, prislovno določilo stopnje, kraja, načina, namena, vzroka, časa ipd. Vsak podstavčni element v e-sperantu je lahko beseda, zaporedje ali podzaporedje. Lastnosti podstavka so občutki, organizacija in razmerje. E-speranto podstavčni občutki so lahko trdilni, vprašalni, velelni in vzklicni. Organizacija podstavka je lahko v obliki prostega stavka, razširjenega stavka ali razširjenega kompleksnega stavka. Razmerja med podstavki v e-sperantu so:

- podstavek AND podstavek (»*Pred petnajstimi leti sem prišel domov in sem ostal doma tri tedne.*«)
- podstavek OR podstavek (»*Greh je greh, če je storjen enkrat ali tisočkrat, če je vsakdanji ali nepoznan.*«)
- podstavek INTENTION podstavek (»*Siloma sem vodil svoje misli na bele ceste, na cvetoče travnike in dišeča polja, da bi ne videl sebe in svojega življenja.*«)
- SEQUENCE podstavek podstavek (»*Prišla je mati; stopala je počasi in varno, v roki je nesla skodelico kave.*«)

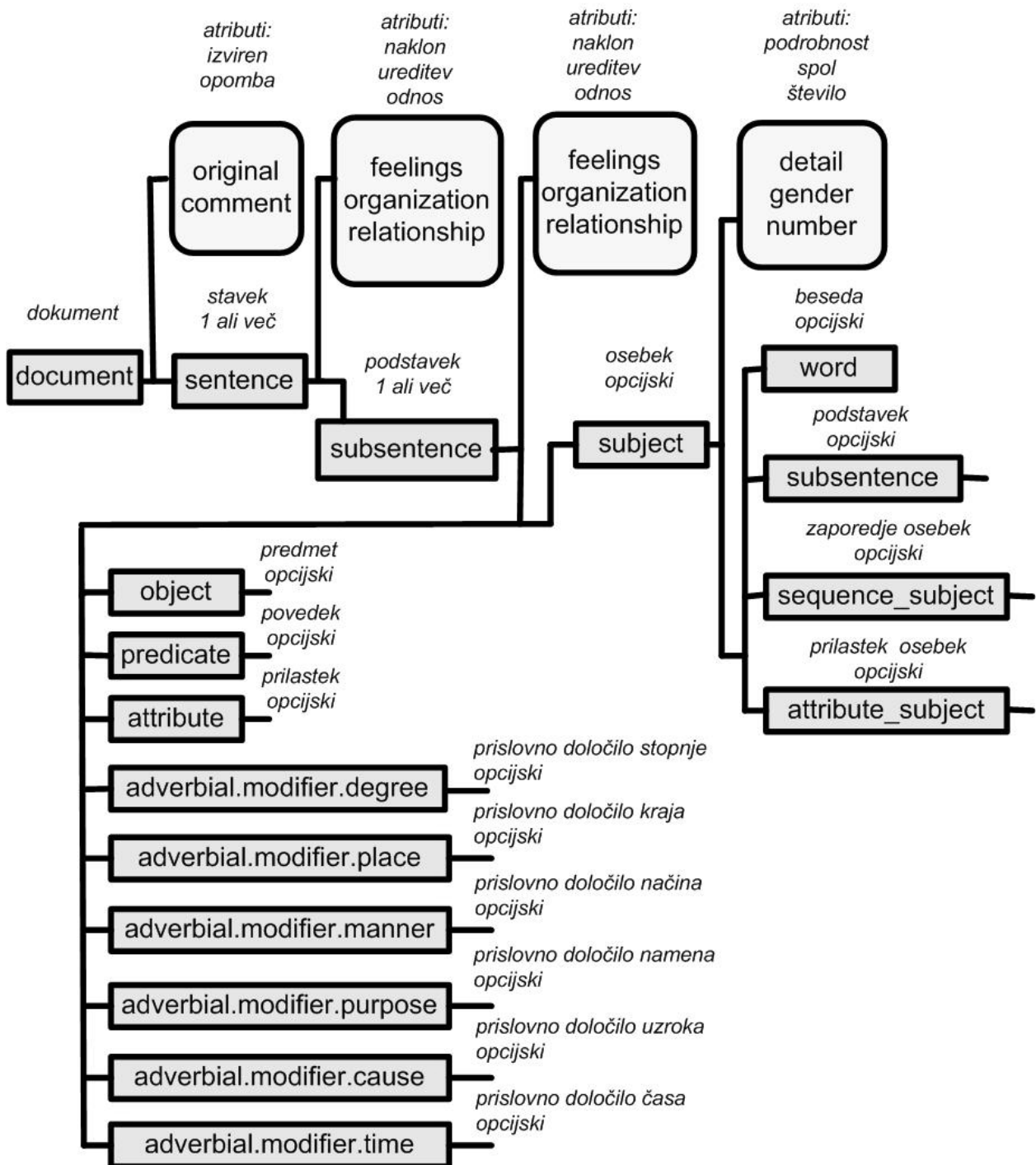
Vsaka beseda je v e-sperantu shranjena v osnovni obliki, glagoli v nedoločniku, samostalniki v ednini in podobno. Osebek, predmet, prilastek in prislovno določilo imajo lahko v e-sperantu naslednje karakteristike:

- podrobnosti (tu se navajajo različne podrobnosti, kot je recimo v pri lastnih imenih podatek, da se beseda ne prevaja, temveč se samo prenese v drug jezik v svoji izvorni obliki),
- spol (moški/ženski), le kjer ta označuje dejanski spol in ni le del slovnice jezika, in
- število (ednina/dvojina/množina).

Karakteristike povedka v e-sperantu so naslednje:

- čas (pretekli/sedanji/prihodnji),
- oblika (če je izvorni glagol uporabljen v zanikani obliki),
- podrobnosti (posamezne pomembne podrobnosti, ki se nanašajo na uporabo glagola v stavku, kot na primer, da se glagol uporablja kot del prilastka).

Slovnici organizacija e-speranta je deloma nakazana na sliki 1.



Slika 1: Slovnična organizacija e-speranta. Dokument je sestavljen iz stavkov, stavki iz podstavkov in podstavki iz stavčnih členov. Stavčni člen je lahko osebek, predmet, povedek, prilastek, prislovno določilo stopnje, kraja, načina, namene, vzroka, časa ipd.

Figure 1. E-speranto grammar organization. A document can be a set of sentences and each sentence a set of sub sentences. Each subsentence can have elements like subject, object, predicate, attribute, adverbial modifier of the degree, place, manner, purpose, cause, time, etc.

Lastnosti, določene za vsako besedo v e-sperantu, morajo omogočati, da se besede prevajajo in

razlagajo pravilno in da se pomen lahko natančno prenaša v poljuben naravni jezik.

4 Skladnja e-speranta

Kot osnovo za skladnjo e-speranta smo izbrali XML. XML je primerno strukturiran in standardiziran, kar je pomembno za avtomatsko obdelavo; obenem pa so dokumenti XML berljivi. Zadnje je pomembno predvsem pri razvoju različnih aplikacij, saj bistveno olajšajo razvoj, kar se je že večkrat potrdilo, predvsem v razvoju interneta. Vsebina dokumenta XML je opremljena z značkami (angl. tagi). V dokumentu XML lahko v nasprotju z dokumenti v HTML (Hyper Text Markup Language – jezik za označevanje dokumentov) definiramo poljubno število oznak. Vsaki oznaki lahko pripišemo več lastnosti (atributov), vsebina vsake oznake pa lahko vsebuje nove oznake. Te lastnosti skladnje XML zadoščajo za strukturiranje dokumenta v skladu s slovnico, ki je bila predstavljena v prejšnjem razdelku.

Omenimo tudi, da je XML združljiv s HTML, kar omogoča vključevanje e-speranta v spletne strani, kar je eden izmed ciljev razvoja e-speranta.

Slovnicična pravila e-speranta se lahko opišejo v shemi XML, na podlagi katere je mogoče avtomatsko preverjati skladnjo dokumenta. Del sheme XML e-speranta je podan na sliki 2, vendar pa je shema še v razvoju. V nadaljnjem bomo po potrebi dodajali ali brisali posamezne elemente, da bi povečali sposobnost zajemanja pomena. Celotno shemo oziroma njeno trenutno stanje je mogoče na zahtevo dobiti od avtorjev, v kratkem pa bo na voljo tudi na spletni strani e-speranta <http://www.e-speranto.si>, ki je trenutno še v izdelavi.

```
<xs:element name="subsentence">
  <xs:complexType>
    <xs:all>
      <xs:element ref="subject" minOccurs="0"/>
      <xs:element ref="object" minOccurs="0"/>
      <xs:element ref="predicate" minOccurs="0"/>
      <xs:element ref="attribute" minOccurs="0"/>
      <xs:element ref="adverbial.modifier.degree" minOccurs="0"/>
      <xs:element ref="adverbial.modifier.place" minOccurs="0"/>
      <xs:element ref="adverbial.modifier.manner" minOccurs="0"/>
      <xs:element ref="adverbial.modifier.purpose" minOccurs="0"/>
      <xs:element ref="adverbial.modifier.cause" minOccurs="0"/>
      <xs:element ref="adverbial.modifier.time" minOccurs="0"/>
    </xs:all>
    <xs:attributeGroup ref="subsentence_att"/>
  </xs:complexType>
</xs:element>
<xs:attributeGroup name="subsentence_att">
  <xs:attribute name="address" type="xs:string"/>
  <xs:attribute ref="feelings"/>
  <xs:attribute ref="organization"/>
  <xs:attribute ref="relationship"/>
</xs:attributeGroup>
<xs:element name="subject">
  <xs:complexType>
    <xs:choice>
      <xs:element ref="word" minOccurs="0"/>
      <xs:element ref="subsentence" minOccurs="0"/>
      <xs:element ref="sequence_subject" minOccurs="0"/>
      <xs:element ref="attribute_subject" minOccurs="0"/>
    </xs:choice>
    <xs:attributeGroup ref="subject_att"/>
  </xs:complexType>
</xs:element>
```

Slika 2: Del sheme XML, ki opisuje osnovno strukturo stavka v e-sperantu.

Figure 2. A part of XML scheme which describes basic structure of e-speranto sentence.

Primer zapisa stavka, prevedenega iz slovenskega jezika v e-speranto, je prikazan na sliki 3.

```
<?xml version="1.0" encoding="UTF-8"?>
<root xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="C:\Documents and Settings\sanidao\My Documents\E-speranto\26-1.xsd">
  <document>
    <sentence relationship="SEQUENCE of subsentences">
      <subsentence feelings="declarative" organization="simple">
        <subject number="singular"><word>mother</word></subject>
        <predicate tense="present"><word>to step</word></predicate>
      </subsentence>
      <subsentence feelings="declarative" organization="compound">
        <subject number="singular"><word>she</word></subject>
        <predicate tense="past"><word>to walk</word></predicate>
        <adverbial.modifier.manner number="singular"><sequence_adverbial><word>slow</word><word>safe</word></sequence_adverbial></adverbial.modifier.manner>
      </subsentence>
      <subsentence feelings="declarative" organization="compound">
        <subject number="singular"><word>she</word></subject>
        <predicate tense="past"><word>to carry</word></predicate>
        <adverbial.modifier.place number="singular"><word>hand</word></adverbial.modifier.place>
        <object comment="collocation cup-of-coffe" number="singular"><sequence_object><word>cup</word><word>coffe</word></sequence_object></object>
      </subsentence>
    </sentence>
  </document>
</root>
```

Slika 3: E-speranto zapis stavka: *Prišla je mati; stopala je počasi in varno, v roki je nesla skodelico kave.*

Figure 3. E-speranto record of the sentence: *Prišla je mati; stopala je počasi in varno, v roki je nesla skodelico kave.*

5 E-speranto in Eclipse

Pri razvoju e-speranta nujno potrebujemo ustrezna razvojna orodja. Odločili smo se za izdelavo integriranega razvojnega okolja (Integrated Development Environment, IDE), ki je namenjeno pisanju besedil v e-sperantu in temelji na razvojnem orodju Eclipse [8].

Eclipse je raztegljivo odprtokodno razvojno orodje za gradnjo razvojnih okolij in aplikacij, ki temeljijo na platformi rich-client. Platforma sama po sebi ne ponuja velike funkcionalnosti; funkcionalnost se doseže na podlagi modela vtičev (vtič je prenosni modul, ki se dinamično vklaplja v platformo). Z dodajanjem vtičev določeno okolje pridobiva funkcionalnosti, ki jih končni uporabnik potrebuje za izvajanje določenih opravil. Povezava med vtiči se doseže z uporabo razširitvenih točk (posebej definiranih razširitvenih mest v programski kodi), ki jih spremlja aplikacijski programski vmesnik. Prek vtičnic ponujajo vtiči svoje storitve naprej drugim, ki ponujene funkcije ponudnikov nadgrajujejo, s tem pa se funkcionalno širi tudi platforma. Eclipse je odprto razširljivo razvojno okolje. Zaradi svoje modularnosti, prilagodljivosti, preprostega razširjanja in vgrajene podporne infrastrukture za gradnjo razvojnih okolij, se je Eclipse izkazal kot ustrezna osnova za izgradnjo okolja, ki podpira razvoj e-speranta.

Razvojno okolje e-speranto trenutno sestavljata dva vtiča: urejevalnik XML in orodje za delo s slovarji. Funkcionalnost okolja vključuje: hitro sestavljanje dokumentov e-speranta s pomočjo sprotnih predlogov vsebine in preverjanja skladnosti strukture dokumenta na podlagi izbrane sheme; tudi hitro dostopanje do slovarjev in njihovo učinkovito uporabo med sestavljanjem in prevajanjem dokumenta.

V prihodnje bo razvojno okolje razširjeno z rezultati tretje (tolmači) in četrte (prevajalniki) faze razvoja e-speranta. Pričakujemo, da okolje, ki temelji na razvojnem orodju Eclipse, čez čas lahko postane osrednje podporno orodje za aplikacije, ki so povezane z večjezikovnim prevajanjem. Tudi orodja za razvoj bodo v prihodnje dostopna na spletni strani e-speranta.

6 Faze razvoja e-speranta

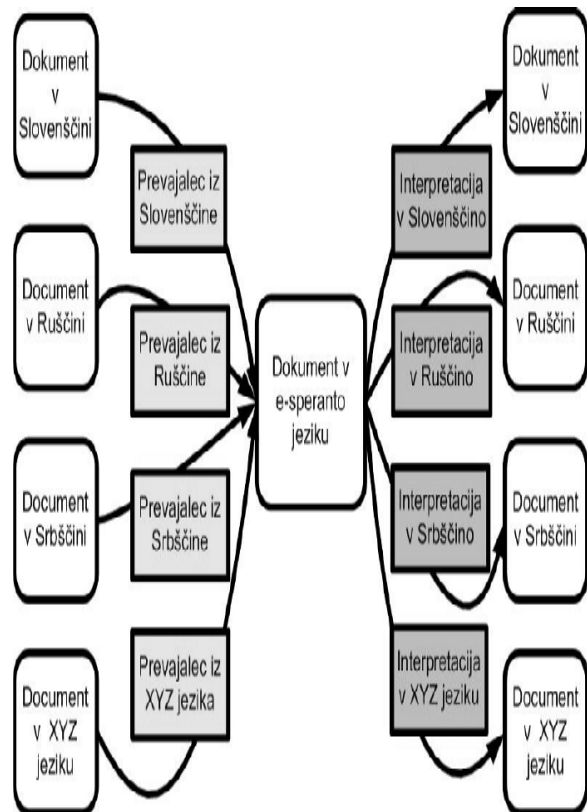
Razvoj e-speranta bo predvidoma potekal v štirih fazah:

1. zasnova jezika e-speranto (slovnica, skladnja, slovarji),

2. izdelava integriranega okolja in orodij za pisanje in preverjanje dokumentov v e-sperantu,
3. razvoj tolmačev (interpretov) e-speranta v naravnih jezikih in
4. razvoj prevajalnikov iz naravnih jezikov v e-speranto.

Prva in druga faza sta se že začeli, kot je to opisano v prejšnjih razdelkih.

Tolmačenje e-speranta v naravnih jezikih je nekoliko preprostejša naloga kot prevajanje iz naravnih jezikov v e-speranto (saj je v prvem primeru izvorno besedilo strukturirano in računalniku prijazno, kar močno olajša postopek). Izdelava ustreznih tolmačev, ki bi jih lahko vgradili v spletne brskalnike, že zadošča za prikaz spletne strani v e-sperantu. Razvoj tolmačev je tretja faza razvoja, razvoj prevajalnikov pa je predviden šele v četrti fazi. Prevajalniki e-speranto bodo delovali v smeri naravni jezik – e-speranto. Ko/če bodo na voljo tudi prevajalniki iz naravnih jezikov v e-speranto, pa bo to poleg ustvarjanja besedil e-speranto v domačem jeziku omogočilo tudi večjezično prevajanje med različnimi pari jezikov, kot je to prikazano na sliki 4, kar je tudi končni cilj razvoja e-speranta.



Slika 4: Večjezično prevajanje s pomočjo e-speranta.
Figure 4. Multilingual translation process with e-speranto.

7 Sklep

Ideja izdelave unikatnega računalniškega jezika za večjezikovno prevajanje prinaša možnost, da ljudje berejo dokumente, napisane v različnih naravnih jezikih, v svojem maternem jeziku.

Ta prispevek predstavlja e-speranto, formalni računalniški jezik, namenjen avtomatskemu prevajanju večjezikovnih dokumentov. Temeljni gradbeni element v e-sperantu je stavek, skladnja e-speranta temelji na XML. Razvoj e-speranta vključuje naslednje štiri faze: razvoj skladnje, slovnice in temeljnega besedišča; razvoj orodja za pisanje besedil v e-sperantu; razvoj prevajalcev (za prevajanje iz e-speranta v različne naravne jezike); in razvoj avtomatskega prevajalnika (neposredno iz naravnih jezikov v e-speranto). Razvoj avtomatskih prevajalcev za naravne jezike odpira možnost za avtomatsko večjezikovno prevajanje, kar je končni cilj e-speranta.

Prednost pristopa k večjezičnemu prevajanju prek vmesnega jezika je predvsem v tem, da mora vsaka jezikovna skupina razviti zgolj tolmač e-speranta v svojem jeziku in prevajalnik iz svojega jezika v e-speranto. Tako odpade poraba razvoja prevajalnikov za vsak par jezikov.

Da bi si olajšali začetek razvoja tretje in četrte faze e-speranta, načrtujemo preizkusiti idejo e-speranta najprej za slovenski, ruski in srbski jezik, ker ti trije jeziki spadajo v isto jezikovno skupino (slovanski jeziki), imajo podobno slovnico in besednjak; obenem pa je vsak materni jezik vsaj enega od avtorjev.

8 Viri in literatura

- [1] Extensible Markup Language, www.w3.org, dostopano 1.4.2007
- [2] Presis, Amebis d.o.o, presis.amebis.si, dostopano 1.4.2007
- [3] Systran, www.systransoft.com, dostopano 1.4.2007.
- [4] Promt, www.promt.ru/ru/index.php, dostopano 25.3.2007.
- [5] Free Translation.com, www.freetranslation.com, dostopano 10.3.2007.
- [6] D. Harlow: The Sixteen Rules of Esperanto Grammar, donh.best.vwh.net/Esperanto/rules.html, dostopano 1.4.2007.
- [7] J. Hana: Esperanto Grammar, <http://www.ling.ohio-state.edu/~hana/esr/index.html>, dostopano 27.2.2007.
- [8] Eclipse.org, <http://www.eclipse.org>, dostopano 1.4.2007.
- [9] F. Amerio, G. Bonvecchiato, G.C. Fighiera, Esperanto: Data and Facts, 2nd edition, FEI-Milan 2002.

Sanida Omerović je leta 2005 diplomirala na Fakulteti za elektrotehniko (v Beogradu, Srbija). Trenutno dela magistrsko nalogo na Fakulteti za elektrotehniko (v Ljubljani) in dela v Laboratoriju za komunikacijske naprave (Fakulteta za elektrotehniko). Na raziskovalnem področju se ukvarja z avtomatskim prevajanjem, še posebej s sistemi za pridobitev znanja.

Grega Jakus je absolvent univerzitetnega študija elektrotehnike smeri Telekomunikacije na Fakulteti za elektrotehniko (v Ljubljani). Njegovo področje dela obsega proučevanje možnosti uporabe odprtokodnih razvojnih orodij pri izdelavi razvojnih okolij in aplikacij z obogatnimi odjemalci.

Tatjana Filimonova je leta 2004 doktorirala na Filološki fakulteti (MGU, Rusija). Na Fakulteti za elektrotehniko dela kot mladi raziskovalec v Laboratoriju za komunikacijske naprave. Na raziskovalnem področju se ukvarja s računalniškim jezikoslovjem, leksikologijo in leksikografijo.

Sašo Tomažič je redni profesor Univerze v Ljubljani na Fakulteti za elektrotehniko. Je predstojnik Laboratorija za komunikacijske naprave in predstojnik Katedre za telekomunikacije. Njegovo raziskovalno delo vključuje raziskave pri obdelavi signalov, varnosti v telekomunikacijah, elektronskega poslovanja in informacijskih sistemov.