

LATVIJAS UNIVERSITĀTE
DATORIKAS FAKULTĀTE

**LATVIEŠU VALODAS SINTAKTISKĀ ANALIZATORA
"ČANKERIS" MODERNIZĀCIJA**

MAGISTRA DARBS

Autors: Zane Siliņa
Stud. apl. Nr. zk09018
Darba vadītājs: profesors Dr.sc.comp. Guntis Bārzdīņš

RĪGA 2015

ANOTĀCIJA

Darbā tiek apskatīts Latvijas Universitātes Mākslīgā Intelektā laboratorijā izstrādāts latviešu valodas sintaktiskais analizators ‘Čankeris’. Analizatoram nepieciešama modernizācija, lai mazinātu analīžu daudznozīmību. Tiek veikta programmas divu algoritmiski atšķirīgu versiju salīdzināšana, lai izvēlētos tālākai modernizācijai piemērotāko. Tiek atrasts veids kā izvēlētajā programmā iekļaut mašīnmācīšanās metodes. Pēc programmas uzlabošanas tiek veikti testi un rezultātu analīze.

Atslēgvārdi: prologs, sintaktiskais analizators, mašīnmācīšanās, shift-reduce, vārdu attēlošana ar vektoriem

ANOTATION

The paper discusses modernization of Latvian language syntactic analyzer "Čankeriš" developed in Latvian University laboratory of Artificial Intelligence. Modernisation s needed to reduce result ambiguity. Comparison is made between two algorithmically different versions of program , to select the most suitable for further modernization. The most appropriate way to include machine learning methods in selected program is found. After improvement tests are carried out and analysis made.

Atslēgvārdi: Prolog, dependency parser, machine learning, shift-reduce, native language processing, word representations as vectors

AUTOREFERĀTS

Darbā tiek pielietota mašīnmācīšanās latviešu valodas sintaktiskajam analizatoram "Čankeris", tā mazinot rezultāta daudznozīmību. Analizatoram ir unikāla pieeja rezultātu iegūšanā, bet tas nespēj tos automatizēti sintaktiski kategorizēt, kas programmu padara par novecojušu uz līdzīgu fona. Modernizācijā tiek izmantota arī vārdu līdzības noteikšana pēc to vektoriem un fails ar latviešu valodas vārdu reprezentācijām vektoros, tiek izmants pirmo reizi vispār.

Literatūras izpēte ir tikusi veikta, bet dodot diezgan vispārīgu skatījumu par tēmu, kurā notiek pētījums – datorlingvistiku un mašīnmācīšanos. Bet darba mērķis bija spēt nodrošināt programmu ar modernizāciju no izstrādes viedokļa, lai vispār modernizācija būtu iespējams veikt. Kā teorija, kas neparādās darbā, tiek skaitīta arī esošās realizācijas apgūšana un programmēšanas valodas mācīšanās.

Problēma ir izklāstīta skaidri, tāpat tiek arī detalizēti skaidrots izvēlētais risinājums - idejiski un realizācija. Teksta apjoms ir salīdzinoši mazs, bet koncentrēts ar informāciju.

Praktiskais darbs satāv no trīs daļām, ir salīdzinātas un analizētas programmas versijas, veikta modernizācijas projektēšana un izpilde, un tiek pētīti modernizētās programmas testēšanā sniegtie rezultāti, salīdzinot ar programmas darbību pirms tās un dažādos modernizācijas līmeņos.

Modernizāciju testējā autore, apstrādājot dabīgu testa datu kopu. Analizēti 174 teikumi, 104 ar analizējamiem rezultātiem. Par tās veiksmīgumu tiek lemts pēc testa datu analīzes, tiek salīdzināti rezultāti pirms modernizācijas un pēc tās.

Darbs tika formatēts pēc norādītajām prasībām, korektu izpildi nodrošināja darba pārlasīšana.

Visas idejas, kas tika aizgūtas pētot citu autoru darbus, izmantotās metodes, pieminētās pieejas, rīku apraksti un situācijas apraksti, tiek pamatotas ar atsaucēm uz izmantoto avotu. Norādītajās atusaucēs ir ne tikai rakstītā pamatojums, bet arī papildus informācija par tēmām, kas darbā iespējams nav pietiekami plaši izklāstītas.

SATURS

IEVADS	6
1. Sintaktiskā analīze, veidi un analizatori	7
1.1. Sintaktiskā analīze	7
1.2. Analīze	9
1.3. Analizatori	9
2. „Čankeri”	11
2.1. Analizatora realizācija	11
2.2. ”Čankera” versijas	12
2.3. Vērtēšanas kritēriji un vērtēšana	14
2.4. Rezultāti	15
2.5. Secinājumi	19
3. Analizatora ” Čankeri” modernizācija	20
3.1. Piedāvātā metode	20
3.2. Pāru meklēšana pēc vārdu līdzības	21
3.3. Realizācija	21
4. Modernizētā ” Čankera” novērtēšana	25
4.1. Testēšanas plāns	25
4.2. Rezultātu analīze	26
SECINĀJUMI	31
IZMANTOTĀ LITERATŪRA UN AVOTI	32
PIELIKUMI	34
1.pielikums	35
2.pielikums	39
3.pielikums	44
4.pielikums	50

IEVADS

Ja latviešu valoda vēlas izdzīvot nākotnes pasaulē, tai ir jārada datorlingvistiskā infrastruktūra, jārada programmas un rīki, kas ļautu datoram saprast latviešu valodu [1]. Šajā darbā tiek apskatīts viens no rīkiem ar šādu potenciālu - Matemātikas un informātikas institūta Mākslīgā intelekta laboratorijā izveidotais latviešu valodas sintaktiskais analizators "Čankeris". Analizators veic teikumu morfoloģisko un sintaktisko analīzi — nosaka vārda morfoloģisko formu un, balstoties uz tām, nosaka iespējamo teikuma sintaktisko struktūru[2].

Sintaktisko analizatoru lielākā problēma ir augstā analīzes daudznozīmība[3], apskatītajā programmā tā ir īpaši augsta, jo programma atrod visus iespējamus analīzes rezultātus, nevis tikai vienu. Potenciāli ir lielāka iespējamība iegūt pareizo rezultātu, nevis iegūt tikai vienu, ko programma novērtējusi kā pareizāko, kad tiek atrasts tikai viens rezultāts. Mašīnmācīšanās dotu iespēju programmai uzlabot daudznozīmības problēmu - saprast, kurš no variantiem ir ticamākais. Līdzīgas funkcionalitātes programmas jau darbojas, izmantojot mašīnmācīšanās metodes sintaktiskā līmenī, bet "Čankerim" nav piemēlēta realizācija. Programma iespējams ir unikāla starp pārējām, jo vismaz darbā aplūkoti analizatori, visi atrod vienu labāko variantu, bet tai nepieciešams samazināt sintaktisko daudznozīmību.

Darba mērķis ir apgūt loģisko programmēšanas valodu *prolog* un analizatora "Čankeris" realizāciju, lai varētu to papildināt ar mašīnmācīšanās pieeju un uzlabot programmas sniegto rezultātu, pievienojot aprakstīto metodi.

Darba sākumā lasītājs tiek iepazīstināts ar jēdzienu sintaktiskā analīze priekš cilvēka un datora un eksistējošiem citiem analizatoriem. Tālāk tiek apskatīta "Čankera" realizācija un divas jaunākās programmas versijas, kas atšķirās ar izmantoto pamatalgoritmu teikumu analizēšanā. Tiek veikta to novērtēšana, lai atrastu efektīvāko uzlabošanai. Tiek piedāvāta realizācija kā programmā iekļaut mašīnmācīšanos un tās apraksts. Darba noslēgumā tiek veikta modernizētās programmas testēšana un rezultātu novērtēšana.

1. SINTAKTISKĀ ANALĪZE, VEIDI UN ANALIZATORI

Pirmās nodaļas mērķis ir sniegt īsu pārskatu, kas ir latviešu valodas sintaktiskā analīze, kādas zināšanas ir nepieciešamas, lai to varētu veikt dators, kā arī dod nelielu ieskatu par izstrādātiem analizatoriem.

1.1. Sintaktiskā analīze

Sintaktiskā analīze pēc akadēmiskās terminu bāzes ir "sintaktisko sakaru, jēdzienisko attieksmju un sintaktisko kategoriju noteikšana sintakses vienībās – vārdu savienojumā un teikumā"[4]. Zemāk tiek aprakstītas sintaktiskās kategorijas un vārdšķiras, ar kurām tās var tikt izteiktas. Kā arī mazs ieskats par vārdšķirām latviešu valodā .

Sintaktisko kategoriju pamatā ir vārdu formas un vārdu savstarpējās saistīšanās iespējas. Pēc nozīmības tos iedala virslocekļos un palīglocekļos. Virslocekļi ir teikuma priekšmets, izteicējs. Tabulā 1.1 ir kopsavilkums par teikuma virslocekļiem - nosaukums, jautājums uz kādu tas teikumā atbild un vārdšķiras, ar kurām to izsaka.[5]

1.1. tabula

Teikuma virslocekļu pārskata tabula[5]

Nosaukums un pasvītrojums	Jautājumi	Vārdšķira, ar kuru virsloceklis ir izteikts
Teikuma priekšmets, pasvītro ar vienu līniju	kas?	lietvārds, vietniekvārds, īpašības vārds, darbības vārds divdabja formā, skaitļa vārds, darbības vārds nenoteiksmē
Izteicējs, pasvītro ar dubultlīniju	ko dara?kāds ir? kas ir?	darbības vārds, lietvārds vai vietniekvārds ar saitiņu , īpašības vārds, lokāms divdabis, skaitļa vārds ar saitiņu un apstākļa vārds

Teikuma virslocekļus apraksta pārējie teikuma locekļi - palīglocekļi.

Galvenie palīglocekļu veidi ir apzīmētājs, pielikums, papildinātājs, apstākļi un dubultloceklis. Tabulā 1.2 kopsavilkums par teikuma palīglocekļiem - nosaukums, vārdšķira, kurai loceklis var būt pakārtots, tā loma teikumā un jautājums uz kādu tas teikumā atbild.

1.2. tabula

Teikuma palīglocekļu pārskata tabula[6]

Teikuma palīglocekļa nosaukums	Kam palīgloceklis pakārtots	Ko palīgloceklis izsaka	Raksturīgie jautājumi
Apzīmētājs	lietvārdam vai tā aizstājējam	pazīmi	kāds? kāda? kurš? kura? cik? kā?
Papildinātājs	parasti darbības vārdam	Paskaidro darbības vārdu; norāda, uz ko vērsta darbība vai ar ko to veic	ko? kam? ar ko? par ko? pret ko? uz ko? no kā? ko darīt? kā?
Apstākļi	darbības vārdam, īpašības vārdam, apstākļa vārdam	īpašību, pazīmi, kas piemīt darbībai vai stāvoklim	jautājumi atkarīgi no apstākļa nozīmes
Pielikums	lietvārdam, vietniekvārdam	paskaidro ko priekšmetisku	–
Dubultloceklis	lietvārdam vai vietniekvārdam un darbības vārdam vienlaikus	izsaka personas vai priekšmeta stāvokli vai pazīmi kādas norises laikā	–

Vienu teikuma loekli var veidot dažādu vārdšķiru vārdi un nevis tikai viens, bet arī vairāki vārdi. Vārdus grupē vārdšķirās pēc vairākām pazīmēm: vārdu vispārīgās nozīmes - nozīmes, kas piemīt vārdiem līdzās to konkrētajām nozīmēm, vārdu morfoloģisko formu īpatnībām, vārdu sintaktiskā lietojuma. Latviešu valodā izdala 10 vārdšķiras:

- lietvārds,
- īpašības,
- skaitļa,
- vietniekvārds,
- darbības vārds,
- apstākļa vārds,
- prievārds,
- saiklis,
- partikula,
- izsaukmes vārds.[7]

Pēc vārdu nozīmes īpatnībām un to izmantojuma teikumā izšķir divas vārdšķiru grupas: patstāvīgos vārdus, vārdi var būt par teikuma locekļiem, un palīgvārdus, parasti nevar būt par teikuma locekļiem. Palīgvārdi - prievārdi, saikļi un partikulas.[7]

1.2. Analīze

Sintaktiskā analīze tiek balstīta uz vārdu formām., kuras cilvēkam ir intuitīvi skaidras. Tiek viennozīmīgi atpazīti patstāvīgie vārdi - darbības vārdi no lietvārdiem, no īpašības vārdiem, palīgvārdi ir noteikti. Tāpēc cilvēks sintaktiskajā analīzē neizjūt tik lielu nepieciešamību pēc morfoloģijas zināšanu pielietošanas, nav nepieciešamības katru vārdu morfoloģiski analizēt. Datoram nepiemīt šāda intuīcija, tam vajag likumus, kā atpazīt vienu vārdšķiru no otras. Vārdam, vispirms ir jānosaka vārdšķira, un tikai tad var veikt sintaktisko analīzi. [3]

Latviešu valoda ir sarežģīta, jo bieži ir gadījumi, kad vārdšķiru var noteikt tikai pēc pārējā teikuma konteksta - vārdi māja (ēka) un māja (māja ar roku), datoram ir pilnīgi vienādi, cilvēkam, zinot kontekstu, to atšķiršana problēmas nesagādā. Daudznozīmība ir ne tikai morfoloģiskā līmenī, bet arī sintaktiskā. Ja morfoloģiskās atšķirības cilvēks intuitīvi izjūt, tad sintaksē jau rodas problēmas. Piemēram, palīglocekļu sintaktiskās funkcijas ne vienmēr ir viennozīmīgi atdalāmas un nosakāmas. Iespējami arī tādi teikuma palīglocekļi, kas pēc savām sintaktiskajām attieksmēm saistās ar vairākiem teikuma locekļiem vai izsaka vairākas nozīmes. Datoram problēmas rodas jau daudz vienkāršākā līmenī, piemēram, lietvārds akuzatīva locījumā varbūt gan laika, gan vietas, apstākļi - dienā, klētī. Lai dators viennozīmīgi atrastu vienīgo pareizo analīzi, tam ir nepieciešamas zināšanas par semantiku vai pietiekami liela sintaktsko zināšanu bāze, no kuras veikt secinājumus, lai izvēlētos pareizo rezultātu.[3,6]

1.3. Analizatori

Šobrīd eksistē ļoti daudz sintaktisko analizatoru realizācijas, piemēram, *MaltParser*, *Stanford Parser* un *Mate Tools*. Pēc analizatoru apskates tiek secināts, ka šie analizatori izdot tikai vienu –pareizāko, analīzes veidu[8, 9, 10], tādējādi ir varbūtība saņemt nepareizu rezultātu, bez alternatīviem analīžu variantiem. ”Čankeris” šajā ziņā ir oriģināls, tas izvada visus iespējamus analīzes veidus, ko var atrast pēc definētajiem sintakses un morfoloģijas likumiem.

Izvadot visus ,lietotājam tiek nodemonstrēta analīzes iespējamā daudznozīmības problēma. Lai mazinātu daudznozīmību, analizatoros tiek iekļauta mašīnmācīšanās, kas ļauj programmām pēc padotas informācijas veikt spriedumus par analīzes rezultātu pareizību [11]. ”Čankerī” ir iekļauta statiskā mašīnmācīšanās, bet tikai morfoloģiskā līmenī – „rezultāti tiek kategorizēti pēc doto formu biežuma morfoloģiskajā korpusā”[12] . Angļu valodas analizatoriem ir pieejamas arī

plašas sintaktisko zināšanu datu bāzes, kuras var izmantot, piemēram *Penn TreeBank*[13] . Kopš 2011. gada tiek strādāts pie sintaktiski marķēta korpusa latviešu valodā, bet pagaidām koprus ir neliels un darbs pie tā turpinās[14]. Tāpēc vienīgā pieejmā sintaktisko zināšanu datu bāze latviešu valodai ir jau "Čankerī" iekļautie sintakses likumi.

2. „ČANKERIS”

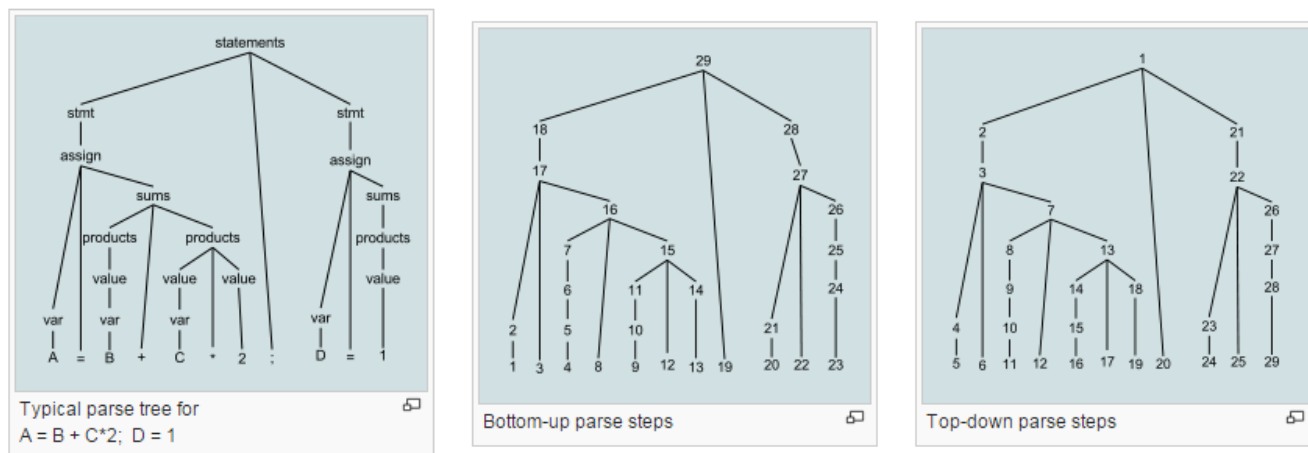
Otrā nodaļā tiek aprakstīta ”Čankera” realizācija, kas tiek apzināta ar lietotāju dokumentācijas un ar izstrādi saistīto rakstu lasīšanu, gan ar programmkoda izpēti. Veikta divu jaunāko algoritmiski atšķirīgu versiju novērtēšana ar mērķi izvēlēties vienu modernizācijai.

2.1. Analizatora realizācija

Latviešu valodas sitnaktiskajais analizators ”Čankeris” ir izstrādāts projekta Semti-Kamols ietvaros. Tā realizācija ietver morfoloģisko analīzi – vārdam tiek noteikta vārdšķira un piekārtoti atbilstošie morfoloģiskie atribūti[15]. Vārda vārdšķira tiek meklēta morfoloģiskajā bibliotēkā, ja tiek analizēts vārds, kas šajā bibliotēkā neietilpst, tiek piedāvāta iespēja minēt vārdšķiru. Sintaktiskā analīze tiek realizēta definējot likumus, kas balstās pēc veiktās morfoloģiskās analīzes. Likumā ietilpst galvenais vārds, sintaktiskā loma un paskaidrojošais vārds. Ar pamatalgoritma palīdzību izveidotie likumi tiek pielietoti teikuma vārdiem, kam jau ir veikta morfoloģiskā analīze. [16]

Galvenā un paskaidrojošā vārda attiecība programmā tiek attēlota ar sarakstu palīdzību – [galvenais vārds[sintaktiskā loma, paskaidrojošais vārds]], programmā katram vārdam ir pievienota arī morfoloģiskā analīze, pamatforma, informācija, vai vārds nācis no vārdnīcas vai ticis minēts. Teikums ir sarakstu struktūra, kas reprezentē koku. Veiksmīgas analīzes gala rezultāts ir saraksts ar katra varianta analīzes koku reprezentējošo sarakstu. Visu variantu atrašanai tiek izmantota *prolog* atsekošana, tas bez papildus funkcionalitātes izstrādes atrod visus iespējamus pareizos risinājumus, kas atbilst morfoloģijas un sintaktisko likumiem pieļautajām variācijām. Ir iespējama ļoti liela daudznozīmība.

sintakses likumiem meklē piesastītu. Izmantojot šo pieeju rezultātā var netikt iegūts koks. Attēlā 2.2. videājā kokā ir parādīta soļu secība, kādā tiktu veikta analīze tā paša attēla kreisās puses koka izteiksmei, izmantojot pieeju no lejas uz augšu.[3,17]



2.2.att., Analīzes koka veidošanas metodes pa soļiem

Jaunākā versija ir balstīta uz Nivres pētījumos aprakstītu algoritmu, kas izmanto tikai bottom up pieeju. Tajā tiek pielietota *shift-reduce* algoritma modifikācija ar nosaukumu – *arc-standart*. Algoritmā tiek izmantots steks S un saraksts W ar ievaddatiem – vārdiem, arku kopa A , kas reprezentēs atkarību grafu, atkarības starp vārdiem. Analizatora stāvokli raksturo trijnieks $\{S, W, A\}$. Izmantojot definētos likumus saistības tiek attēlotas ar arkām uz kreiso, *left_arc* un labo pusi, *right_arc* :

left_arc pievieno arku $n' > n$ no nākamā vārda n' uz steka pirmo vārdu n un izņem n no steka. Lai arka tiktu pievienot jābūt definētam likumam $LEX(n') \leftarrow LEX(n)$ un grafā nedrīkst atrasties arka $n' > n$, jo katrs vārds var būt atkarīgs tikai no viena cita vārda.

right_arc pievieno arku $n > n'$ no steka pirmā vārda n uz nākamo vārdu n' sarakstā un ieliek n' stekā. Lai arka tiktu pievienot jābūt definētam likumam $LEX(n) \rightarrow LEX(n')$ un grafā nedrīkst atrasties arka $n > n'$, jo katrs vārds var būt atkarīgs tikai no viena cita vārda. [18]

Algoritms izsauc četras funkcijas - *left_arc*, *right_arc*, *reduce*, *shift*. Atkarībā no vārda stekā un sarakstā, un definētajiem likumiem tiek pievienota konkrētā arka, *reduce* – izņem vārdu no steka, ja tā galvenais vārds jau ir noskaidrots un pievienots grafā, *shift* – ieliek vārdu stekā un pāriet uz nākamo vārdu sarakstā. [18] Šajā risinājumā rodas problēma, ka stekā var palikt vairāk kā viens elements, kam nav atrasts galvenais vārds, tad šie elementi tiek pievienoti mākslīgi izveidotai koka saknei, kurai pieliek pārējos elementus. Šāda situācija, kad mākslīgi veidotajai saknei ir pievienoti vairāki elementi, ir mazticama – šādā situācijā netiek iegūta koka struktūra.

Algoritmam ir jaunāka versija – *arc-eager*, kas risina šo problēmu, vēlreiz meklējot stekā esošajiem vārdiem galvenos vārdus. [19] Teikumā nevar būt situācija, ka vārdam nav galvenā vārda, tāpēc pareizajiem risinājumiem stekā paliek tikai viens elements.”Čankerim” automātiski veic atsekošanu – atrodot visus pareizos variantus, tāpēc jaunā versija nav nepieciešama.

2.3. Vērtēšanas kritēriji un vērtēšana

Kā galvenie kritēriji programmu vērtēšanai tiek izvirzīti ātrums un analīzes pilnība – jo tikai pilnas analīzes gadījumā, tiek meklēta visi iespējamie analīzes varianti. Variantu skaits tiek apskatīts atkarībā pret vārdu skaitu teikumā. Tas dod priekšstatu par teikumu garumu, kādā daudznozīmība sāk strauji palielināties, un vārdu skaitu teikumā, kurā variantu skaits ir reālā laikā pārskatāms.

Analīzei tiek izmantoti 3 autores sastādīti teikumi, lai analizētu specifiskus gadījumus - vielīdzīgu teikumu locekļu analīze, mākslineciskiem teikuma līdzekļiem bagāta teikuma analīze un teikums ar palīgvārdiem. Lai novērtētu programmu darbību uz reāla teksta tiek izmantoti teikumi no Edvard Virzas daiļdarba „Straumēni” . Analizators pašreiz ir rezlīzēts, lai analizētu tikai vienkāršus paplašinātus teikumus[20]. Tiek ņemti tikai tādi, tāpēc tie nav secīgi, viens otram grāmatā sekojoši. Lai notestētu dažāda garuma teikumu analīzi, katrs no testa teikumiem tiek izanalizēts pilns un tad analizēts izņemot vārdu vai vārdus, līdz paliek tikai izteicējs, pēc iespējas saglabājot jēgu, bet neuzstādot to kā pašmērķi.

Testēšanas rezultāti tiek fiksēti tabulā, kuras struktūra redzama tabulā 2.1.

2.1. tabula

Versiju vērtēšanas rezultātu pieraksta tabula

Piemēra nr.	Teikums	Vārdu skaits	Variantu skaits "Ziemassvētkiem"	"Ziemassvētki"	Variantu skaits "Vasarai"	"Vasara"

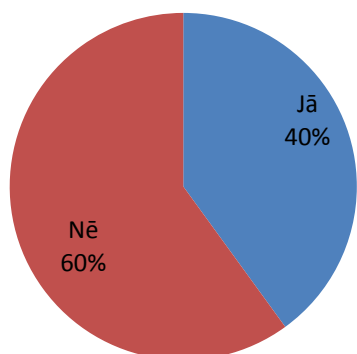
Vērtēšanas procesā tiek izmantoti izstrādātāju dotie versiju nosaukumi „Ziemassvētki” ir ar projektā veidoto algoritmu, ”Vasara” ir ar *shift-reduce* algoritmu.

2.4. Rezultāti

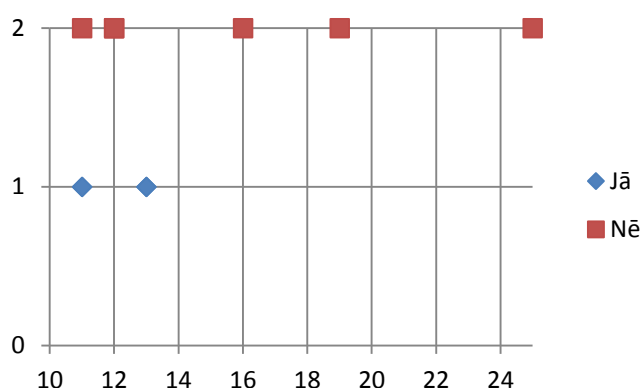
Iegūtie rezultāti tiek apkopoti diagrammās. Nodaļa tiek organizēta struktūrā - vispirms īss apraksts, kas tiek attēlots diagrammās un kāpēc veidota šāda diagramma, tad diagramma vai diagrammas un no tās veiktie secinājumi. Kopumā tiek veikta 10 teikumu analīze. Analizētie teikumi turpmāk tiek pieminēti, izmantojot to kārtas numurus, kas redzams zemāk esošajā teikumu uzskaitījumā.

1. Vasaras dienā sulīgi zaļiem zāles stiebriem klātā pļavā draiskā nevainībā spēlējās kucēni.
2. Vakara un rīta gaisma apspīd Jāņa un Ilzes virtuves un viesistabas logus.
3. Vectēvs ar mazmeitu noguruma dēļ uz skolu gāja ļoti lēni.
4. Tāpēc iešana pa atmiņu pēdām ir iešana pa miroņu valstību.
5. "Neskaitāmiem pavedieniem viņi tos piesaista pie mājas senām nelaimēm un seniem priekiem." [20]
6. "Sveķaina un sausa meža gaisa vietā jūs apdveš zāļu velgā dvaša." [20]
7. "Pie griestiem karājās rūpīgi savērtas vilnas, pakulu un linu dzijas vai arī vēl nesavērtu linu grīstes." [20]
8. "Atmiņa arī nav vairāk nekā cits kā pagājušās dzīves atspīdums mūsu prātā." [20]
9. "No tādiem pat kokiem kā kūts bija taisīta arī klēts, liela ēka divām telpām, vecu un kuplu ābeļu apēnota." [20]
10. "Šajā klētī stāvēja skapji, dzelzīm apkaltas zaļas lādes, tīnes un no saknēm vai liepas koka skaliem pīti meitu pūri, pilni palagiem, deķiem, vilainēm un krekliem." [20]

Pirmais rezultāts ir analizēšanas novērtējums. Tiek apskatīts pilnībā izanalizēto teikumu skaits un pilnas analīzes atkarība pēc vārdu skaita teikumā



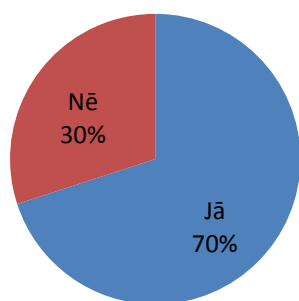
2.3. att., Diagramma atbild uz jautājumu: "Vai diagramma tika izanalizēta pilnībā"



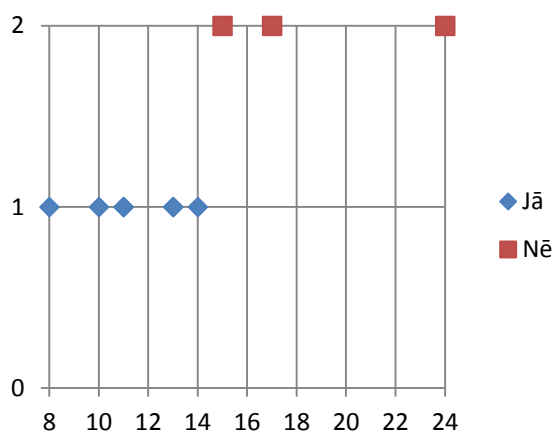
2.4.att., Pilnas analīzes izpilde atkarībā no vārdu skaita teikumā

Pēc šī testa nevar veikt versiju salīdzināšanu pēc pilnām analīzēm, jo sniegtie rezultāti ir vienādi, tāpēc attēloti kopā. Vasara sniedz mazliet labāku rezultātu, jo spēja apturēt analīzi, ziņojot, ka vārds „savērtps” netiek atpazīts, bet, atzīmējot darbības vārdu minēšanu, tā ar analīzi dotajā laikā arī netiek galā. Tiek secināts, ka programmas, lai gan atpazīst vietniekvārdus un partikulas, tomēr nespēj tos sintaktiski iekļaut teikumā. Tiek veiktas teikumu modifikācijas, izņemot partikulas un aizvietojojot vietniekvārdus, un vēlreiz veikta analīze. Teikumu modifikācijas:

- Neskaitāmiem pavedieniem ~~viņi~~ aizgājēju veli ~~tos~~ jaunos saimniekus piesaista pie mājas senām nelaimēm un seniem priekiem.
- Sveķaina un sausa meža gaisa vietā jūs apdveš zāļu velgā dvaša.
- Pie griestiem karājās rūpīgi savērptas vilnas, pakulu un linu dzijas vai ~~arī~~ vēl nesavērptu linu grīstes.
- Atmiņa ~~arī~~ nav vairāk ~~nekas cits~~ kā pagājušās dzīves atspīdums ~~mūsu~~ prātā.
- No tādiem ~~pat~~ kokiem kā kūts bija taisīta ~~arī~~ klēts, liela ēka divām telpām, vecu un kuplu ābeļu apēnota.
- Šajā klētī stāvēja skapji, dzelzīm apkaltas zaļas lādes, tīnes un no saknēm vai liepas koka skaliem pīti meitu pūri, pilni palagiem, deķiem, vilainēm un krekliem.



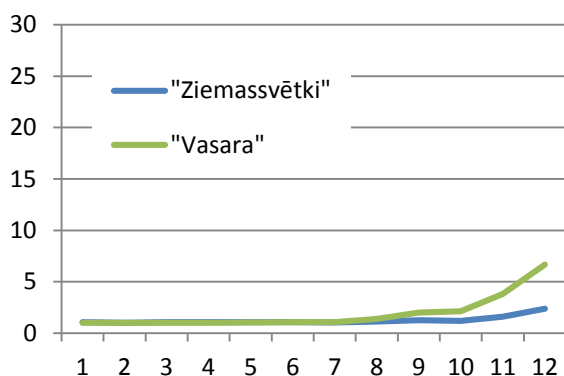
2.5. att., Diagramma atbild uz jautājumu
 "Vai teikums pēc modifikācijas
 tika izanalizēts pilnībā"



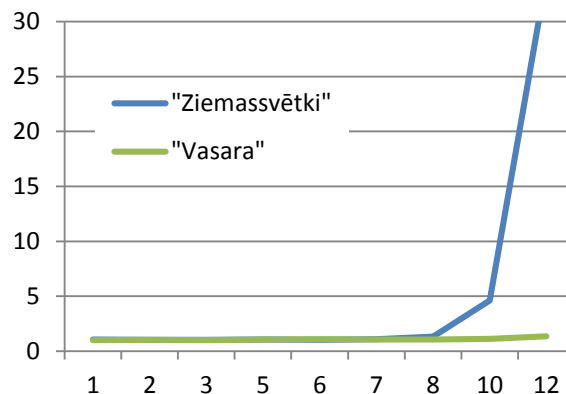
2.6.att., Pilnas analīzes izpilde atkarībā no
 vārdu skaita teikumā
 pēc teikumu modifikācijas

Pēc modifikācijām statistika tiek uzlabota un šī darba ietvaros izmantošanai pieņemts secinājums, ka analizatori nespēja izanalizēt teikumu, ja tā garums pārsniedz 14 vārdus. Var izdarīt arī rupju pieņēmumu, ka 60% teikumu ir nepieciešama modifikācija, un tā palielina pilnas analīzes iespēju par 30%. Divu no grāmatas izvēlēto teikumiem modifikācijas tiek paturētas, lai tiem būtu iespējama pilna analīze un tālākie mērījumi varētu tikt veikti trīs sastādītiem un trīs izvēlētiem teikumiem. Joprojām tiek uzskatīts, ka arī modificētie teikumi ir dabīgi, jo tiek aizvietoti vietniekvārdi ar atbilstošajiem vārdiem no konteksta.

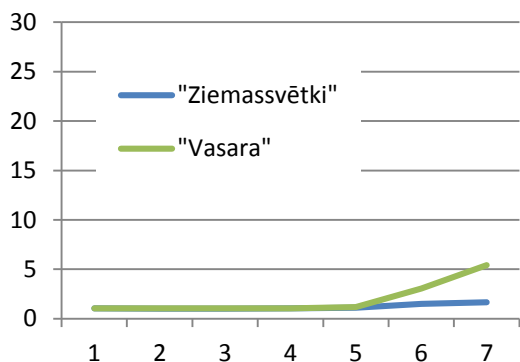
Nākamais apskatītais kritēris ir ātrums. Tiek veidota diagramma katram teikumam ar abu versiju sniegto rezultātu. Lai būtu uzskatāmāk salīdzinājums, diagrammas ir iekļautas mazā formātā un to X ass maksimālā vērtības pieņemta visām vienāda.



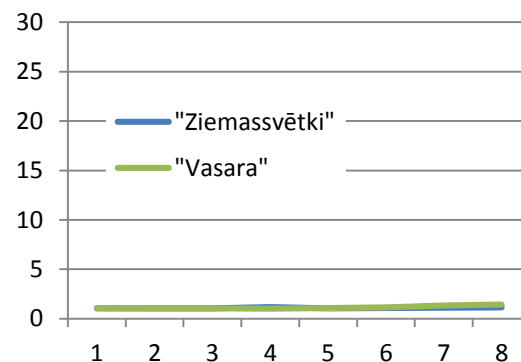
2.7. att., Ātrumu salīdzinājums 1.teikumam



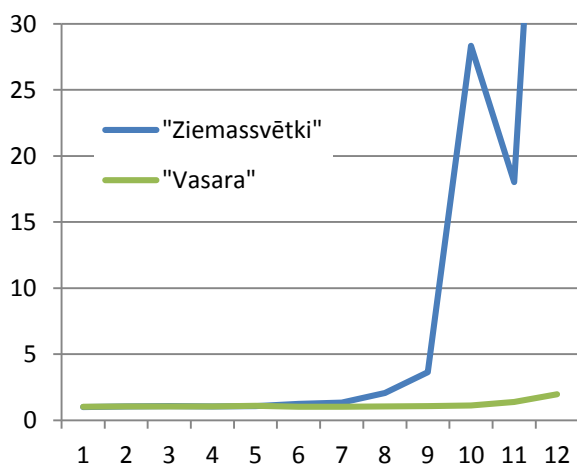
2.8. att., Ātrumu salīdzinājums 2.teikumam



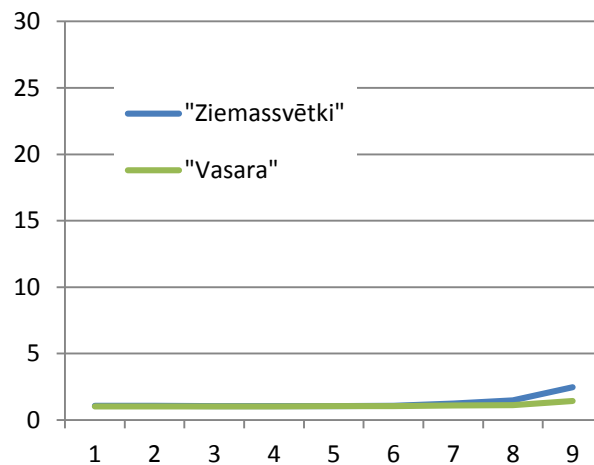
2.9. att., Ātrumu salīdzinājums 3.teikumam



2.10. att., Ātrumu salīdzinājums 4.teikumam



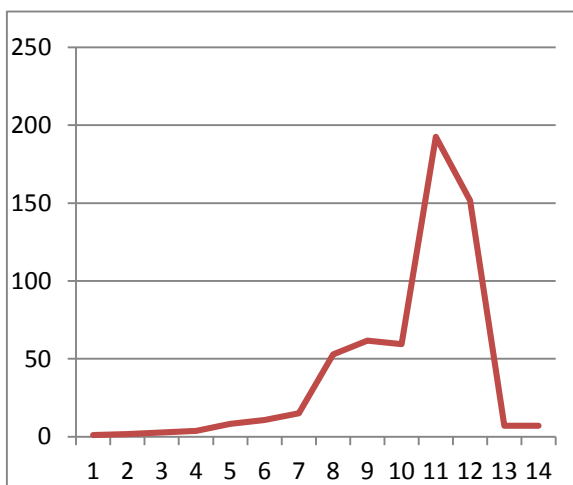
2.11. att., Ātrumu salīdzinājums 5.teikumam



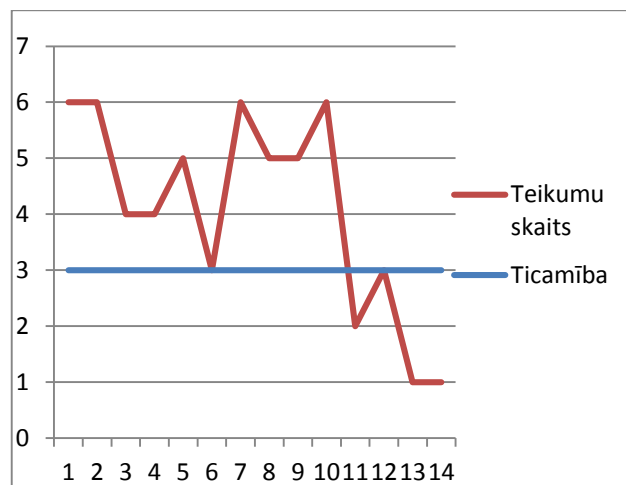
2.12. att., Ātrumu salīdzinājums 6.teikumam

Tieši pusei – 3 teikumiem, Ziemassvētki ir ātrāki un otriem 3 Vasara. Bet, divās no diagrammām - 2. teikuma un 5.teikuma, ātrumu starpība straujais pieaugums - Ziemassvētki kļūst izteikti lēnāki, nevar netikt ņemts vērā. Vasara sniegums ātruma vērtēšanā ir labāks.

Veiksmīgas modernizācijas gadījumā, lai novērtētu programmas spēju atrast pareizo variantu, lietotājam būs jāspēj noteikt vai ir izvēlēts pareizais variants. Tāpēc tiek apskatīts pie kāda vārdu skaita variantu skaits sāk strauji augt. Izveidotās diagrammas attēlo vidējo variantu daudzumu atkarībā no vārdu skaita teikumā un tā ticamību. Rezultāts ir ticams -virs zilās līnijas, ja variantu vidējā summa ir aprēķināta no vismaz 3 teikumiem – vismaz 3 teikumiem ir bijusi analīze ar šādu vārdu skaitu.



2.13. att., Ātrumu salīdzinājums 5.teikumam



2.14. att., Ātrumu salīdzinājums 5.teikumam

2.5. Secinājumi

Lai gan "Čankera" jaunā versija ir devusi labāku rezultātu, salīdzinot ātrumus, rezultāti tiek aplūkoti teikuma pilnas analizēšanas un variantu skaitu analīzes rezultātu kontekstā. No tiem tiek secināts, ka testēšanai ir jāizvēlas teikumi, kas ir īsāki vismaz par 15 vārdiem un, lai rezultātu pareizības vērtēšana nebūtu ļoti sarežģīta, būtu vēlams izvēlēties teikumus, kuri ir līdz 8 vārdiem gari. Šādā kontekstā aplūkojot ātrumu salīdzināšanu, tiek secināts, ka ātrumu strauja atšķirība pieaug pēc 8 vārdiem un, izvēloties aprakstīt garuma teikumus, derētu abas programmas. Bet rezultāts, kas netika aplūkots diagrammas veidā, Vasaras versiju izvirza kā vienīgo kandidātu – analizējot teikumu „Tāpēc iešana pa atmiņu pēdām ir iešana pa miroņu valstību.”[20], projekta versija rezultātos piedāvāja dublikātus, pat četrreiz vairāk rezultātu. Pilna rezultātu tabula aplūkojama 1. pielikumā.

3. ANALIZATORA "ČANKERIS" MODERNIZĀCIJA

Trešajā nodaļā tiek izklāstīta modernizācijas pamatideja un aprakstīta tās realizācija un iekļaušana analizatorā "Čankeris". Nodaļā iekļauti tikai programmas galveno funkciju kodu un to definīcijas ar darbībām, nav iekļautas tās, kas apstrādā tukšus sarakstus vai inicializācijas funkcijas, ja tās veic tikai mainīgā inicializāciju vai vērtības piešķiršanu, tās iespējams aplūkot programmas failos, 4. pielikums. Veikta modernizētās programmas testēšana un rezultātu novērtēšana.

3.1. Piedāvātā metode

Vēl viena metode, kā analizē iekļaut mašīnmācīšanos ir izvadīt noteiktu skaitu "pareizāko variantu", tad rezultātiem pielieto kārtošanas metodi [21]. Tādas parasti izmanto informācijas izgūvē. Piemēram, meklējot dokumentus pēc atslēgas vārdiem, rezultāti tiek kārtoti no atslēgas vārdiem atbistošākā uz mazāk atbilstošiem.[22] "Čankerim" tiek pielāgot modifikācija, kas ļaus izcelt potenciāli pareizo un piešķirt variantiem atšķirīgu ticamību, tā mazinot daudznozīmību. Tiek izmantota uzraudzītā mašīnmācīšanās, kas notiek lietotājam ievadot mācību datus un tā ļaujot mācīties [11], liela izmēra sintaktiski marķēta korpusavēl nav, lai varētu mēģināt iekļaut neuzraudzītu mašīnmācīšanos.

Citos analizators, piemēram, *MSTParser*, ir iespēja padot pareizi marķētu tekstu. Šiem analizatoriem saskarne veidota caur termināli un teikumi netiek padoti dabīgā veidā, ne brīdī, kad programmai tiek apmācīta, ne arī analīzei, informācija atrodama, lejupielādētajam rīkam pievienotajā *README* failā [23]. "Čankeris" izmanto tīmekļa saskarni un teikumi tiek padoti dabīgā veidā un lietotājam nav iespējas padot paša marķētus teikumus. Risinājumā tiek izmantota visu variantu izvade – programma var tikt mācīta atzīmējot pareizo no tās izvadītajiem variantiem. Tas ir arī ērtāk kā izstrādāt īpašus testa datus, analizators teksta marķēšanu jau ir veicis. Pieejas trūkums – situācija, kad analizators nav spējis atrast pareizo risinājumu. Bet šī problēma ir jārisina ar sintaktisko likumu definēšanu, jo analizē nav bijusi veiksmīga tāpēc, ka likumi tādu nepieļauj. Šādu datu manuāla saglabāšana arī neko nedotu, ja likumu dēļ nav iespējams tāds analīzes rezultāts.

Tālāk tiek aprakstīta pareizo analīžu bāzes struktūru. Tiek izvēlēta forma, kādā vislabāk saglabāt datus, lai tie būtu pēc iespējas ērtāk izmantojami un realizācija nebūtu ļoti sarežģīta. Rezultāts tiek saglabāts sadalīts pa pāriem [galvenais vārds[sintaktiskā loma, paskaidrojošais

vārds]]. Tas ir ērti, jo teikumā katram vārdam ir piekārtots vārds, ko tas paskaidro, izteicējs paskaidro teikuma beigu pieturzīmi, tāpēc būtībā teikums ir pāru saraksts, kas tiek sakārtots programmā hierarhiski vienā sarakstā, lai varētu iegūt koka struktūru. Nākamajās analīzēs tiek pārbaudīts, vai tās vārdu pāri jau ir pieejami zināšanu bāzē, un skaitīts pāru skaits, kas sakrīt, tā piekārtojot katram teikumam ticamības koeficientu – atrasto pāru skaitu.

3.2. Pāru meklēšana pēc vārdu līdzības

Saglābājot tikai konkrētus pārus, faila izmēram būtu jābūt diezgan liels, lai tā ietekme būtu jūtama. Lai ļautu sasniegt iespaidīgākus rezultātus ar mazu failu, tiek pievienots fails ar vārdu reprezentācijām vektoros, kas ļaus arī mācību faila vārdiem jēdzieniski tuvus vārdus atpazīt, kā mācību failos jau iekļautus,[24].

Fails ar vārdu attēlojumu vektoros ir iegūts ar *word2vec* rīka palīdzību, kas no teksta korpusa izgūst šādu failu – *embedding.pl*, fails atrodams 4.pielikumā.[25] Latviešu valodai tā pielietota pirmo reizi tikai pirms dažām nedēļām, kad tika izveidots izmantotais fails. Tas iegūts ar *word2vec* rīku un Latviešu valodas līdzsvarotā korpusa[26].

Līdzība starp vārdiem tiek aprēķināta, izrēķinot kosinusu starp vārdu vektoriem[27]. Ja tiktu meklēts vārds Latvija, kā tuvi vārdi tiktu atpazīti citu valstu nosaukumi – Lietuva, Igaunija, Austrālija.

3.3. Realizācija

Realizācija sastāv no divām daļām – mācību datu saglabāšana un mācīšanās. Programmas saskarne tiek papildināta ar vēl vienu teksta lauku, un pogu "Pareizais, attēls 3.1.

Latviešu valodas gramatikas analizators "Čankeriš"

LU MII :: SemTi-Kamols :: vers. 2014.07.11 ^{alfa} (Vasaras versija ar ARC-EAGER)

Teikums:

Analīze: Lietot morf. analizatoru Meklēt deminutīvus Ātļaut vokatīvu Noklusētās pazīmes Visi varianti Pilna analīze

Izdrukas: Vienkāršot sintaktiskās lomas Pazīmju kodi Morf. analizatora izsaukumi Rezultāti txt formātā

Miņšana: Lietvārdi Darbības vārdi Divdabji Īpašības vārdi

3.1.att., "Čankera" lietotāja saskarne pēc modernizācijas

Teksta laukā lietotājs ievada pareizā varianta numuru un ar pogu iniciē veic rezultāta saglabāšanu. Analizējto jaunu teikumu, jāpārlicinās, ka varianta laukā ir 0, citādi tiek uzrez saglabāts jaunās analīzes attiecīgais variants.

Pareizais teikums ar funkciju tiek sadalīts pa pāriem un tad saglabāts failā. Tas tiek realizēts ar vairākām funkcijām. Sākumā tiek izsaukta funkcija, kurai padots rezultātu saraksts un pareizās versijas numurs *learn_results_html_write*, attēls 3.2. Ja ir izvēlēts varianta numurs, tiek izsaukta funkcija *cleanx*, attēls 3.4., kas dala rezultātu pa pāriem, *newPair* attēls 3.4., kas pārus saglabā dinamiskajā atmiņā, lai atvieglotu pāru meklēšanu, rēķinot variantu koeficientus. Rezultāts no atmiņas saglabāts failā. Faila ielasīšana tiek paskaidrota pie mācīšanas realizācijas apraksta.

```
learn_results_html_write([R|VerResults],ItemNumber,Version):-
atom_number(Version,X),
X>0,
cleanx(top,Ver,Z),
newPair(Z),
open('fails',write,Stream,[encoding(utf8)]),set_output(Stream),
listing(memory),
close(Stream),
!.
```

3.2.att., "Čankera" funkcija, kas realizē pareizā analīzes rezultātu saglabāšanu failā

```
cleanx(HLEMMA,IN,CC):-append(A,[[[x_.,LEMMA_.,_.,F_]B].IN],!,
append(A,B,C),
cleanxx(LEMMA,C,CC).

cleanx(HLEMMA,IN,[[[HLEMMA,LOMA,LEMMA]CC])):-
append(A,[[[LOMA_.,LEMMA_.,_.,F_]B].IN],
append(A,B,C),
cleanxx(LEMMA,C,CC).

cleanxx(LEMMA,[A|AA],BBB):-cleanx(LEMMA,A,B),
cleanxx(LEMMA,AA,BB),
append(B,BB,BBB).
```

3.3.att., "Čankera" funkcija, kas realizē pareizā analīzes sadalīšanu pāros

```
newPair([Pair|List]):-assertz(memory(Pair)),
newPair(List).
```

3.4.att., "Čankera" funkcija, kas realizē pāru saglabāšanu dinamiskajā atmiņā

Tā kā mācīšanās laikā programm strādā ar atmiņu, nevis ar pareizo pāru failu, tad, atverot programmu, faila saturs tiek ielasīts atmiņā. Koeficienti tiek rēķināti ar funkciju *koef*, attēls 3.5, tajā tiek atrasts arī lielākais koeficients. Ja ir vienādi koeficienti, tiek ņemts pirmais, tā arī izmantojot jau iekļauto uz morfoloģiju balstīto mācīšanos. Tālākās lietošanas ērtībām koeficienti tiek piekārtoti katram variantam kā pirmais elements, un kā visu rezultātu pirmais elements tiek pielikts varianta ar visu augstāko koeficientu numurs. Funkcija rekursīvi pārbauda visus rezultātus rezultātu sarakstā. Tie tiek sadalīti pa pāriem ar jau apskatīto *cleanx* funkciju, tad ar *nPair*, attēls 3.6., tiek izsaukta funkcija, kas izsauc funkciju *mem*, attēls 3.7, kas palielina koeficientu par 1, ja vārdi tiek atrasti uzreiz mācīšanās failās vai par līdzības koeficientu, ja tiek izmantots *embedding* fails. Līdzības koeficients tiek rēķināts visos gadījumos, kad pāris netiek uzreiz atrasts atmiņā. Ja viens no vārdiem ir, tad tā tiek rēķināta līdzība starp teikuma vārdu un otru faila vārdu. Tiek pārbaudītas abas situācijas, ka galvenais vārds neatbilst pāru failā esošajam un ka paskaidrojošais vārds neatbilst. Kā arī gadījums, ka neviens no teikuma vārdiem nav atmiņā.

```

koef([Result | Results], ItemNumber, [Result1 | ResultsWK], Max, K) :-
    Result=[Ver|_],
    cleanx(top, Ver, Z),
    flag(skait, _, 0),
    nPair(Z),
    flag(skait, Koef, Koef),
    append([Koef], Result, Result1),
    (Koef>Max -> New=Koef, L=ItemNumber; New=Max, L=K),
    NextItemNumber is ItemNumber + 1,
    koef(Results, NextItemNumber, ResultsWK, New, L).

```

3.5.att., "Čankera" funkcija, kas realizē visu variantu koeficientu aprēķināšanu

```

nPair([]).
nPair([Pair|List]):- mem(Pair), nPair(List).

```

3.6.att., "Čankera" funkcija, kas nodrošina visu pāru meklēšanu atmiņā

Atšķirībā no visām pārējām modernizācijas funkcijām, koda daļu , kas veic vārdu līdzības pārbaudīšanu izstrādāja profesors Dr.sc.comp. Guntis Bārzdiņš. embedding fails tika iegūts tikai 14. maijā un, lai varētu pētījumā iekļaut arī tā darbības novērtēšanu, bija nepieciešams ātri faila integrācija.

```

mem([BaseForm,[A,B],BaseForm2):-memory([BaseForm,[A,_],BaseForm2]),
flag(skait,Koef,Koef+1),!.

mem([BaseForm,[A,B],BaseForm2):-memory([BaseForm,[A,_],_]),
e(BaseForm2,V2),findall(ZZ,(memory([BaseForm,[A,_],BaseForm3]),e(BaseForm3,V3)
,cossim(V2,V3,ZZ)),Z),max_list(Z,R),flag(skait,Koef,Koef+R),!.

mem([BaseForm,[A,B],BaseForm2):-memory([_,[A,_],BaseForm2]),
e(BaseForm,V2),findall(ZZ,(memory([BaseForm3,[A,_],BaseForm2]),e(BaseForm3,V3)
,cossim(V2,V3,ZZ)),Z),max_list(Z,R),flag(skait,Koef,Koef+R),!.

mem([BaseForm,[A,B],BaseForm2):-e(BaseForm,V0),e(BaseForm2,V2),
findall(ZZ,(memory([BaseForm1,[A,_],BaseForm3]),e(BaseForm1,V1),e(BaseForm3,V3),
cossim(V2,V3,Z2),cossim(V0,V1,Z1),ZZ is Z1*Z2),Z),
max_list(Z,R),flag(skait,Koef,Koef+R),!.

```

**3.7.att., "Čankera" funkcija, kas realizē pāru meklēšanu dinamiskajā atmiņā,
arī izmantojot vārdu līdzību**

Rezultāti tiek izvadīti pievienojot katram variantam tā koeficientu. Variantam ar lielāko koeficientu , tas ir izcelts treknrakstā. Vēstures uzgalabāšanai ir izveidots fails, "testResult", kurā pareizajam variantam tiek saglabāts, pirmais izvadītais variants, pareizākais variants, pareizais variants un izvadīto variantu skaits.

4. MODERNIZĒTĀ "ČANKERA" NOVĒRTĒŠANA

Ceturtajā nodaļā tiek aprakstīts testēšanas plāns, kā tiek izvēlēti un sagatavoti testpiemēri, parādīts kā tiks reģistrēti rezultāti, rezultāti aprakstošas diagrammas un no tām veiktie secinājumi.

4.1. Testēšanas plāns

Veiksmīgai un pēc iespējas precīzākai testēšanai testa teikumi tiek izvēlēti no dabīga teksta. Pēc otrā nodaļā izdarītajiem secinājumiem testpiemēru teksti tiek ņemti no grāmatas, kas paredzēta, lai bērni mācītos lasīt - Mārītes Oses "Protu, protu!". Tā satur īsus stāstiņus, kas pārsvarā veidoti no vienkāršiem paplašinātiem un vienkāršiem nepaplašinātiem teikumiem. Tiek analizēti visi teikumi pēc kārtas, arī salikti teikumi - statistikai. Tiešās runas tiek sadalītas divos teikumos – piebilde ir viens, tiešā runa – otrs teikums. Ja teikumos ir vietniekvārdi, partikulas vai vārdi vecās izloksnēs, tiek veikta modifikācija un atkārtota analīze. Visas modifikācijas arī tiek fiksētas. Tiek veikta analīze 174 teikumiem, lai būtu veikta pilnīga analīze vismaz 100 teikumiem un stāsts tiktu izanalizēts pilnībā, rezultāti itek fiksēti tabulā ar struktūru kā parādīts 4.1 tabulā.

4.1. tabula

Modernizācijas rezultātu tabulas ststrukturā[4]

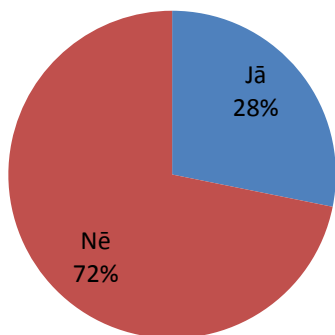
Teikums	Pilna analīze	Ieteiktais	Pareizais	Rezultāts	Koeficients

Tiek fiksēts vai ir veikta pilna analīze, analīzes variantu skaits, "pareizākā" varianta numurs "Ieteiktais", pareizais variants – 0, ja programma nav spējusi piedāvāt pareizu rezultātu. Kolonnā „Rezultāts” tiek atbildēts uz jautājumu "Vai programma norādīja pareizo variantu?", blakus kolonna Lai varētu noteikt vārdu salīdzināšanas izmantošanu tiek veikta arī "pareizākā" varianta koeficienta fiksēšana.

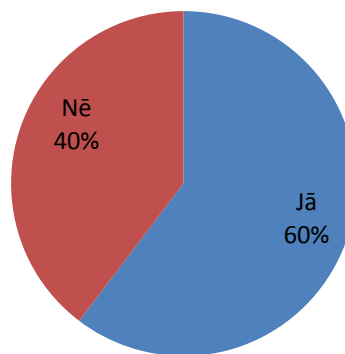
4.2. Rezultātu analīze

Sākumā tiek vērtēta programmas darbība kopumā, salīdzinot 2.nodaļas analīzē iegūtos rezultātus par pilnas analīzes veikšanu ar rezultātiem, kas iegūti testā ar lielāku testa datu kopu. Pēc tam veikta modernizācijas vērtēšana. Pilns testa rezultātu fails pieejams 3.pielikumā.

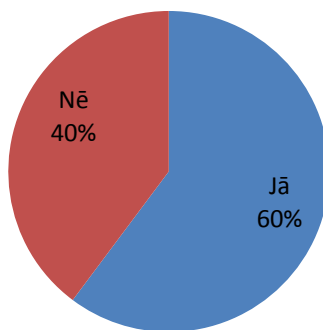
Tiek apskatīts pilnībā izanalizēto teikumu skaits pirms modifikācijām un pēc modifikācijām. Oriģinālie un modificētie teikumi parādīti 2.pielikumā.



4.1. att., Diagramma atbild uz jautājumu
"Vai pirms teikuma modifikācijas
teikumam bija iespējama pilna analīze?"
par visiem analizētajiem teikumu



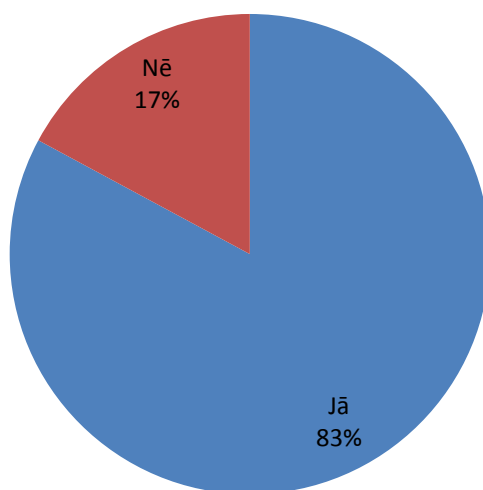
4.2. att., Diagramma atbild uz jautājumu
"Vai pēc teikumu modifikācijas
teikumam bija iespējama pilna analīze?"
par visiem analizētajiem teikumiem



4.3.att., Diagramma atbild uz jautājumu
"Vai pēc modifikācijas, teikumam bija iespējama pilna analīze?"
par modificētajiem teikumiem

Otrajā nodaļā pilna analīze pirms modifikācijas bija 40% teikumu rādītājs ir krities par 12% - 28% izanalizēti, 4.1 attēls. Uzlabojums pēc modifikācijas ir gandrīz tāds pats – iepriekš bija 30%, tagad ir 32%, pēc modifikācijas izanalizēti 60%, attēls 4.2. Skaitļos: no 174 teikumiem pirms modifikācijas tika izanalizēti 49 -28%, neizanalizēti 125-72%, modifikācija bija iespējama 93 teikumiem, 60%, attēls 4.3, jeb 56 teikumiem pēc tās tika veiksmīgi veikta pilna analīze, rezultātā, palielinot kopējo pilnībā izanalizēto teikumu skaitu par 30% - līdz 105.

Tālāk gandrīz visa analīze notiek ar pilnībā analizētajiem teikumiem. Tiek apskatīta analizatora analizēšanas spēja, izveidojot diagrammu par gadījumiem, kad "Čankeris" ir atradis pareizo analīzes rezultātu.



4.4.att., Diagramma atbild uz jautājumu

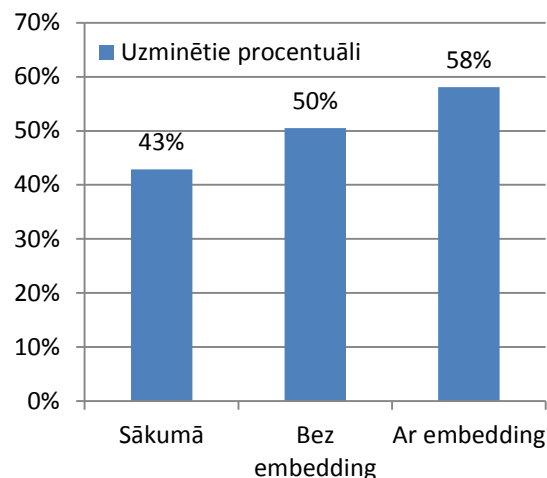
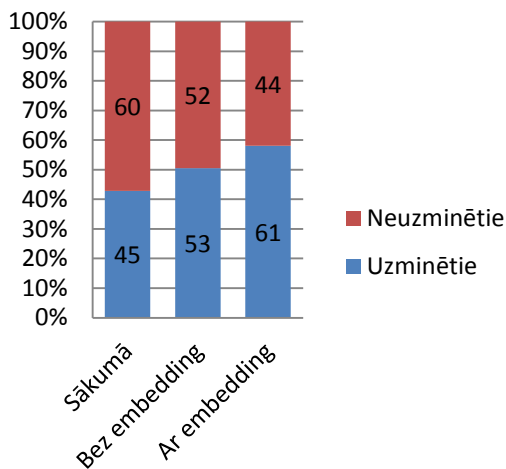
"Vai teikuma analīzē tika piedāvāts pareizs analīzes variants?"

Analizators, kā redzams 4.4 attēlā, 83% gadījumos ir spējis piedāvāt pareizu analīzes rezultātu. Šie skaitļi attēlo arī teikumu skaitu, kas tiek iemācīti analizatoram.

Tiek ņemts vērā fakts, ka par rezultāta pareizību spriež autore, kurai nav specializēta izglītība šajā jautājumā. Lai skaitļi būtu pilnībā korekti, šo testu būtu jāveic cilvēkam ar padziļinātām zināšanām par sintaktisko analīzi ar konsultāciju iespēju no izstrādātāju puses, kas spēj izskaidrot iekšējai lietošanai domātos apzīmējumus un to pielietojumu.

Modernizācijas rezultātā uzlabojums tiek mērīts pēc mācīšanās faila izveides un tad pievienojot arī embedding failu. Kopumā koeficienti tika piekārtoti 115 teikumiem, tiek skaitīti

arī teikumi, kas netika pilnībā analizēti. 47 teikumiem neizmantojot embedding failu, 108 – ar embedding failu, skaitļos teikumi pārklājas. Tiek apskatīts kopumā, cik teikumiem tika pareizi atrasts rezultāts, neizmantojot nevienu no pievienotajām modernizācijām, izmantojot tikai tiešos pārus un izmantojot klāt vārdu līdzības meklēšanu.

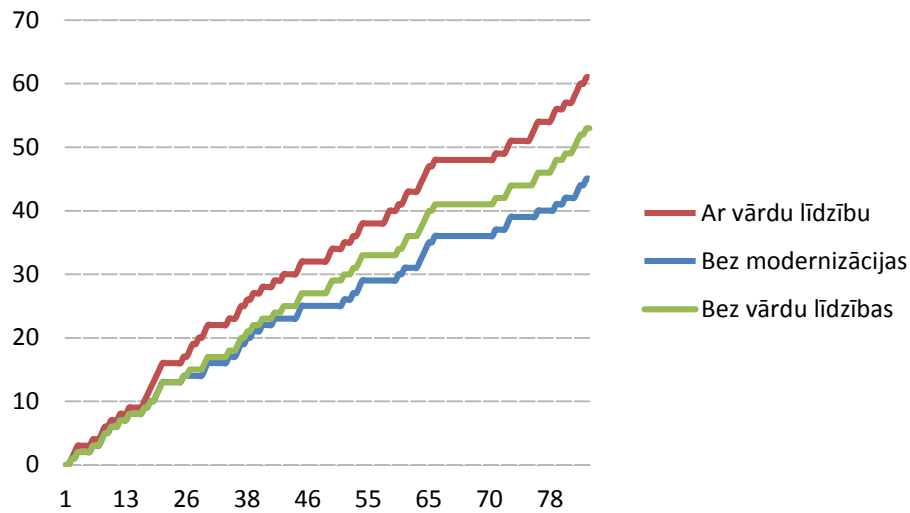


4.5.att., Diagramma parāda pareizi ieteikto un nepareizi ieteikto rezultātu skaitu sadalījumu, ja netiek izmantota modernizācija, bez vārdu līdzības meklēšanas un ar to

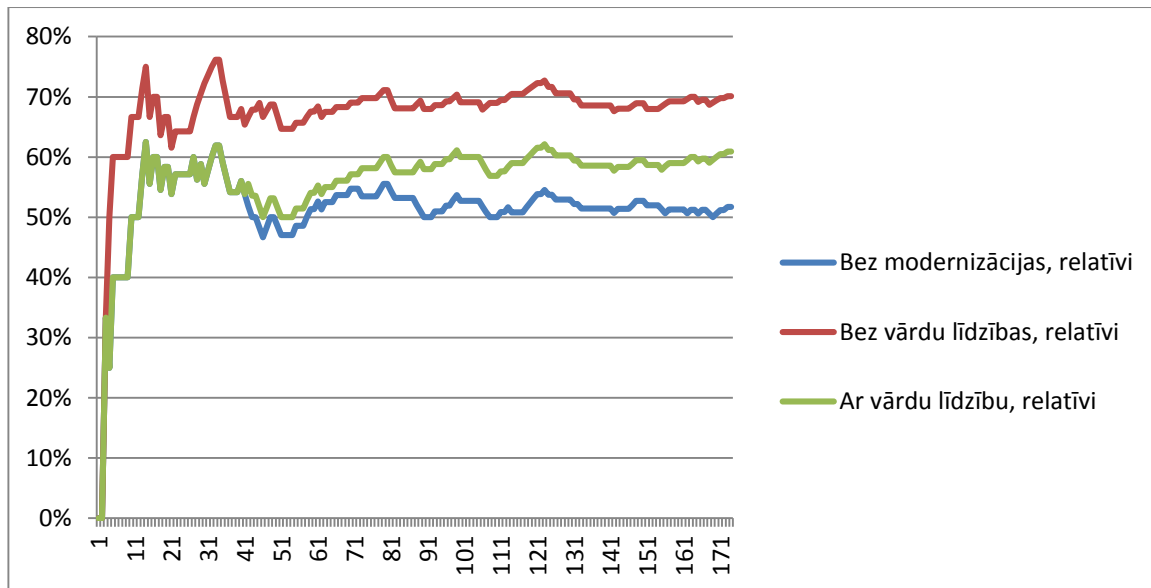
4.6. att., Diagramma parāda pareizi ieteikto rezultātu skaitu procentuāli, ja netiek izmantota modernizācija, bez vārdu līdzības meklēšanas un ar to

Pēc 4.5 attēla un 4.6 attēla var secināt, ka vidēji ar pāru mācīšanos programma tiek uzlabota par 10% un pievienojot arī vārdu līdzības meklēšanu vēl par 10% - kopumā programmas darbības uzlabota par ~20 procentiem. Šie rezultāti, kā 4.5. attēla skaitļos var redzēt, no visiem analizētajiem teikumiem.

Tiek vērtēts arī pareizi izvēlēto variantu pieaugums atkarībā no iemācīto teikumu daudzuma.



4.7. att., Ieteikto rezultātu skaita pieaugums pret iemācīto teikumu skaita, ja netiek izmantota modernizācija, bez vārdu līdzības meklēšanas un ar to



4.7. att., Ieteikto rezultātu skaits relatīvi pret iemācīto teikumu skaita, ja netiek izmantota modernizācija, bez vārdu līdzības meklēšanas un ar to

Kā 4.6. attēlā var redzēt pareizo ieteikto rezultātu skaits pieaug, izmantojot vārdu līdzības meklēšanas iekļaušanu sākumā daudz straujāk, pēc tam pieaugums izlīdzinās. Izmantojot tikai pāru failu pieaugumam sākumā ir tāds pats kā bez modernizācijas, bet iemācīto teikumu skaitam

pieaugot starpība arī pieaug, līdz visu trīs versiju starpība arī iemācīto teikumu skaitam pieaugot ir gandrīz konstant. Apskatot rezultātus relatīvi, 4.7. attēlā redzams, ka līdzības faila izmantošana ļoti ātri un salīdzinoši vienmērīgāk sasniedz savu maksimālo rezultātu ~70 pareizību, pie 71 iemācīta teikumu. Pirmo reizi 70% sliekšnis tiek sasniegts, kad ir iemācīti 10 teikumi. Izmantojot tikai pāru meklēšanu rezultāts sākumā ir identisks ar bez modernizācijas programmu, bet arī pie 71 iemācītā teikuma, tas sasniedz savu maksimumu 60% pareizību. Kritums pēc 41 iemācīta faila, gan ir lielāks kā izmantojot līdzību meklēšanu, bet mazāks kā nemodernizētajai versijai, kura pēc tam vairs nesasniedz modernizētās programmas rezultātu un paliek pie apmēram 50% pareizības. Var secināt, ka vārdu līdzības meklēšana ir divkārtējusi konkrēto pāru uzlabojumu – ar pāriem uzlabojumiem ir 10%, pievienojot arī līdzību meklēšanu, divreiz lielāks -20%.

Pārsteidzoši ir tas, ka rezultāti pēc tam paliek nemainīgi, tiek sagaidīts, ka abu modernizēto versiju pareizi ieteiktie rezultāti turpinās vēl pieaugt palielinot failus, problēma varētu būt, ka tiek saglabāts par maz informācijas par vārdu, tikai vārda pamatforma. Veicot testu tika ievērots, ka vismaz divās reizēs visaugstākais koeficients bija diviem rezultātiem ar vienīgo atšķirību darbības vārda morfoloģiskajos atribūtos – vai darbības vārds ir pārejošs vai nepārejošs un tādā situācijā tiek izvēlēts pirmais variants. Tāpat rezultātu iespējams varētu uzlabot pašiem vārdu pāriem pievienojot ticamību, šobrīd var gadīties situācija, ka tiek saglabāti vārdu pāri ar atšķirīgām sintaktiskajām lomām un tad to analīzes variantiem, balstoties tikai uz šo pāri, būs vienāds rezultāts. Ļoti iespējams, ka kāds no vārdu pāriem tiek izmantots daudz biežāk un tad ir arī daudz lielāka iespējamība, ka tieši tas analīzes variants būs pareizāks.

SECINĀJUMI

- Mēģinājums modernizēt latviešu valodas sintaktiskā analizatoru "Čankeris", iekļaujot tajā mašīnmācīšanos ir veiksmīgs – programma pēc lietotāja saglabātiem pareiziem rezultātiem, spēj izdarīt secinājumus nākamajās analīzēs.
- Jaunākās versijas analizatora analīzes pieeja tiek atzīta par veiksmīgāku, tā sniedz venmērīgu rezultātu analīzes ātrumā un netiek konstatēti gadījumi ar vienādu rezultātu atkārtošanos.
- Programmas daudznozīmība tika veiksmīga samazināta, piešķirot analīzes variantiem koeficientus atšķirībā no to ticamības.
- Modernizācijas rezultātā programmas ieteiktais rezultāts par 20% biežāk ir pareizs.
- Vārdu līdzības meklēšanas iekļaušana divkāršoja tikai ar konkrētu pāru meklēšanu iegūto rezultātu.
- Pareizi ieteikto variantu skaits relatīvi pēc ~70 iemācīto teikumu skaita vairs neaug, ieteiktais rezultāts ir pareizs apmēram 70% gadījumu.
- Esošo rezultātu iespējams var uzlabot saglabātajiem vārdiem, pievienojot morfoloģiskos atribūtus un vārdu pāriem ticamības koeficientu.

IZMANTOTĀ LITERATŪRA UN AVOTI

1. A. Spektors "Latviešu valodas datorfonda izveide", "Latvijas Zinātņu akadēmijas vēstis. Sērija A", *Nr 2, 2001*, 74- 82. lpp.
2. SemTi Kamols projekta latviešu valodas gramatiskais analizētājs [tiešsaiste].[atsauce 01.05.2015]. Pieejams: <http://www.semti-kamols.lv/chunker-doc.pdf>
3. Latviešu valoda digitālajā vidē: datorlingvistika, Informatīvi izglītojoša semināru cikla materiāli [tiešsaiste]. Rakstu krājums. Rīga : LVA, 2012. 186 lpp. Pieejams: http://www.valoda.lv/Petijumi/Valodas_situacijas_izpete/mid_510. ISBN 978-9984-815-57-2
4. Akadēmiskā terminu datubāze AkadTerm [tiešsaiste].[atsauce 01.05.2015]. Pieejams: <http://termini.lza.lv/term.php?term=sintaktisk%C4%81%20anal%C4%ABze&list=anal%C4%A Bze&lang=LV>
5. Virtuālā skola/latviešu valoda [tiešsaiste]. [atsauce 01.05.2015]. Pieejams <http://valoda.ailab.lv/latval/vidusskolai/SINTAKSE/virslocekli.htm>
6. Virtuālā skola/latviešu valoda [tiešsaiste]. [atsauce 01.05.2015]. Pieejams <http://valoda.ailab.lv/latval/vidusskolai/SINTAKSE/paligloc.htm>
7. Virtuālā skola/latviešu valoda [tiešsaiste]. [atsauce 01.05.2015]. Pieejams <http://valoda.ailab.lv/latval/vidusskolai/morfol/vardsk.htm>
8. MaltParser, User guide [tiešsaiste].[atsauce 01.05.2015]. Pieejams <http://www.maltparser.org/intro.html>
9. StanfordParser online demo parser [tiešsaiste].[atsauce 01.05.2015]. Pieejams [.http://nlp.stanford.edu:8080/parser/](http://nlp.stanford.edu:8080/parser/)
10. MateTool online demo parser [tiešsaiste].[atsauce 01.05.2015]. Pieejams <http://en.semepar.ims.uni-stuttgart.de/>
11. P. Lison, "An introduction to machine learning", INF5820 course, Language Technology Group (LTG) Department of Informatics, Univ. Of Oslo, 2012.
12. N. Gruzītis, "Tekstu analīze", Specseminārs Web science, Latvijas Universitāte, 2012.
13. Penn Tree Bank [tiešsaiste].[atsauce 03.05.2015]. Pieejams <https://www.cis.upenn.edu/~treebank/>
14. UL IMCS Latvian Treebank [tiešsaiste].[atsauce 03.05.2015]. Pieejams <http://eksperimenti.ailab.lv/tred/index.html>

15. SemTi Kamols projekta latviešu valodas morfoloģisko pazīmju kopa [tiešsaiste].[atsauce 01.05.2015]. Pieejams: http://www.semti-kamols.lv/doc_upl/TagSet.pdf
16. G. Barzdins, N. Gruzitis, G. Nespore, B. Saulite (2007). "Dependency-Based Hybrid Model of Syntactic Analysis for the Languages with a Rather Free Word Order", Joakim Nivre, Heiki-Jaan Kaalep, Kadri Muischnek and Mare Koit (Eds.) (NODALIDA 2007) Conference Proceedings, *May 2007*, pp. 13-20.
17. T.Sharma, S. Das, V. Bhalla, "Parsing – a brief study", *International Journal of Research (IJR)*, Vol-1, Issue-8, September 2014.
18. P.M. Nugues, "An introduction to Language Processing with Perl and Prolog", Springer, 2006, p. 303-310.
19. J. Nivre, D. Fernandez-Gonzalez, "Arc-Eager Parsing with the Tree Constraint", *Computational Linguistics*, Vol. 40, No. 2, June 2014, pp259-267.
20. E. Virza, "Straumēni", Rīga, Liesma, 1989, 13-30lpp.
21. F. Sangati, W. Zuidema, R. Bod, "A generative re-ranking model for dependency parsing", Proceedings of the 11th International Conference on Parsing Technologies (IWPT-09: 7-9 October 2009), 2009, pp. 238-241.
22. H. Li, "A Short Introduction to Learning to Rank", *Ieice trans. Inf. & syst.*, Vol.e94-d, No.10, October 2011.
23. MSTParser README fails [tiešsaiste].[atsauce 02.05.2015]. Pieejams <http://www.seas.upenn.edu/~strctrln/MSTParser/README>
24. P.D. Turney, P. Pantel, "From Frequency to Meaning: Vector Space Models of Semantics", *Journal of Artificial Intelligence, Research* 37, 2010, pp141-188
25. Summary about word2vec tool [tiešsaiste].[atsauce 14.05.2015]. Pieejams <https://code.google.com/p/word2vec/>
26. Līdzsvarots mūsdienu latviešu valodas tekstu korpuss [tiešsaiste].[atsauce 01.05.2015]. Pieejams <http://www.korpuss.lv/>
27. C. Perone, "Machine Learning :: Cosine Similarity for Vector Space Models (Part III)" [tiešsaiste].[atsauce 01.05.2015]. Pieejams <http://blog.christianperone.com/?p=2497>
28. M.Ose, "Protu, protu!", *Zvaigzne*, 1994, 4-30 lpp.

PIELIKUMI

1. pielikums

”Čankera” versiju salīdzināšanas rezultātu tabula

Piemēra nr.	Teikums[20]	Vārdu skaits	Variantu skaits „Ziemassvētkiem”	„Ziemassvētki”	Variantu skaits „Vasarai”	„Vasara”
1	Spēlējās.	1	1	1,06	1	1,02
	Spēlējās kucēni.	2	1	1,01	1	1,02
	Ļāvā spēlējās kucēni.	3	3	1,08	3	1,02
	Dienā ļāvā spēlējās kucēni.	4	9	1,07	9	1,02
	Dienā klātā ļāvā spēlējās kucēni.	5	18	1,06	18	1,05
	Dienā stiebriem klātā ļāvā spēlējās kucēni.	6	27	1,07	27	1,06
	Dienā zāles stiebriem klātā ļāvā spēlējās kucēni.	7	36	1,05	36	1,08
	Dienā zāles stiebriem klātā ļāvā nevainībā spēlējās kucēni.	8	216	1,14	216	1,38
	Dienā zāles stiebriem klātā ļāvā draiskā nevainībā spēlējās kucēni.	9	216	1,26	216	2,01
	Dienā zaļiem zāles stiebriem klātā ļāvā draiskā nevainībā spēlējās kucēni.	10	162	1,21	162	2,14
	Dienā sulīgi zaļiem zāles stiebriem klātā ļāvā draiskā nevainībā spēlējās kucēni.	11	378	1,6	378	3,81
	Vasaras dienā sulīgi zaļiem zāles stiebriem klātā ļāvā draiskā nevainībā spēlējās kucēni.	12	432	2,39	432	6,69
2	Apspīd.	1	1	1,06	1	1
	Apspīd logus.	2	1	1,02	1	1,03
	Apspīd virtuves logus.	3	2	1,04	2	1
	Apspīd virtuves un viesistabas logus.	5	3	1,05	3	1,06
	Gaisma apspīd virtuves un viesistabas logus.	6	4	1,03	4	1,09
	Gaisma apspīd Jāņa virtuves un viesistabas logus.	7	10	1,1	10	1,07
	Vakara gaisma apspīd Jāņa virtuves un viesistabas logus.	8	10	1,32	10	1,06

	Vakara un rīta gaisma apspīd Jāņa virtuves un viesistabas logus.	10	10	4,61	10	1,13
	Vakara un rīta gaisma apspīd Jāņa un Ilzes virtuves un viesistabas logus.	12	16	34,04	16	1,36
3	Gāja.	1	2	1,05	2	1,05
	Vectēvs gāja.	2	2	1,03	2	1,05
	Vectēvs noguruma dēļ gāja.	4	2	1,02	2	1,04
	Vectēvs noguruma dēļ gāja lēni.	5	9	1,06	9	1,04
	Vectēvs ar mazmeitu noguruma dēļ gāja lēni.	7	18	1,1	18	1,19
	Vectēvs ar mazmeitu noguruma dēļ uz skolu gāja lēni.	9	54	1,49	54	3,05
	Vectēvs ar mazmeitu noguruma dēļ uz skolu gāja ļoti lēni.	10	90	1,67	90	5,43
4	ir.	1	1	1,05	1	1,07
	ir iešana.	2	4	1,04	2	1,04
	iešana ir iešana.	3	3	1,04	1	1,04
	iešana pa pēdām ir iešana.	5	8	1,17	2	1,03
	iešana pa pēdām ir iešana pa valstību.	7	16	1,06	4	1,09
	iešana pa pēdām ir iešana pa miroņu valstību.	8	16	1,07	4	1,15
	iešana pa atmiņu pēdām ir iešana pa miroņu valstību.	9	16	1,1	4	1,33
	Tāpēc iešana pa atmiņu pēdām ir iešana pa miroņu valstību.	10	16	1,16	4	1,42
5	Piesaista.	1	1	1,01	1	1,02
	Veļi piesaista.	2	1	1,05	1	1,05
	Veļi piesaista pie nelaimēm.	4	1	1,08	1	1,05
	Veļi piesaista pie nelaimēm un priekiem.	6	1	1,04	1	1,05
	Veļi piesaista pie mājas nelaimēm un priekiem.	7	4	1,06	4	1,09
	Veļi saimniekus piesaista pie mājas nelaimēm un priekiem.	8	4	1,23	4	1,03

	Pavedieniem veļi saimniekus piesaista pie mājas nelaimēm un priekiem.	9	4	1,35	4	1,03
	Pavedieniem veļi saimniekus piesaista pie mājas senām nelaimēm un priekiem.	10	7	2,07	7	1,05
	Pavedieniem veļi saimniekus piesaista pie mājas senām nelaimēm un senienu priekiem.	11	7	3,64	7	1,07
	Neskaitāmiem pavedieniem veļi saimniekus piesaista pie mājas senām nelaimēm un senienu priekiem.	12	7	28,33	7	1,12
	Neskaitāmiem pavedieniem veļi jaunus saimniekus piesaista pie mājas senām nelaimēm un senienu priekiem.	13	7	18,03	7	1,38
	Neskaitāmiem pavedieniem aizgājēju veļi jaunus saimniekus piesaista pie mājas senām nelaimēm un senienu priekiem.	14	7	71,18	7	1,96
6	Apdveš.	1	1	1,06	1	1,03
	Apdveš dvaša.	2	1	1,06	1	1,01
	Vietā apdveš dvaša.	3	3	1,04	3	1,03
	Gaisa vietā apdveš dvaša.	4	3	1,05	3	1,03
	Sveķaina gaisa vietā apdveš dvaša.	5	3	1,04	3	1,05
	Sveķaina un sausa gaisa vietā apdveš dvaša.	7	6	1,06	6	1,04
	Sveķaina un sausa meža gaisa vietā apdveš dvaša.	8	18	1,23	18	1,1
	Sveķaina un sausa meža gaisa vietā apdveš zāļu dvaša.	9	18	1,47	18	1,11
	Sveķaina un sausa meža gaisa vietā apdveš zāļu valgā dvaša.	10	72	2,46	72	1,43
8	Pie griestiem karājās rūpīgi savērptas vilnas, pakulu un līnu dzijas vai arī vēl nesavērptu līnu grīstes.	-	-	-	-	-

9	Atmiņa arī nav vairāk nekā cits kā pagājušās dzīves atspīdums mūsu prātā.	-	-	-	-	-
10	No tādiem pat kokiem kā kūts bija taisīta arī klēts, liela ēka divām telpām, vecu un kuplu ābeļu apēnota.	-	-	-	-	-
11	Klētī stāvēja skapji, dzelzīm apkaltas zaļas lādes, tīnes un no saknēm vai liepas koka skaliem pīti meitu pūri, pilni palagiem, deķiem, vilainēm un krekliem.	-	-	-	-	-

2.Pielikums

”Čankera” modernizācijas testa datu modifikācijas

Teikums[28]	Modificēts teikums
Agra rīta stunda.	
Mazais Jurītis iet uz skolu.	
Uz ielas ir liela rosība.	
Ļaudis steidzas uz darbu.	
Te zēnu uzrunā tēvocis Uldis.	
Dēls, kur tu tik agri?	Kur tik agri?
Vai arī uz darbu?	Vai uz darbu?
Jurītis pasmaida un atbild.	
Uz skolu.	
Bet tēvocis saka.	Tēvocis saka.
Nu, tad jau uz darbu.	Tad jau uz darbu.
Grāmatai nav mutes, taču viņa runā.	Grāmatai nav mutes, taču grāmata runā.
Viņa stāsta pasakas un stāstus.	Grāmata stāsta pasakas un stāstus.
Skolotāja lasa.	
Mēs uzmanīgi klausāmies.	Uzmanīgi klausāmies.
Grāmata runā ar skolotājas muti.	
Es jau protu lasīt.	Jau protu lasīt.
Es lasu uzdoto stāstiņu.	Lasu uzdoto stāstiņu.
Grāmata runā ar manu muti.	Grāmata runā ar skolēna muti.
Skolā ir daudz grāmatu.	
Tās ir arī mūsu klasē.	Grāmatas ir skolēnu klasē.
Audzinātāja saka.	
Mīliet un saudzējiet grāmatas.	
Tramvajs priecājas, ja bērni ir piesardzīgi.	
Tas pagaida, kamēr tie iekāpj.	Tramvajs pagada, kamēr bērni iekāpj.
Pavisam maziem bērniem tramvajs ļauj kopā ar māmiņām iekāpt pa priekšējām durvīm.	Pavisam maziem bērniem tramvajs ļauj kopā ar māmiņām iekāpt pa tramvaja priekšu.
Dažreiz tramvajs steidzas.	
Strādnieki jāved uz rūpnīcu.	
Strādnieki nedrīkst nokavēt darbu.	
Skolēni un skolotāji jāaizved uz skolu.	
Viņi nedrīkst nokavēt pirmo stundu.	Skolēni un skolotāji nedrīkst nokavēt pirmo stundu.
Tramvajs ir ļoti labs.	
Tas rūpējas par mums.	Tramvajs rūpējas par mums.

Kad esam iekāpuši tramvajā, viņš ļauj skatīties pa logu.	Kad esam iekāpuši tramvajā, tramvajs ļauj skatīties pa logu.
Visas mājas tad skrien ātri garām.	Mājas tad skrien ātri garām.
Koki ietves malā arī skrien.	Koki ietves malā arī skrien.
Mums kļūst ļoti jautri.	Kļūst ļoti jautri.
Kad tramvajs pietur pie skolas, mēs izkāpjam.	Kad tramvajs pietur pie skolas, izkāpjam.
Apejam tramvaju gar priekšpusi un dodamies uz skolu.	Apejam tramvaju gar priekšpusi un dodames uz skolu.
Visi bērni raksta.	Bērni raksta.
Viņi grib rakstīt glīti, glīti.	Bērni grib rakstīt glīti, glīti.
Ivars arī grib rakstīt glīti.	Ivars grib rakstīt glīti.
Viņš sēž taisni.	Ivars sēž taisni.
Elkoņi zēnam novietoti uz galda.	
Ivars velk lēni no vienas līnijas līdz otrai.	Ivars velk lēni no vienas līnijas līdz otrai līnijai.
Burti iznāk vienādi un gludi.	
Rindu pēc rindas veidojas rūpīgs rakstu darbs.	Veidojas rūpīgs rakstu darbs.
Ivars smaida.	
Skolotāja arī smaida un saka.	Skolotāja smaida un saka.
Ivar, tu šodien esi nopelnījis piecinieku.	Šodien esi nopelnījis piecinieku
Upe, lietus laikā izkāpdama no krastiem, uzsaukusi strautam.	Upe, lietus laikā izkāpdama no krastiem, esot uzsaukusi strautam.
Kāds tu mazs un tik niecīgs.	Mazs un niecīgs.
Tevi neviens neievēro.	Neievēro. Paliek tikai izteicējs, tiek uzskatīts, ka pilna analīze nav iespējama.
Ar mani ir pavisam cita lieta.	Ar upi ir cita lieta.
Strauts atteicis.	
Nelielies ar savu lielumu un varenumu.	Nelielies ar lielumu un varenumu.
Ja nebūtu mazu strautu, nebūtu arī Tevis.	Ja nebūtu mazu strautu, nebūtu arī upes.
Uz lauka ganījās zosu bars.	
Lapsa to ieraudzīja.	Lapsa baru ieraudzīja.
Viņai iekārojās zosu cepeša.	Lapsao sakārojās zoss cepeti.
Lapsa klusām pielīda pie vienas zoss.	
Te pēkšņi kāds sāka svilpot.	Te pēkšņi krūze sāka svilpot.
Zosis aizbēga.	
Lapsa ieraudzīja krūzi, kurā vējš svilpoja.	
Ak tad tu esi tā spēlmane.	Ak tad esi tā spēlmane.
Lapsa iesaucās.	
Nu, pagaid tikai.	Pagaidi tikai.
Lapsa pieskrēja pie krūzes un uzmauca sev osu kaklā.	Lapsa pieskrēja pie krūzes un uzmauca osu kaklā.
Nu tā nesa krūzi uz upi slīcināt.	Nu nesa krūzi uz upi slīcināt.
Lapsa nonāca pie upes.	

Viņa uzkāpa uz laipas un laida krūzi ūdenī.	Lapsa uzkāpa uz laipas un laida krūzi ūdenī.
Krūze piesmēlās un sāka vilkt lapsu līdz.	
Nu lapsa sāka lūgties.	
Krūzīt, krūzīt.	
Es pa jokam, tu no tiesas.	Lapsa pa jokam, krūze no tiesas.
Bet nelīdzēja nekāda lūgšanās.	
Krūze piesmēlās pilna un ierāva lapsu dziļā ūdenī.	
Dzīvojis reiz puisis.	
Viņš bijis liels koklētājs un dziedātājs.	Puisis bija liels koklētājs un dziedātājs.
Kad viņš koklējis un dziedājis, putni bariem laidušies klausīties.	Kad puisis koklējis un dziedājis, putni bariem laidušies klausīties.
Puisi reiz naktī mežā apmaldījies.	
Viņš iekritis dziļā vilku bedrē.	Puisis iekritis dziļā vlku bedrē.
Dziļumā vilks jau gatavojies koklētājam uzbrukt.	
Puisi sācis koklēt un dziedāt.	
Vilks kaucis dziesmai līdz.	Vilks kaucis dziesmai līdz.
Tas nedarījis nekā ļauna.	Vlks nedarījis nekā ļauna.
Koklētājs gaismiņā izlēcis no bedres un gājis uz mājām.	
Tas pavadījis lielo kolētāju līdz vārtiem.	Vilsk esot pavadījis lielo koklētāju uz māju.
Vilks nemaz negribējis vairs uz mežu atpakaļ griezties.	Vilks nemaz negribēja vairs uz mežu atpakaļ griezties.
Piemājas dārziņā aug trīs ābeles.	
Viena ābolu pilna.	
Zem augļu svara tās zari nolīkuši.	Zem augļu svara ābeles zari ir nolīkuši.
Ābele lūdzas.	
Ņem nost manus smagos ābolus.	Ņem nost smagos ābolus.
Citādi es lūzīšu.	Citādi lūzīšu.
Mikiņš noplūc pašu sārtāko ābolu.	
Viņš ar nazi pārgriež ābolu un nogaršo.	Mikiņš ar nazi pārgriež ābolu un nogaršo.
Zēns nopūšas.	
Sēkliņas vēl nav gluži brūnas.	
Mikiņš atbalsta zarus ar gariem, zarainiem kokiem.	
Nu zari nelūzīs.	
Pa ielu gāja vecenīte un nesa grozu.	
Grozā bija āboli.	
Kaut es dabūtu vienu.	Kaut dabūtu vienu.
Nodomāja Hugo.	
Zēns klusu piezagās no mugurpuses.	
Viņš paņēma no groza ābolu.	Zēns paņēma no groza ābolu.
Hugo to steidzīgi iebāza kabatā.	Zēns ābolu steidzīgi iebāza kabatā.

Vecenīte neko nemanīja.	
Viņa apstājās un sacīja.	Vecenīte apstājās un sacīja.
Panāc šurp, puisīt.	Panāc šurp.
Viņa izvēlējās vislielāko ābolu un sniedza zēnam.	Vecenīte izvēlējās vislielāko ābolu un sniedza zēnam.
Ēd.	
Tas ir no mana dārza.	Ābols ir no vecenītes dārza.
Zēnam likās, ka ābols viņa kabatā kļūst karsts kā uguns.	Zēnam likās, ka ābols kabatā kļūst karsts kā uguns.
Zēns steidzīgi iemeta grozā nozagto ābolu un aizskrēja.	
Hugo atskrēja mājas un sāka raudāt.	
Kādam ķēniņam bija skaista meita.	
Reiz meita saslimst.	
Ķēniņš aicina visādus ārstus.	
Neviens nespēj meitai palīdzēt.	Ārsti nevar meitai palīdzēt.
Pēdīgi kāds saka, ka vajagot zelta ābolu.	Pēdīgi saka, ka vajagot zelta ābolu.
Tas meitu izārstēšot.	Zelta ābols meitu izārstēs.
Ķēniņš izlaiž pa visu zemi ziņu.	Ķēniņš izlaiž pa zemi ziņu.
Kurš zelta ābolu atnesīs, dabūs meitu par sievu.	
Tajā zemē dzīvoja tēvs ar trim dēliem.	Zemē dzīvoja tēvs ar trim dēliem.
Kādu rītu tēvs pieceļas un ierauga durvju priekšā ābeli ar trim zelta āboliem.	Rītā tēvs pieceļas un ierauga durvju priekšā ābeli ar trim zelta āboliem.
Tēvam skaidrs, kas darāms.	
Viņš noplūc ābolus un iedod katram dēlam pa vienam.	Tēvs noplūc ābolus un iedod dēliem pa vienam.
Ejiet, dēli, uz pili un izmēģiniet laimi.	Ejiet uz pil un izmēģiniet laimi.
Pirmais gāja vecākais dēls.	Sākumā gāja vecākais dēls.
Viņš paņēma ceļam līdzi gabalu maizes.	Vecākais dēls paņēma ceļam līdzi gabalu maizes.
Zelta ābolu viņš ielika kabatā.	Zelta ābolu vecākais dēls ielika kabatā.
Dēls iet, iet līdz ieiet biežā mežā.	
Vecākajam dēlam gribās ēst.	
Viņš salūko gabalu maizes un sāk to tiesāt.	Vecākais dēls salūko gabalu mazes un sāk maizi tiesāt.
Te, kur gadījies, kur ne, nāk sirms vecītis.	
Tas lūdz, lai šim iedodot kādu garoziņu.	Vecītis lūdz, lai iedodot kādu garoziņu.
Vecākais dēls ij nedomā ar vecīti dalīties kumosā.	Vecākais dēls nedomā ar vecīti dalīties ar kumosu.
Strādā pats, tad nebūs maize jādiedelē.	
To dzirdēdams, vecītis neko nesaka un pazūd.	Dēlu dzirdēdams, vecītis neko nesaka un pazūd.
Otrā dienā vecākais dēls nonāk ķēniņa pilī.	
Sulaiņi to aizved pie ķēniņa.	Sulaiņi dēlu aizved pie ķēniņa.
Ķēniņš vaicā, ko šis vēloties.	Ķēniņš vaicā, ko dēls vēloties.

Dēls saka, ka esot atnesis zelta ābolu.	
Gribot princesi par sievu.	
Vispirms parādi ābolu.	
Ķēniņš pavēl.	
Vecākais dēls tūlīt bāzīs roku kabatā, lai izvilktu ābolu.	
Tavu brīnumu .	
Ābols kabatā pārvērties par čūsku.	
Ķēniņš par šādu mānīšanos liek iemest vecāko dēlu cietumā.	Ķēniņš par mānīšanos liek iemest vecāko dēlu cietumā.
Nu iet laimi izmēģināt vidējais dēls.	
Bet arī vidējam dēlam klājas gluži tāpat kā vecākajam.	Vidējam dēlam klājas gluži tāpat kā vecākajam.
Nu jaunākais dēls posās uz pili.	
Ceļam viņš iebāza maizi kabatā.	Ceļam jaunākais dēls iebāza maizi kabatā.
Turpat dēls ielika arī zelta ābolu.	Turpat dēls ielika zelta ābolu.
Dēls gāja, gāja līdz nonāca biežā mežā.	
Viņš sadomāja apēst maizi.	Dēls sadomāja apēst maizi.
Tikko dēls sāka ēst, klāt sirms vecītis.	
Viņš lūdz iedot kādu garoziņu.	Vecītis lūdz iedot kādu garoziņu.
Esot ļoti izsalcis.	
Dēls atdeva visu maizi.	
Še, ēd vesels tētiņ.	
Es jau paēdu.	Jau paēdu.
Vecītis pateicās par labo sirdi.	
Arī uz priekšu dari otram tikai labu.	Uz priekšu dari otram labu.
Tad tev dzīvē labi klāsies.	Tad dzīvē labi klāsies.
Otrā rītā dēls nonāca ķēniņa pilī.	
Viņš iedeva ķēniņam zelta ābolu.	Dēls iedeva ķēniņam zelta ābolu.
Ķēniņš nespēja prieku valdīt.	Ķēniņš nevarēja prieku valdīt.
Viņš paņēma zelta ābolu un gāja pie meitas.	Ķēniņš paņēma zelta ābolu un gāja pie meitas.
Meita ieraudzīja ābolu.	
Tūlīt viņa no gultas ārā un veselāka par veselu.	Tūlīt meita no gultas ārā un veselāka par veselu.

3.pielikums

Modernizēta "Čankera" testēšanas rezultātu tabula

Teikums	Pilna analīze	Ieteiktais	Pareizais	Rezultāts	Koeficients
Agra rīta stunda.	Jā	1	4	Nē	0,00E+00
Mazais Jurītis iet uz skolu.	Jā	1	3	Nē	0,00E+00
Uz ielas ir liela rosība.	Jā	1	1	Jā	2,05E-03
Ļaudis steidzas uz darbu.	Jā	3	3	Jā	2,90E-03
Te zēnu uzrunā tēvocis Uldis.	Jā	1	1	Jā	0,00E+00
Kur tik agri?	Jā	1	0	Nē	0,00E+00
Vai uz darbu?	Nē	-	-	-	0,00E+00
Jurītis pasmaida un atbild.	Nē	-	-	-	2,57E-04
Uz skolu.	Nē	-	-	-	0,00E+00
Tēvocis saka.	Jā	1	1	Jā	1,08E-03
Tad jau uz darbu.	Nē	-	-	-	0,00E+00
Grāmatai nav mutes, taču grāmata runā.	Nē	-	-	-	0,00E+00
Grāmata stāsta pasakas un stāstus.	Jā	1	1	Jā	1,55E-03
Skolotāja lasa.	Jā	1	1	Jā	1,29E-03
Uzmanīgi klausāties.	Jā	1	5	Nē	9,10E-04
Grāmata runā ar skolotājas muti.	Jā	1	1	Jā	4,19E-03
Jau protu lasīt.	Jā	1	0	Nē	1,15E-03
Lasu uzdoto stāstiņu.	Jā	1	3	Nē	1,00E+00
Grāmata runā ar skolēna muti.	Jā	1	1	Jā	3,00E+00
Skolā ir daudz grāmatu.	Nē	-	-	-	0,00E+00
Grāmatas ir skolēnu klasē.	Jā	1	3	Nē	1,00E+00
Audzinātāja saka.	Jā	1	1	Jā	1,00E+00
Mīliet un saudzējiet grāmatas.	Nē	-	-	-	0,00E+00
Tramvajs priecājas, ja bērni ir piesardzīgi.	Nē	-	-	-	1,34E-01
Tramvajs pagada, kamēr bērni iekāpj.	Nē	-	-	-	1,34E-01
Pavisam maziem bērniem tramvajs ļauj kopā ar māmiņām iekāpt pa tramvaja priekšu.	Jā	241	0	Nē	2,25E-03
Dažreiz tramvajs steidzas.	Jā	1	1	Jā	1,00E+00
Strādnieki jāved uz rūpnīcu.	Jā	3	3	Jā	3,59E-03
Strādnieki nedrīkst nokavēt darbu.	Jā	1	1	Jā	0,00E+00
Skolēni un skolotāji jāaizved uz skolu.	Jā	2	2	Jā	3,78E-03
Skolēni un skolotāji nedrīkst nokavēt pirmo stundu.	Jā	1	1	Jā	1,00E+00

Tramvajs ir ļoti labs.	Jā	1	1	Jā	0,00E+00
Tramvajs rūpējas par mums.	Jā	1	1	Jā	2,38E-03
Kad esam iekāpuši tramvajā, tramvajs ļauj skatīties pa logu.	Nē	-	-	-	2,51E-08
Mājas tad skrien ātri garām.	Jā	6	11	Nē	1,50E-03
Koki ietves malā arī skrien.	Jā	7	6	Nē	1,00E+00
Kļūst ļoti jautri.	Jā	6	3	Nē	2,55E-03
Kad tramvajs pietur pie skolas, izkāpjām.	Jā	1	0	Nē	0,00E+00
Apejam tramvaju gar priekšpusi un dodames uz skolu.	Nē	-	-	-	0,00E+00
Bērni raksta.	Jā	1	1	Jā	2,47E-03
Bērni grib rakstīt glīti, glīti.	Jā	1	2	Nē	1,00E+00
Ivars grib rakstīt glīti.	Jā	2	2	Jā	1,00E+00
Ivars sēž taisni.	Jā	3	3	Jā	1,01E-03
Elkoņi zēnam novietoti uz galda.	Jā	1	0	Nē	1,08E+00
Ivars velk lēni no vienas līnijas līdz otrai līnijai.	Jā	15	31	Jā	4,57E-03
Burti iznāk vienādi un gludi.	Jā	1	7	Nē	5,16E-04
Veidojas rūpīgs rakstu darbs.	Jā	1	1	Jā	7,92E-04
Ivars smaida.	Jā	1	1	Jā	1,93E-03
Skolotāja smaida un saka.	Nē	-	-	-	6,76E-07
Šodien esi nopelnījis piecinieku	Jā	1	10	Nē	0,00E+00
Upe, lietus laikā izkāpdama no krastiem, esot uzsaukusi strautam.	Jā	1	10	Nē	5,97E-07
Mazs un niecīgs.	Nē	-	-	-	0,00E+00
Neievēro. Paliiek tikai izteicējs, tiek uzskatīts, ka pilna analīze nav iespējama.	Nē	-	-	-	0,00E+00
Ar upi ir cita lieta.	Nē	-	-	-	-3,43E-08
Strauts atteicis.	Jā	1	1	Jā	7,49E-04
Nelielies ar lielumu un varenumu.	Nē	-	-	-	0,00E+00
Ja nebūtu mazu strautu, nebūtu arī upes.	Nē	-	-	-	0,00E+00
Uz lauka ganījās zosu bars.	Jā	1	1	Jā	2,19E-03
Lapsa baru ieraudzīja.	Jā	1	1	Jā	0,00E+00
Lapsao sakārojās zoss cepeti.	Nē	-	-	-	0,00E+00
Lapsa klusām pielīda pie vienas zoss.	Jā	1	1	Jā	3,71E-03
Te pēkšņi krūze sāka svilpot.	Jā	1	6	Nē	1,44E-03
Zosis aizbēga.	Jā	1	1	Jā	1,20E-03
Lapsa ieraudzīja krūzi, kurā vējš svilpoja.	Nē	-	-	-	1,00E+00
Ak tad esi tā spēlmane.	Nē	-	-	-	0,00E+00
Lapsa iesaucās.	Jā	1	1	Jā	3,27E-03
Pagaidi tikai.	Nē	-	-	-	0,00E+00

Lapsa pieskrēja pie krūzes un uzmauca osu kaklā.	Nē	-	-	-	0,00E+00
Nu nesa krūzi uz upi slīcināt.	Jā	6	0	Nē	3,58E-03
Lapsa nonāca pie upes.	Jā	1	1	Jā	2,65E-03
Lapsa uzkāpa uz laipas un laida krūzi ūdenī.	Nē	-	-	-	0,00E+00
Krūze piesmēlās un sāka vilkt lapsu līdz.	Nē	-	-	-	0,00E+00
Nu lapsa sāka lūgties.	Jā	2	2	Jā	1,00E+00
Krūzīt, krūzīt.	Nē	-	-	-	0,00E+00
Lapsa pa jokam, krūze no tiesas.	Nē	-	-	-	0,00E+00
Bet nelīdzēja nekāda lūgšanās.	Nē	-	-	-	0,00E+00
Krūze piesmēlās pilna un ierāva lapsu dziļā ūdenī.	Nē	-	-	-	0,00E+00
Dzīvojis reiz puisis.	Jā	1	1	Jā	1,41E-03
Puisis bija liels koklētājs un dziedātājs.	Jā	1	1	Jā	0,00E+00
Kad puisis koklējis un dziedājis, putni bariem laidošies klausīties.	Nē	-	-	-	0,00E+00
Puisi reiz naktī mežā apmaldījies.	Jā	5	3	Nē	3,06E-03
Puisis iekritis dziļā vīķu bedrē.	Jā	1	3	Nē	4,97E-07
Dziļumā vilks jau gatavojies koklētājam uzbrukt.	Nē	-	-	-	3,10E+00
Puisi sācis koklēt un dziedāt.	Nē	-	-	-	5,05E+00
Vilks kaucis dziesmai līdzī.	Jā	1	0	Nē	2,58E-03
Vilks nedarījis nekā ļauna.	Nē	-	-	-	1,08E+00
Koklētājs gaismiņā izlēcis no bedres un gājis uz mājām.	Nē	-	-	-	0,00E+00
Vilsk esot pavadījis lielo koklētāju uz māju.	Jā	2	2	Jā	2,00E+00
Vilks nemaz negribēja vairs uz mežu atpakaļ griezties.	Jā	2	2	Jā	1,00E+00
Piemājas dārziņā aug trīs ābeles.	Jā	1	5	Nē	1,08E-03
Viena ābolu pilna.	Nē	-	-	-	0,00E+00
Zem augļu svara ābeles zari ir nolīkuši.	Jā	1	0	Nē	1,00E+00
Ābele lūdzas.	Jā	1	1	Jā	2,14E-03
Ņem nost smagos ābolus.	Jā	1	0	Nē	7,59E-04
Citādi lūzīšu.	Nē	-	-	-	0,00E+00
Mikiņš noplūc pašu sārtāko ābolu.	Jā	1	1	Jā	3,31E+00
Mikiņš ar nazi pārgriež ābolu un nogaršo.	Nē	-	-	-	0,00E+00
Zēns nopūšas.	Jā	1	1	Jā	0,00E+00
Sēkliņš vēl nav gluži brūnas.	Jā	1	1	Jā	1,00E+00
Mikiņš atbalsta zarus ar gariem, zarainiem kokiem.	Jā	15	8	Nē	1,05E-04

Nu zari nelūzīs.	Nē	-	-	-	0,00E+00
Pa ielu gāja vecenīte un nesa grozu.	Nē	-	-	-	1,00E+00
Grozā bija āboli.	Jā	2	0	Nē	1,00E+00
Kaut dabūtu vienu.	Nē	-	-	-	0,00E+00
Nodomāja Hugo.	Nē	-	-	-	0,00E+00
Zēns klusu piezagās no mugurpuses.	Jā	1	4	Nē	0,00E+00
Zēns paņēma no groza ābolu.	Jā	4	4	Jā	3,39E-03
Zēns ābolu steidzīgi iebāza kabatā.	Jā	1	9	Jā	0,00E+00
Vecenīte neko nemanīja.	Nē	-	-	-	0,00E+00
Vecenīte apstājās un sacīja.	Nē	-	-	-	4,14E-06
Panāc šurp.	Jā	1	1	Jā	1,13E-03
Vecenīte izvēlējās vislielāko ābolu un sniedza zēnam.	Nē	-	-	-	0,00E+00
Ēd.	Jā	1	1	Jā	1,29E-03
Ābols ir no vecenītes dārza.	Jā	4	4	Jā	2,00E+00
Zēnam likās, ka ābols kabatā kļūst karsts kā uguns.	Nē	-	-	-	0,00E+00
Zēns steidzīgi iemeta grozā nozagto ābolu un aizskrēja.	Nē	-	-	-	1,87E-03
Hugo atskrēja mājas un sāka raudāt.	Nē	-	-	-	0,00E+00
Kādam ķēniņam bija skaista meita.	Jā	1	1	Jā	1,00E+00
Reiz meita saslimst.	Jā	1	1	Jā	1,90E-03
Ķēniņš aicina visādus ārstus.	Jā	1	1	Jā	5,65E-03
Ārsti nevar meitai palīdzēt.	Jā	1	1	Jā	1,00E+00
Pēdīgi saka, ka vajagot zelta ābolu.	Nē	-	-	-	0,00E+00
Zelta ābols meitu izārstēs.	Jā	1	1	Jā	7,09E-04
Ķēniņš izlaiž pa zemi ziņu.	Jā	2	7	Nē	2,47E-03
Kurš zelta ābolu atnesīs, dabūs meitu par sievu.	Nē	-	-	-	0,00E+00
Zemē dzīvoja tēvs ar trim dēliem.	Jā	5	7	Nē	1,00E+00
Rītā tēvs pieceļas un ierauga durvju priekšā ābeli ar trim zelta āboliem.	Nē	-	-	-	0,00E+00
Tēvam skaidrs, kas darāms.	Nē	-	-	-	0,00E+00
Tēvs noplūc ābolus un iedod dēliem pa vienam.	Nē	-	-	-	1,00E+00
Ejiet uz pil un izmēģiniet laimi.	Nē	-	-	-	5,06E-07
Sākumā gāja vecākais dēls.	Jā	1	2	Nē	1,00E+00
Vecākais dēls paņēma ceļam līdz gabalu maizes.	Jā	1	0	Nē	1,00E+00
Zelta ābolu vecākais dēls ielika kabatā.	Jā	1	6	Nē	1,00E+00
Dēls iet, iet līdz ieiet biežā mežā.	Nē	-	-	-	1,00E+00
Vecākajam dēlam gribās ēst.	Jā	1	0	Nē	2,00E+00

Vecākais dēls salūko gabalu mazes un sāk maizi tiesāt.	Nē	-	-	-	5,48E-07
Te, kur gadījies, kur ne, nāk sirms vecītis.	Nē	-	-	-	0,00E+00
Vecītis lūdz, lai iedodot kādu garoziņu.	Nē	-	-	-	3,22E-06
Vecākais dēls nedomā ar vecīti dalīties ar kumosu.	Jā	55	0	Nē	1,00E+00
Strādā pats, tad nebūs maize jādiedelē.	Nē	-	-	-	0,00E+00
Dēlu dzirdēdams, vecītis neko nesaka un pazūd.	Nē	-	-	-	0,00E+00
Otrā dienā vecākais dēls nonāk ķēniņa pilī.	Jā	2	3	Nē	2,00E+00
Sulaiņi dēlu aizved pie ķēniņa.	Jā	1	1	Jā	2,00E+00
Ķēniņš vaicā, ko dēls vēloties.	Nē	-	-	-	5,28E-04
Dēls saka, ka esot atnesis zelta ābolu.	Nē	-	-	-	1,99E-03
Gribot princesi par sievu.	Nē	-	-	-	0,00E+00
Vispirms parādi ābolu.	Jā	1	1	Jā	1,58E-03
Ķēniņš pavēl.	Jā	1	1	Jā	3,44E-03
Vecākais dēls tūlīt bāzīs roku kabatā, lai izvilkto ābolu.	Nē	-	-	-	1,00E+00
Tavu brīnumu .	Nē	-	-	-	0,00E+00
Ābols kabatā pārvērties par čūsku.	Jā	1	5	Nē	0,00E+00
Ķēniņš par mānīšanos liek iemest vecāko dēlu cietumā.	Jā	14	0	Nē	1,00E+00
Nu iet laimi izmēģināt vidējais dēls.	Jā	1	0	Nē	2,00E+00
Vidējam dēlam klājas gluži tāpat kā vecākajam.	Jā	3	0	Nē	1,24E-03
Nu jaunākais dēls posās uz pili.	Jā	11	11	Jā	6,58E-03
Ceļam jaunākais dēls iebāza maizi kabatā.	Jā	3	3	Jā	2,00E+00
Turpat dēls ielika zelta ābolu.	Jā	1	1	Jā	4,00E+00
Dēls gāja, gāja līdz nonāca biežā mežā.	Nē	-	-	-	1,00E+00
Dēls sadomāja apēst maizi.	Jā	1	0	Nē	1,73E-03
Tikko dēls sāka ēst, klāt sirms vecītis.	Nē	-	-	-	9,22E-04
Vecītis lūdz iedot kādu garoziņu.	Jā	1	0	Nē	1,50E-03
Esot ļoti izsalcis.	Jā	2	2	Jā	1,00E+00
Dēls atdeva visu maizi.	Jā	1	1	Jā	1,76E-03
Še, ēd vesels tētiņ.	Nē	-	-	-	0,00E+00
Jau paēdu.	Jā	1	2	Nē	1,36E-03
Vecītis pateicās par labo sirdi.	Jā	1	1	Jā	4,94E+00
Uz priekšu dari otram labu.	Nē	-	-	-	1,04E+00
Tad dzīvē labi klāsies.	Jā	3	21	Nē	1,61E-03
Otrā rītā dēls nonāca ķēniņa pilī.	Jā	3	3	Jā	4,00E+00
Dēls iedeva ķēniņam zelta ābolu.	Jā	1	1	Jā	1,01E+00

Ķēniņš nevarēja prieku valdīt.	Jā	1	1	Jā	1,00E+00
Ķēniņš paņēma zelta ābolu un gāja pie meitas.	Nē	-	-	-	2,00E+00
Meita ieraudzīja ābolu.	Jā	1	1	Jā	1,00E+00
Tūlīt meita no gultas ārā un veselāka par veselu.	Nē	-	-	-	0,00E+00

Programmas diska apraksts

Darbā iekļauts disks ar modernizētās programmas kodu. Failā README tiek aprakstīts kā startēt programmu.

Maģistra darbs: **latviešu valodas sintaktiskā analizatora "Čankeris" modernizācija**

Ar savu parakstu apliecinu, ka pētījums veikts patstāvīgi, izmantoti tikai tajā norādītie informācijas avoti un iesniegtā darba elektroniskā kopija atbilst izdrukai.

Autors: _____
(Autora paraksts)

Ar savu parakstu apliecinu, ka esmu lasījis augstāk minēto maģistra darbu un atzīstu to par **p i e m ē r o t u / n e p i e m ē r o t u** (nevajadzīgo svītrot) aizstāvēšanai Latvijas Universitātes datorzinātņu maģistrantūrā.

Darba vadītājs: _____
(Vadītāja paraksts)

Darbs iesniegts maģistrantūras sekretariātā _____.
(Iesniegšanas datums)

Ar šo es apliecinu, ka darba elektroniskā versija ir augšupielādēta LU informatīvajā sistēmā.

Studiju metodiķe: _____
(Metodiķes paraksts)

Recenzents: _____
(Akad. amats, zin. grāds, vārds, uzvārds)

Darbs aizstāvēts maģistra gala pārbaudījuma komisijas sēdē

_____ prot. Nr. _____
(Darba aizstāvēšanas datums)

Komisijas sekretārs: _____
(Sekretāra paraksts)