



دانشگاه تهران

تحصیلات تكمیلی

پایان نامه کارشناسی ارشد در زبانشناسی همگانی

عنوان:

پردازش گروه های نحوی و جملات ساده خبری

زبان فارسی در پرولوگ بر اساس برنامه کمینه گرا

استاد راهنما:

دکتر عباسعلی آهنگر

استاد مشاور:

دکتر محمد علی زهرا زاده

تحقیق و نگارش:

زینب طباخی

بهمن ۱۳۹۱

لهم إني
أعوذ بِكَ مِنْ شَرِّ
مَا أَنْتَ مَعَهُ
أَنْتَ أَعْلَمُ

بسمه تعالی

این پایان نامه با عنوان پردازش جملات ساده و مرکب زبان فارسی در پرولوگ بر اساس برنامه کمینه گرا قسمتی از برنامه آموزشی دوره کارشناسی ارشد زبانشناسی همگانی توسط دانشجو زینب طباخی با راهنمایی استاد پایان نامه دکتر عباسعلی آهنگر تهیه شده است. استفاده از مطالب آن به منظور اهداف پژوهشی با ذکر مرجع و اطلاع کتبی به حوزه تحصیلات تكمیلی دانشگاه سیستان و بلوچستان مجاز می باشد.

نام و امضا دانشجو

این پایان نامه واحد درسی شناخته می شود و در تاریخ توسط هیئت داوران بررسی و درجه به آن تعلق گرفت.

تاریخ	امضاء	نام و نام خانوادگی
-------	-------	--------------------

استاد راهنما: دکتر عباسعلی آهنگر

استاد مشاور: محمد علی زهراء زاده

: داور ۱

: داور ۲

نماینده تحصیلات تکمیلی:



تعهدنامه اصالت اثر

اینجانب زینب طباخی تعهد می کنم که مطالب مندرج در این پایان نامه حاصل کار پژوهشی اینجانب است و به دستاوردهای پژوهشی دیگران که در این نوشته از آن استفاده شده است مطابق مقررات ارجاع گردیده است. این پایان نامه پیش از این برای احراز هیچ مدرکی هم سطح یا بالاتر ارایه نشده است.

کلیه حقوق مادی و معنوی این اثر متعلق به دانشگاه سیستان و بلوچستان می باشد.

نام و نام خانوادگی دانشجو: زینب طباخی

امضاء:

تقدیم به پدر و مادرم

که وجودم از بود آنهاست

و همسر مهربانم که در همه مراحل زندگی یار و

یاوری دلسوز بوده است

سپاسگزاری

جناب آقایان...دکتر آهنگر..... ودکتر زهرا زاده..... اساتید راهنمای و مشاورم:

شما روشنایی بخش تاریکی جان هستید و ظلمت اندیشه را نور می بخشد. چگونه سپاس گویم مهربانی و لطف شما را که سرشار از عشق و یقین است؟ چگونه سپاس گویم تأثیر علم آموزی شما را که چراغ روشن هدایت را بر کلبه‌ی محقر وجودم فروزان ساخته است؟ آری در مقابل این همه عظمت و شکوه شما مرا نه توان سپاس و نه کلام وصف است.

چکیده

بررسی مختصر تاریخچه نرم افزار های مبتنی بر زبان طبیعی بیانگر این مطلب است که جز با داشتن یک انگاره زبانی مناسب طراحی آنها امکانپذیر نخواهد بود. لذا به نظر می رسد برنامه کمینه گرای^۱ چامسکی (۱۹۹۵) بتواند این انگاره را به منظور تدوین برنامه ای که در بر دارنده روش تجزیه کلیه بخش‌های زبانی باشد، فراهم آورد. هدف پژوهش حاضر بررسی انواع گروه های نحوی و جملات ساده خبری در زبان فارسی بر اساس برنامه کمینه گرا است. بخش نرم افزاری این پژوهش زبان برنامه نویسی پرولوگ^۲ است که از جمله زبان های مبتنی بر منطق محسوب می شود. پیکره در نظر گرفته شده در این پژوهش با استفاده از الگوریتم های از پیش تعیین شده خودمفسر^۳ پرولوگ در چهار سطح زبانی توسط چهار تحلیل گر، بر اساس طرح برنامه کمینه گرای چامسکی تجزیه می شوند. به منظور تجزیه هریک از بخش های زبانی مربوط، در پرولوگ، بر اساس برنامه کمینه گرا یک تحلیل گر در نظر گرفته شده است که در بر دارنده دو بخش واژگان^۴ و نظام محاسباتی^۵ است. نتیجه حاصل از این پژوهش حاکی از آن است که برنامه کمینه گرا در مقایسه با سایر انگاره های زبانی، شیوه اقتصادی تری را در زمینه پردازش گروه های نحوی و جملات ساده خبری زبان فارسی فراهم می آورد.

كلمات کلیدی: برنامه کمینه گرا، پرولوگ، واژگان، نظام محاسباتی، گروه های نحوی، جمله ساده خبری.

¹ Minimalist Program

² Prolog

³ Self- Interpreter

⁴ Lexicon

⁵ Computational system

فهرست مطالب

عنوان	
صفحه	
۱	فصل اول: درآمدی بر پژوهش
۲	۱-۱- مقدمه
۲	۱-۲- بیان مساله
۳	۱-۳- پرسش های پژوهش
۴	۱-۴- فرضیه های پژوهش
۴	۱-۵- اهداف پژوهش
۵	۱-۶- ضرورت انجام پژوهش
۵	۱-۷- روش انجام پژوهش
۶	۱-۸- محدودیت های پژوهش
۷	فصل دوم: پیشینه و مفاهیم نظری پژوهش
۸	۲-۱- مقدمه
۸	۲-۲- پیشینه پژوهش
۸	۲-۳- پیشینه عملی پژوهش
۱۰	۲-۴- پیشینه نظری پژوهش
۱۰	۲-۵- تاریخچه پردازش زبان طبیعی
۱۱	۲-۶- ۱-۱-۲- دهه بینش بنیادی: ۱۹۴۰ تا ۱۹۵۷
۱۳	۲-۷- ۱-۲- دو گرایش اصلی: ۱۹۵۷ تا ۱۹۷۰
۱۴	۲-۸- ۱-۲- چهار الگو:
۱۶	۲-۹- ۱-۲- بازگشت به تجربه گرایی و الگوهای کران بسته: ۱۹۸۳ تا ۱۹۹۴
۱۷	۲-۱۰- ۱-۲- ادغام شیوه ها: ۱۹۹۴ تا ۱۹۹۹

عنوان	صفحه
۲-۲-۲-۲-۲- پیدایش پردازش دستوری	۱۷
۲-۲-۲-۲-۱- دستور زبان های نوع ۳	۱۹
۲-۲-۲-۲-۲- دستور زبان های نوع ۲ یا بافت آزاد	۲۰
۲-۲-۲-۳- دستور زبان های نوع ۱	۲۱
۴-۲-۲-۲- دستور زبان های نوع صفر	۲۲
۳-۲-۲-۲- مساله تجزیه دستوری	۲۲
۲-۳-۲-۱- تجزیه دستوری در پرولوگ	۲۳
۳-۲-۲- مفاهیم پژوهش	۲۴
۲-۳-۱- مفاهیم کلی در زبان برنامه نویسی پرولوگ	۲۴
۱-۱-۳-۲- پرولوگ	۲۵
۲-۱-۳-۲- متغیر	۲۷
۳-۱-۳-۲- حقایق	۲۸
۴-۱-۳-۲- فهرست	۳۰
۵-۱-۳-۲- دلایل انتخاب پرولوگ	۳۱
۲-۳-۲- مفاهیمی در زمینه دستور زایشی و دیدگاه کمینه گرایی	۳۳
۱-۲-۳-۲- دستور همگانی (UG)	۳۳
۲-۲-۳-۲- اصول و پارامترها	۳۶
۳-۲-۳-۲- برنامه کمینه گرا	۳۷
۴-۴- چارچوب نظری پژوهش	۴۰
۱-۴-۲- کمینه گرایی و پرولوگ	۴۰
۲-۴-۲- تجزیه پایین به بالا	۴۵
۴۸- فصل سوم: توصیف و تحلیل داده ها	
۱-۳- مقدمه	۴۹
۲-۳- بخش واژگان	۴۹

صفحه	عنوان
۴۹	۱-۲-۳- وازگان
۵۱	۳-۳- گروه های نحوی زبان فارسی
۵۱	۱-۳-۳- فعل و انواع آن در زبان فارسی
۵۲	۱-۱-۳-۳- فعل اصلی
۵۳	۲-۱-۳-۳- فعل رابط
۵۳	۳-۱-۳-۳- فعل معین
۵۳	۴-۱-۳-۳- فعل وجہی
۵۴	۵-۱-۳-۳- عنصر فعلی
۵۴	۲-۳-۳- تحلیل نحوی فعل و گروه فعلی در زبان فارسی
۵۴	۱-۲-۳-۳- پایه فعل
۵۶	۲-۲-۳-۳- گروه فعلی
۵۶	۱-۲-۲-۳-۳- فعل رابط
۵۹	۲-۲-۲-۳-۳- فعل لازم
۶۰	۳-۲-۲-۳-۳- فعل متعددی تک مفعولی
۶۱	۴-۲-۲-۳-۳- فعل متعددی دو مفعولی
۶۲	۵-۲-۲-۳-۳- فعل متعددی مرکب
۶۴	۶-۲-۲-۳-۳- فعل متعددی دارای وابسته حرف اضافه ای
۶۵	۳-۳-۳- انواع صورت های تصریفی فعل در زبان فارسی
۶۶	۱-۳-۳-۳- حال ساده اخباری
۶۶	۲-۳-۳-۳- حال مستمر
۶۷	۳-۳-۳-۳- حال التزامی
۶۷	۴-۳-۳-۳- آینده
۶۸	۵-۳-۳-۳- امر
۶۸	۶-۳-۳-۳- نهی

صفحه	عنوان
۶۸	۷-۳-۳-۳- گذشته ساده
۶۹	۸-۳-۳-۳- گذشته استمراری
۶۹	۹-۳-۳-۳- گذشته مستمر
۶۹	۱۰-۳-۳-۳- گذشته نقلی
۷۰	۱۱-۳-۳-۳- گذشته نقلی استمراری
۷۱	۱۲-۳-۳-۳- گذشته بعيد
۷۱	۱۳-۳-۳-۳- گذشته التزامی
۷۱	۴-۳-۳- گروه اسمی
۷۲	۱-۴-۳-۳- وابسته های پیشین اسم
۷۲	۱-۱-۴-۳-۳- وابسته های وصفی
۷۲	۲-۱-۴-۳-۳- وابسته عدد و ممیز
۷۳	۳-۱-۴-۳-۳- وابسته اشاره
۷۳	۴-۱-۴-۳-۳- وابسته پرسشی
۷۳	۵-۱-۴-۳-۳- وابسته تعجب
۷۳	۶-۱-۴-۳-۳- وابسته نامشخص
۷۴	۲-۴-۳-۳- وابسته های پسین اسم
۷۵	۱-۲-۴-۳-۳- وابسته صفتی
۷۵	۲-۲-۴-۳-۳- وابسته اسمی
۷۶	۳-۲-۴-۳-۳- وابسته بدل
۷۷	۴-۲-۴-۳-۳- گروه کمی نما
۷۸	۵-۲-۴-۳-۳- گروه مالکیت
۷۹	۶-۲-۴-۳-۳- گروه اضافه
۷۹	۷-۲-۴-۳-۳- گروه حرف تعریف
۸۱	۵-۳-۳- گروه صفتی

صفحه	عنوان
۸۲	-۱-۵-۳-۳ روابط دستوری گروه صفتی
۸۳	-۶-۳-۳ گروه حرف اضافه ای
۸۴	-۷-۳-۳ گروه قیدی
۸۵	-۸-۳-۳ جمله
۸۶	-۱-۸-۳-۳ گروه وابسته نما
۸۶	-۲-۸-۳-۳ بند
۸۷	-۴-۳ ساختار کلی برنامه پردازش جملات فارسی بر مبنای کمینه گرایی
۸۹	فصل چهارم: نتیجه گیری و پیشنهادات
۹۰	-۱-۴ مقدمه
۹۰	-۲-۴ مروری بر پرسش ها و فرضیه های پژوهش
۹۰	-۱-۲-۴ پرسش های پژوهش
۹۱	-۲-۲-۴ فرضیه های پژوهش
۹۱	-۳-۴ نتیجه گیری
۹۲	-۱-۳-۴ توانایی برنامه پردازشگر نحوی در پرولوگ بر اساس برنامه کمینه گرا در تشخیص گروه های نحوی و اجزای کلام
۹۳	-۲-۳-۴ تاثیر مثبت پرولوگ و کمینه گرایی بر یکدیگر
۹۳	-۴-۴ پیشنهادات
۹۴	منابع
۹۴	الف- منابع فارسی
۹۵	ب- منابع انگلیسی

فهرست جدول ها

عنوان جدول	صفحه
جدول ۱-۳- تصریف فعل حال ساده اخباری	۶۶
جدول ۲-۳- تصریف فعل حال مستمر	۶۶
جدول ۳-۳- تصریف فعل حال التزامی	۶۷
جدول ۴-۳- تصریف فعل آینده	۶۷
جدول ۵-۳- تصریف فعل امر	۶۸
جدول ۶-۳- تصریف فعل نهی	۶۸
جدول ۷-۳- تصریف فعل گذشته ساده	۶۸
جدول ۸-۳- تصریف فعل گذشته استمراری	۶۹
جدول ۹-۳- تصریف فعل گذشته مستمر	۶۹
جدول ۱۰-۳- پسوندهای رابط	۷۰
جدول ۱۱-۳- تصریف فعل گذشته نقلی	۷۰
جدول ۱۲-۳- تصریف فعل گذشته نقلی استمراری	۷۰
جدول ۱۳-۳- تصریف فعل گذشته بعید	۷۱
جدول ۱۴-۳- تصریف فعل گذشته التزامی	۷۱
جدول ۱۵-۳- وابسته های پیشین اسم	۷۴
جدول ۱۶-۳- وابسته های پسین اسم	۷۵

فهرست شکل ها

عنوان جدول	صفحه
شکل ۲-۱- درخت ساختاری در فهرست پرولوگ	۳۱
شکل ۲-۲- انگاره کلی برنامه کمینه گرا	۳۸
شکل ۲-۳- ساختار کلی نظریه ایکس بار	۴۰
شکل ۳-۱- نمودار درختی گروه فعلی با فعل رابط	۵۷
شکل ۳-۲- پیاده سازی مثال (۱۹) در پرولوگ و تصویر نتیجه آن	۵۸
شکل ۳-۳- نمودار درختی گروه فعلی با فعل لازم	۵۹
شکل ۳-۴- نمودار درختی گروه فعلی با فعل های تک مفعولی	۶۰
شکل ۳-۵- نمودار درختی گروه فعلی با فعل های دو مفعولی	۶۲
شکل ۳-۶- نمودار درختی گروه فعلی با فعل های متعدد مرکب	۶۴
شکل ۳-۷- نمودار درختی گروه فعلی با فعل متعدد دارای وابسته حرف اضافه ای	۶۵
شکل ۳-۸- تجزیه روساختی وابسته بدل	۷۶
شکل ۳-۹- نمودار درختی گروه کمی نما	۷۹
شکل ۳-۱۰- نمودار درختی گروه اسمی	۸۰
شکل ۳-۱۱- تجزیه روساختی گروه صفتی	۸۱
شکل ۳-۱۲- نمودار درختی گروه صفتی	۸۲
شکل ۳-۱۳- نمودار درختی گروه حرف اضافه ای	۸۴

فهرست اختصارات

XP	X Phrase
X'	X-bar
X	X-head
VP, vp	Verb Phrase
V', v'	V-bar
V1, v1	Linking Verb
V2, v2	Intransitive Verb
V3, v3	Monotransitive Verb
V4, v4	Ditransitive Verb
V5, v5	Transitive Compound Verb
V6, v6	Prepositional Verb
DP, dp	Determiner Phrase
D', d'	D-bar
D, d	Determiner
PP, pp	Preposition Phrase
P, p	Preposition
QP, qp	Quantifier Phrase
Q', q'	Q-bar
Q, q	Quantifier
Spec	Specifier
, spec	
&P, &p	And Phrase
&	And-bar
&	And
Ø	Deletion
TP	Tense Phrase
QP, qp	Quantifier Phrase
Q', q'	Q-bar
Q, q	Quantifier

CP,cp	Complement Phrase
C',c'	Complement-bar
C,c	Complement
NP,np	Noun Phrase
N,n	Noun
ZP	Specifier Position
AdjP	Adjective Phrase
,adjp		
Adj'	Adjective-bar
,adj'		
Adj	Adjective
,adj		
AdvP	Adverb Phrase
,advp		
Adv'	Adverb-bar
,adv'		
Adv	Adverb
,adv		
PossP	Possessive Phrase
,possp		
Poss'	Possessive-bar
,poss'		
Poss	Possessive
,poss		
EzafP	Ezafat Phrase
,ezafp		
Ezaf'	Ezafat-bar
,ezaf'		
Ezaf	Ezafat
,ezaf		
YP	Complement Position
S,s	Sister, Small

فصل اول

درآمدی بر پژوهش

۱-۱- مقدمه

این فصل در بر گیرنده کلیات پژوهش است. بر این اساس در قالب بخش هایی شامل: بیان مساله، پرسش های پژوهش، فرضیه های پژوهش، اهداف پژوهش، لزوم انجام پژوهش، روش پژوهش و محدودیت هایی که پژوهش گر در حین انجام پژوهش با آنها مواجه شده است، ارایه می گردد.

۱-۲- بیان مساله

الگو سازی منطقی- زبان شناختی ششمین مرحله از ساخت سامانه های مبتنی بر دانش است که الگوهای آن شباهت بسیاری به نمودارهای ادراکی سوا^۱ دارد. در الگو سازی منطقی- زبان شناختی برای ساخت سامانه های مبتنی بر دانش^۲ (KBS) با توانایی یادگیری، از الگوهای ادراکی روش شناسی سامانه های نرم^۳ (SSM)، منطق گزاره ای و زبان هوش مصنوعی پرولوگ استفاده می شود. بدین ترتیب از اوایل دهه ۷۰ میلادی زبان برنامه نویسی پرولوگ^۴ که زبانی مبتنی بر کاربرد منطق مرتبه اول^۵ در برنامه نویسی رایانه ای بود، توسط کلمر^۶ و راسل^۷ شکل گرفت. زبان برنامه نویسی پرولوگ به صورت کلی یک زبان برنامه نویسی منطقی است که با هوش مصنوعی و زبان شناسی محاسباتی^۸ در ارتباط است (کلاکسین^۹ و ملیش^{۱۰}).

زبان برنامه نویسی پرولوگ ریشه در منطق مرتبه اول داشته و با زبان های برنامه نویسی دیگر تفاوت بسیار دارد و از ابتدا با هدف پردازش زبان طبیعی شکل گرفته است. این زبان با دارا بودن یک الگوریتم جستجوی ابتدا

¹ Sow's Conceptual Graphs

² Knowledge-Based System

³ Soft System Methodology

⁴ Prolog Programming Language

⁵ First-Order Logic

⁶ Alain Colmerauer

⁷ Philippe Russell

⁸ Computational Linguistics

⁹ Clocksin

¹⁰ Mellish

عمقی^{۱۱}، توانایی بسیاری در پردازش زبان طبیعی دارد تا جایی که می‌توان آن را در مقایسه با زبان‌های چون سی^{۱۲} و پاسکال^{۱۳} در پردازش زبان طبیعی قدرتمندتر دانست (کلاکسین و ملیش، ۲۰۰۳؛ میرزائیان، ۱۳۸۷).

یکی از نظریه‌های زبانی که به نظر می‌رسد می‌توان آن را در طراحی برنامه‌هایی که بر مبنای زبان برنامه نویسی پرولوگ نوشته می‌شود به کار گرفت، برنامه کمینه گرای^{۱۴} چامسکی^{۱۵} (۱۹۹۵) است. این برنامه به علت برخورداری از شیوه‌های اقتصاد گرایانه، گزینه مناسبی به شمار می‌آید. به علاوه، علاقه بسیار چامسکی در عمومیت بخشیدن این برنامه به زبان‌های بشری، جذابیت آن را برای استفاده در برنامه‌های رایانه‌ای افزایش می‌دهد.

این برنامه با در بر داشتن شیوه آلفا^{۱۶} بسیاری از خرده قوانین دیدگاه‌های گذشته چامسکی را تحت عنوان یک قانون کلی معرفی کرده است و با توجه به استدلال تیغ اکام^{۱۷}، از بسیاری حشوها و تکرارها در آن چشم پوشی شده است (دبیر مقدم، ۱۳۸۳).

از این رو، به نظر می‌رسد بتوان دیدگاه‌های نظری چامسکی در برنامه کمینه گرا در پرولوگ را برای پردازش گروه‌های نحوی و جملات ساده خبری زبان فارسی به کار برد. بنابر این، این زبان می‌تواند شیوه مناسبی را برای پردازش انواع جملات فارسی فراهم آورد. همچنین به نظر می‌رسد بتوان از الگوریتم جستجوی ابتداء عمقی پرولوگ به جای الگوریتم نویسی‌های پیچیده استفاده کرد.

۱-۳-پرسش‌های پژوهش

پرسش‌های پژوهش حاضر عبارتند از:

- آیا می‌توان اصول برنامه کمینه گرا را در پردازش جملات ساده خبری فارسی به وسیله زبان برنامه نویسی پرولوگ به کار برد؟
- آیا می‌توان اصول برنامه کمینه گرا را در پردازش گروه‌های نحوی زبان فارسی به وسیله زبان برنامه نویسی پرولوگ به کار برد؟

¹¹ Depth-First Search

¹² C

¹³ Pascal

¹⁴ The Minimalist Program

¹⁵ Chomsky

¹⁶ Alpha Mode

¹⁷ Occam's Razor

- ۳- آیا زبان برنامه نویسی پرولوگ شیوه مناسبی را برای پردازش انواع جملات زبان فارسی فراهم می آورد؟
- ۴- آیا می توان از الگوریتم جستجوی ابتداء عمقی پرولوگ به جای الگوریتم نویسی های پیچیده استفاده کرد؟

۴- فرضیه های پژوهش

فرضیه های پژوهش به شرح زیر است:

- ۱- اصول برنامه کمینه گرا را می توان در پردازش جملات ساده خبری فارسی به وسیله زبان برنامه نویسی پرولوگ به کار برد.
- ۲- اصول برنامه کمینه گرا را می توان در پردازش جملات گروه های نحوی زبان فارسی به وسیله زبان برنامه نویسی پرولوگ به کار برد.
- ۳- زبان برنامه نویسی پرولوگ شیوه مناسبی برای پردازش جملات زبان فارسی فراهم می آورد.
- ۴- از الگوریتم جستجوی ابتداء عمقی پرولوگ می توان به جای الگوریتم های پیچیده استفاده کرد.

۵- اهداف پژوهش

اهداف پژوهش حاضر عبارتند از:

- ۱- بررسی اصول کمینه گرا در پردازش جملات ساده خبری فارسی به وسیله زبان برنامه نویسی پرولوگ.
- ۲- بررسی اصول کمینه گرا در پردازش گروه های نحوی زبان فارسی به وسیله زبان برنامه نویسی پرولوگ.
- ۳- بررسی شیوه الگوریتم جستجوی ابتداء عمقی زبان برنامه نویسی پرولوگ برای پردازش جملات ساده و مرکب زبان فارسی.
- ۴- نمایش تاثیرگذاری مثبت دو تفکر کمینه گرایی و برنامه نویسی منطقی بر روی یکدیگر و تدوین یک برنامه با استفاده از زبان برنامه نویسی پرولوگ جهت پردازش گروه های نحوی و جملات ساده خبری زبان فارسی بر اساس دیدگاه کمینه گرا.