

Björn Jónsson

Hagnýt notkun tungutækni í símtölvunarlausnum

Björn Jónsson

Hagnýt notkun tungutækni í símtölvunarlausnum

Inngangur

Grunnur fékk 2,5 milljónir í styrk frá Tungutækniþjóði í framhaldi af umsókn, þar sem heiti verkefnisins var „Hagnýt notkun tungutækni í símtölvunarlausnum“. Markmið verkefnisins var að kanna hagnýtingu tungutæknieininga, sem til eru fyrir íslensku, m.a. í kerfum sem Grunnur hefur þróað og eru í almennri notkun.

Notast skyldi við þá talgreina og talgervla sem aðgangur væri að. Gert var ráð fyrir að notast við talgervil (Snorri) sem til er fyrir íslensku. Til samanburðar var einnig nýttur enskur talgervill frá Microsoft. Við upphaf verkefnisins var ekki til íslenskur talgreinir en stefnt var að því að nýta *open source*-talgreini frá Carnegie Mellon og kenna honum einföld íslensk orð (já, nei, tölur o.s.frv.) sem nauðsynleg eru í símtölvunarlausnum. Með tilkomu Hjal-verkefnisins og þeirrar ákvörðunar Landssíma Íslands að kaupa Scansoft-talgreini með íslensku sem byggist á niðurstöðum Hjal-verkefnisins var eðlilega tekin ákvörðun um að byggja á Scansoft-talgreininum í staðinn. Notast var við ís-

lensku útgáfuna, en sú enska var einnig notuð til hliðsjónar.

Ákveðið var að útbúa einfalt talsetur (e. Voice Portal) sem innihéldi dæmi um virkni eins og upplestur fréttar, upplýsingar um flug, uppflettingu í símaskrá, kaup á vöru og ýmislegt fleira. Því til viðbótar skyldi nýta þessa tækni í tengslum við Tímon og Snældu sem eru símtölvunarlausnir sem Grunnur hafði þróað. Á verkefnatímanum sameinaðist hugbúnaðarsvið Grunnus Trackwell undir nafni Trackwell og var verkefninu haldið áfram í nafni Trackwell.

Forritun var að mestu unnin í Visual C++ 6.0. Þetta gerði okkur kleift að nota talgreinana á lágu plani, þ.e. án millilaga. Að sumu leyti er auðveldara að forrita tungutækniforrit í málum, s.s. Vxml (voice-xml), og útbjuggum við einnig ákveðin tungutæknidæmi með VoiceXML til samanburðar. VoiceXML hentar ágætlega þegar útbúa skal einföld kerfi, en þegar útbúa þarf flóknari kerfi teljum við að alvöru forritunarmál eins og C++ henti betur. Meðal annarra þægilegra möguleika við gerð svona kerfa má nefna SALT (speech-application-langu-

age-tags). Notkun C++ gerði okkur tæknilega mögulegt að rannsaka betur hina ýmsu möguleika tungutækninnar, þar sem við höfum betri stjórn á því sem verið er að vinna með.

Helstu niðurstöður

Gerð þessara lausna tókst vel, en reyndist töluvert meiri vinna en áætlað var, auk þess sem yfir ákveðinn þröskuld var að fara til að byrja með, sérstaklega hvað varðar notkun talgreinanna. Annar tæknilegur þröskuldur var samvirkni talgreinisins og svarþjónskjarna Grunns. Umbreyta þurfti svarþjónskjarnanum til að hann gæti haft talviðmót til viðbótar við hefðbundið DTMF-viðmót (veldu 1 fyrir X, veldu 2 o.s.frv.). Þetta var leyst almennt og er því svarkerfiskjarninn tilbúinn til notkunar fyrir þróun talgreinlausna. Slík sambætting er hins vegar einfaldari ef notast er við VoiceXML. Í öðru lagi kom í ljós, eins og raunar vitað var, að talgreinir er í eðli sínu þung keyrsla í samburði við t.d. svarþjónskerfið sem byggist á DTMF-viðmóti.

Það kom líka skýrt í ljós að notkun tungutæknieininga með símtölvunarlausnum krefst ekki eingöngu forritunarhæfni heldur er þekking á málfræðiþættinum nauðsynleg.

Talgervlar

Notkun talgervla er tæknilega einföld, sér í lagi þeir ef þeir uppfylla SAPI 5-tungutækni-

staðalinn. Vandamálið er að íslenski talgervilinn Snorri er lélegur, reyndar það vöndur að deila má um hvort hann sé nothæfur í almennar lausnir. Snorri venst reyndar og smám saman læra notendur að skilja hann. Undantekningalítið sögðust þó nýir notendur eiga erfitt með að skilja hann. Í það minnsta er hann fráhrindandi og ráðleggjum við þeim sem útbúa vilja símtölvunarlausnir að notast við samsettar og foruppteknar hljóðskrár þar sem slíkt er mögulegt.

Talgreinar

Langmesta vinnan fór í notkun talgreina og fékkst mjög góð reynsla af notkun þessara lausna sem á eftir að nýtast vel við frekari gerð tungutæknilausna. Talgreinirinn frá Scansoft er svokallaður stakorðagreininir og einskorðaðist notkunin við það. Talgreinirinn virkaði vel í flestum tilfellum og sá íslenski betur en sá enski, sem ekki er óeðlilegt þar sem notendur voru allir íslenskir auk þess sem íslenskan er einsleitara tungumál en enska. Það er samt vísbending um að vel hafi tekist til með gerð íslenska talgreinisins og að hann eigi eftir að nýtast vel við gerð mismunandi kerfa.

Eftirfarandi eru nokkur atriði sem rétt er að hafa í huga við gerð slíkra lausna sem snúa að viðmóti við notendur, sem er ákaflega mikilvægur þáttur við gerð tungutæknilausna:

Skilgreining orðasafna. Mikilvægt er að skilgreina hvaða orð eigi að skiljast á mis-

munandi stöðum í svartrénu. Eðlilegt er að hafa jafnvel mörg leyfileg orð fyrir sömu skipun.

Ágiskunarstilling. Hægt er að stilla svokallaðan ágiskunarstuðul sem segir til um hvort reynt sé að giska á um hvaða orð er að ræða ef talgreinirinn er ekki viss um hvað sagt var. Það fer eftir eðli lausna hvernig ber að stilla þennan stuðul.

Staðfestingartextar. Í sumum tilfellum er nauðsynlegt að spyrja notandann hvort hann hafi sagt ákveðið orð til staðfestingar, þetta á ekki síst við þegar um kaup á vöru eða þjónustu er að ræða. Slíkt ferli hægir á vinnslu og því mikilvægt að nota rétt.

Staðfestingartónar. Mikilvægt er að viðmótið virki þægilega á notandann, t.d. að ekki komi langar þagnir meðan kerfið er að sækja gögn, svo dæmi sé tekið. Með því að nota staðfestingartóna og önnur hljóðtákn, t.d. þegar verið er að vinna eða sækja gögn, má gera flæðið mun eðlilegra.

Hjálp á öllum stigum. Æskilegt er að notendur geti fengið aðstoð hvar sem er í kerfinu, einfaldlega með því að segja „hjálp“. Helst þarf hjálpin að vera þannig uppbyggð að hún byggist á því hvar viðkomandi er staddur í valtrénu.

Geta hætt hvar sem er. Á sama hátt er æskilegt að notendur geti hvar sem er sagt „hætta“ til að hætta vinnslu eða fara efst í valtréð.

Þótt notkun talgreina gefi marga nýja möguleika fram yfir hefðbundin svarkerfi eru samt ýmsir vankantar við notkun þessarar tækni. Þar má t.d. nefna eftirfarandi atriði.

Seinlegt að segja talnarunur. Með talgreini þarf vanalega að segja eina tölu í einu. Í flestum tilfellum er fljótlegra að slá inn tölur á takkaborði símtækis.

Staðfestingar. Staðfestingar á svörum eru þreytandi en samt oft nauðsynlegar, t.d. þegar tölur eða upphæðir eru lesnar inn. Innláttur frá takkaborði Símans er öruggari.

Hægvirvari vinnsla. Í mörgum tilfellum er fljótlegra að nota hefðbundin svarkerfi en með talgreini. Þetta á ekki síst við þegar framkvæma þarf sömu aðgerðina oft.

Skilningur. Framleiðendur talgreina lofa að þeir skilji það sem sagt er við þá í 97% tilfella. Dregið er í efa að slíkur skilningur sé fyrir hendi, a.m.k. kemur í ljós að ef um bakgrunnshávaða er að ræða minnkar þessi skilningur. Á sama hátt virðist sem skilningur úr GSM-símum sé lakari en úr borðsímum. Hefðbundin svarkerfi, þar sem tölur eru slegnar inn á takkaborði, hafa 100% skilning.

Notkunarmöguleikar þessara tungutæknilausna eru miklir og má telja upp fjölda möguleika í því sambandi. Á einfaldan hátt má þó skipta þessum þjónustum í nokkra flokka:

- Viðbót við núverandi svarkerfislausnir (upplýsinga- og fréttabjónustur)

- Þjónustur sem spara mannskap (síma-skrá, þjónustuver)
- Tekjuskapandi lausnir (kaup og sala á ýmsum vörum og þjónustu)

Lokaorð

Í upphafi var gerð ákveðin verk- og tíma-áætlun sem lögð var fram með umsókn. Verkefnið stóð yfir í u.þ.b. eitt og hálf ár með hléum. Þetta er ákveðin seinkun sem var vísitandi gerð þar sem beðið var eftir að íslenskur talgreinir yrði tilbúinn til prófana. Jafnframt var haldið utan um þá vinnu sem fór í verkefnið en hún fór allnokkuð fram úr áætlun, aðallega þar sem forsendur breyttust á verktímanum.

Að verkinu komu eftirfarandi starfsmenn Grunns:

Ívar Ragnarsson
Orri Eiríksson
Björn Jónsson

Það er mat okkar að þetta verkefni hafi tekist vel og muni nýtast okkur og vonandi öðrum sem stefna að gerð símtölvunarlausna sem nýta eiga sér tungutæknieiningar. Tungutækni í símtölvunarlausnum er mjög öflug viðbót en ýmsa pytti ber að varast til að árangurinn verði góður.