

# بررسی روش‌های تشخیص خودکار و رفع ابهام معنایی مفاهیم استعاری و دارای مجاز

بهرام وزیرنژاد	محمد بحرانی	هادی عبدی قویدل
استادیار	استادیار	دانشجوی کارشناسی ارشد
دانشگاه صنعتی شریف	دانشگاه صنعتی شریف	دانشگاه صنعتی شریف
<a href="mailto:bahram@sharif.ir">bahram@sharif.ir</a>	<a href="mailto:bahrani@sharif.ir">bahrani@sharif.ir</a>	<a href="mailto:hadi_abdighavidel@mehr.sharif.ir">hadi_abdighavidel@mehr.sharif.ir</a>

## چکیده

رفع ابهام معنایی از متون به یکی از چالش‌های مهم برای ماشین مبدل شده است. با اینکه امروزه ابزارهای متعددی وجود دارند که هر کدام به درک بهتر توسط ماشین کمک می‌کنند، اما رفع ابهام موجود در جملاتی که دارای مفاهیم استعاری و مجاز هستند هنوز به عنوان مسأله‌ای حل نشده باقی مانده و از تیررس تحقیق و تفحص رایانشی برای زبان فارسی به دور مانده‌اند. زیرا تاکنون این گونه ابهامات توسط زبان‌شناسان متبحر در علم منطق و روان‌شناسی به صورت دستی رفع می‌شد، درحالی‌که هرچه به سمت خودکار شدن تولید و درک زبان نزدیک می‌شویم لازم است این ابهامات نیز به صورت خودکار شناخته شده و درک و تفسیر شوند. مقاله حاضر ضمن معرفی تئوری کلی زبان استعاری و اهمیت آن برای ماشین، روش‌های مختلف رایانشی تشخیص و رفع ابهام معنایی عبارات استعاری و دارای مجاز را شرح می‌دهد.

**کلیدواژه‌ها:** رفع ابهام معنایی، مفاهیم استعاری و مجازی، زبان فارسی، درک زبان

## ۱- مقدمه

مکالمات روزمره بشر هیچگاه سرد و بی‌روح نبوده و همیشه همگام با تفکرات خلاقانه اوست. این تفکرات اغلب به صورت زبان استعاری<sup>۱</sup> نمود می‌یابند. زبان استعاری را زبانی می‌نامند که با استفاده از صنایع ادبی<sup>۲</sup>

---

<sup>۱</sup> figurative language

<sup>۲</sup> figures of speech

به بیان مفهومی به غیر از معنای تحت‌اللفظی آن می‌پردازد. گاه از زبان استعاری، زبان مجازی نیز یاد می‌کنند که ما در این مقاله برای امتناع از خلط مباحث و مفاهیم تنها از آن به زبان استعاری یاد می‌کنیم. با استفاده از زبان استعاری بشر برای ذهنیت خویش که معمولاً یک مفهوم عقلی و نامحسوس است، عینیتی را که غالباً پدیده‌ای محسوس است بر می‌گزینند. این نوع زبان را نمی‌توان تنها مختص مجموعه متون ادبی دانست، زیرا در مکالمات رایج و متون غیرادبی زبان نیز بارها شاهد این پدیده هستیم. بر خلاف زبان تحت-اللفظی که درک آن عملی بس آسان است، درک زبان استعاری بدین گونه نیست. معنای نهفته در این زبان را نمی‌توان از معنای اجزای آن دریافت. به عنوان مثال عبارت "خلیل حافظ می‌خواند." دارای مفهوم مجازی است. بدین معنا که حافظ در این عبارت در غیر معنی حقیقی‌اش یعنی کتاب حافظ به کار رفته است. یا در مثالی دیگر، عبارت "خلیل ماشین را جراحی کرد." دارای مفهوم استعاری است. در این عبارت جراحی کردن به جای تعمیر کردن بر حسب یک نوع تشابه به کار رفته است. تشخیص و درک چنین عباراتی، امروزه به یکی از چالش‌های اساسی علم پردازش زبان طبیعی مبدل گشته است و سیستمی که قابلیت درک و تفسیر چنین زبانی را در متون نامحدود دارا باشد، ابزاری بسیار با ارزش و پرکاربرد برای علم پردازش زبان طبیعی و زبان‌شناسی رایانشی به حساب خواهد آمد.

زبان استعاری از قسمت‌های مختلفی نظیر تشبیه، استعاره، مجاز، مبالغه، تعریض، تلمیح و غیره تشکیل شده که در این مقاله هدف خود را تنها در رابطه با تحقیق و مطالعه راجع به مفاهیم استعاری و دارای مجاز معطوف ساخته‌ایم.

در زبان انگلیسی و برخی زبان‌های دیگر، تحقیقات متعددی در این زمینه انجام شده است، اما اهمیت این زمینه تحقیقاتی برای علم پردازش زبان طبیعی و زبان‌شناسی رایانشی در ایران ناشناخته بوده و بدان خاطر مطالعاتی در این رابطه صورت نپذیرفته است. اگر قابلیت پردازش مفاهیم استعاری و دارای مجاز به سیستم‌هایی مانند بازیابی اطلاعات، ترجمه ماشینی، استخراج اطلاعات، پاسخگوی خودکار، دسته‌بندی متون، درک گفتار، خوشه‌بندی متون، خلاصه‌سازی و تفسیر متون و همچنین عقیده‌کاوی یا تحلیل مقصود در زبان فارسی تزریق شود، دقت عملکرد و میزان کارایی آنها به مقدار قابل توجهی افزایش خواهد یافت.

اخیراً یک موسسه آمریکایی موسوم به یارپا<sup>۱</sup> در حال انجام پروژه‌ای است که طبق گفته بایلیس<sup>۲</sup> (۲۰۱۱) قرار است در طی آن برای تجزیه و تحلیل مسائل فرهنگی و اتخاذ سیاست‌های بهتر دست به ساخت نرم‌افزاری برای تجزیه و تحلیل زبان استعاری بزند. اهداف کلی این برنامه بدین ترتیب است:

<sup>۱</sup> US Intelligence Advanced Research Project Activity

<sup>۲</sup> Bayliss

۱- مشخص کردن اینکه مردم چگونه در مورد مسائل پیچیده فکر کرده و زبان استعاری چه تأثیری بر عقاید آنها دارد.

۲- مشخص کردن پیچیدگی معنای مرتبط با انواع موضوع.

۳- نشان دادن معانی و جهان‌بینی‌های آشکار و استنتاجی گروه‌ها یا افراد خاص، ویژگی ناهمخوانی‌های موجود در مسائل اجتماعی.

۴- بیان ماهیت ظاهری و باطنی گروه‌های اجتماعی و سیاسی و فهم مشکلات و درگیری‌های روانی.

این برنامه باید در ایجاد چارچوب نظام‌مندی کمک کند که بتوان استعاره‌های مرتبط با ابعاد مختلف یک مساله تحلیلی را سازماندهی کرد و همچنین باید دست به ایجاد پایگاه دانشی از استعاره‌ها بزند که در آن تمامی استعاره‌ها و اطلاعات مربوط به آن برای دسترسی و استفاده‌های آتی ذخیره شده باشند. موسسه یارپا در معرفی اولیه پروژه خود روش‌های کلی مطالعه زبان استعاری را بدون پرداختن به شرح مفصل آن به پنج بخش تقسیم کرده است. مقاله حاضر با الهام از این روش‌ها، چارچوب کلی خود را در سه بخش تشخیص و تفسیر و ارزیابی به تفصیل مورد تحلیل و بررسی قرار می‌دهد.

## ۲- روش‌های مختلف تشخیص زبان استعاری

### ۲-۱- روش مبتنی بر پایگاه دانش

در پایگاه داده یا دانش معمولاً مفاهیم و اطلاعات خاص با اصول مهندسی ذخیره می‌شوند و برای اهداف خاص در آینده مورد استفاده قرار می‌گیرند. زبان‌شناسان و متخصصین حوزه ادبیات مجموعه‌ای از اصطلاحات استعاری شناخته شده برای اهل زبان را جمع‌آوری و آنها را در دست مهندسی قرار می‌دهند تا آنها این مجموعه اصطلاحات را با شیوه‌های مهندسی دانش سازماندهی کرده و در پایگاهی ذخیره کنند. در تشخیص مبتنی بر پایگاه دانش، روش به صورت جستجویی است. نخست متن بر طبق منطق مرتبه اول و یا غیره به صورت قاب‌های معنایی نمایش داده می‌شود و سپس اگر متن یا جمله حاوی مفهوم استعاری موجود در پایگاه دانش باشد، عمل تشخیص و تفسیر با توجه به مدخل‌های موجود در این پایگاه دانش رخ می‌دهد. در غیر اینصورت متن فاقد مفهوم استعاری خواهد بود.

پروژه ای‌تی‌تی‌متا<sup>۱</sup> (جان باردن<sup>۲</sup>: دانشگاه بیرمنگام) اولین قدم را در راه جمع‌آوری اصطلاحات استعاری برداشته است. این پروژه بانک دانشی<sup>۳</sup> ایجاد کرده که در آن نمونه‌هایی از فرایندهای واقعی گفتمان که حاوی زبان استعاری است به چشم می‌خورد. این پایگاه دانش شامل ۱۱۰۰ مورد نمونه‌های متنی و ۶۵ مورد نمونه گفتاری است.

## ۲-۲- روش مبتنی بر هستان‌شناسی

توماس گروبر (۱۹۹۳: ۲۲۰-۱۹۹) آنتولوژی یا هستان‌شناسی را در علم رایانه، ذکر مشخصات صریح و رسمی مفاهیم مشترک تعریف کرده است. بدین معنی که روابط و مفاهیم موجود کاملاً روشن تعریف شده باشند و ماشین بتواند آنها را با حداکثر سهولت ممکن بخواند و منحصر به فرد یا گروه خاصی نباشد. هستان‌شناسی مورد استفاده در علم زبان‌شناسی رایانشی، الهام گرفته از شیوه فرهنگ لغت نویسی است و روابط بین کلمات را در سطحی وسیع‌تر برای زبان نشان می‌دهد. وردنت<sup>۴</sup> و فریم‌نت<sup>۵</sup> و اخیراً کانسپت‌نت<sup>۶</sup> نمونه‌های هستان‌شناسی‌های رایج در جهان هستند که از هرکدام به طریقی برای تشخیص خودکار و درک زبان استعاری استفاده شده‌اند.

فاس<sup>۷</sup> (۱۹۹۳: ۹۰-۴۹) نخستین شخصی بود که با تکیه بر روابط موجود در هستان‌شناسی به تحقیق و مطالعه در این زمینه پرداخت. وی با استفاده از معنی‌شناسی تلفیقی<sup>۸</sup> نوعی هستان‌شناسی طراحی کرد که در آن قاب‌های معنایی بیش از ۵۰۰ کلمه در دسترس بود. سپس در قالب برنامه متا۵ بخشی از فرایند هم‌سنجی را انجام داده و طی عمل غربال‌گری، تمایز بین تحت‌اللفظی بودن<sup>۹</sup>، مجاز، استعاره و بی‌قاعدگی را از طریق نوع رابطه بین کلمات مشخص کرد.

نزدیک ده سال بعد از فاس، ویم پیترز<sup>۱۰</sup> (۲۰۰۴) با استفاده از وردنت تحقیقات جدیدی را آغاز کرد. وی در رساله خود بر اساس قاعده تکرر معنایی با قاعده و بی‌قاعدگی که توسط آپرسجان<sup>۱۱</sup> مطرح شده است، به دنبال

<sup>1</sup> Att-Meta

<sup>2</sup> John Barden

<sup>3</sup> <http://www.wcsbhamacuk/jab/ATT-Meta/Databank>

<sup>4</sup> WordNet

<sup>5</sup> FrameNet

<sup>6</sup> ConceptNet

<sup>7</sup> Fass

<sup>8</sup> collative semantics

<sup>9</sup> literal: حقیقی یا غیر ادبی

<sup>10</sup> Wim Peters

<sup>11</sup> Apresjan

گره‌های سطح بالایی بود که نقش والد<sup>۱</sup> را بر عهده دارند و دارای نوادگانی هستند که مفاهیم مشترکی دارند. نهایتاً نتیجه‌گیری آنها به این ترتیب بوده که این گره‌ها در اکثر موارد یا معرف مفاهیم استعاری هستند و یا مفاهیم دارای مجاز.

## ۲-۳- روش پیکره بنیاد

این روش از پیکره برچسب خورده برای فرایند تشخیص استفاده می‌کند. این نوع پیکره‌ها بر سه شکل هستند:

- ۱- پیکره پارس‌شده طبق نظریه وابستگی: هسته و وابسته را در گروه‌های نحوی از جمله تا کلمه مشخص می‌کند.
  - ۲- پیکره برچسب‌خورده با نقش‌های تتا و یا حالت‌های معنایی: به عنوان مثال نشان دادن نقش‌های کنشگر و کنش پذیر.
  - ۳- پیکره برچسب‌زده‌شده با زبان استعاری: این نوع پیکره‌ها توسط زبان‌شناسان متبحر در علم منطق و روان‌شناسی آموزش داده می‌شوند.
- این پیکره‌ها یا دارای یک ویژگی نظیر ویژگی اشاره شده در مورد اول و یا به صورت ترکیبی هستند. در نهایت پس از طی عملیات آماری، الگوهای به‌کار رفته در جملات استعاری تحلیل می‌شوند و بر اساس آن داده‌های آزمون مورد ارزیابی قرار می‌گیرند. در ادامه به روش‌های ترکیبی اشاره می‌شود.

## ۲-۴- مدل ترکیبی مبتنی بر هستان‌شناسی و پیکره بنیاد

این مدل دو روش مبتنی بر هستان‌شناسی و پیکره‌بنیاد را تلفیق کرده و نهایتاً با استفاده از روش‌های آماری عملیات تشخیص را پیاده‌سازی می‌کنند.

نخستین روش ترکیبی شش سال بعد از ویم پیترز با ورود فریم‌نت به حوزه پردازش زبان طبیعی بوده است. گدیگان<sup>۲</sup> و همکاران (۲۰۰۶) آیتم‌های لغوی مربوط به فریم خاص را از فریم‌نت استخراج کردند و سپس جملات حاوی این آیتم‌های لغوی را در پیکره پن‌تری-بنک<sup>۳</sup> جستجو و برچسب استعاره را برای موارد

---

<sup>1</sup> ancestor

<sup>2</sup> Gedigan

<sup>3</sup> Penn Tree Bank

استعاره‌ای بر آنها زدند. در نهایت با آموزش این پیکره توانستند به صورت آماری استعاری بودن یا نبودن را تعیین کنند.

کریشناکوماران<sup>۱</sup> و ژو<sup>۲</sup> (۲۰۰۷) با الهام از روش سیستم تروفای<sup>۳</sup> بیرک<sup>۴</sup> و سارکار<sup>۵</sup> (۲۰۰۶) که مفاهیم تحت-لفظی و غیر تحت‌اللفظی فعل را استخراج می‌کرد، به مفاهیم دیگر اجزای سخن نظیر اسم و صفت نیز پرداختند. کریشناکوماران و ژو با استفاده از روابط شمولیت و بسامد کلمات دنگاشتی مفاهیم استعاری را در سطح جمله پیش‌بینی کردند. مثلاً با داشتن مفهوم استعاره‌ای در جملات به این نتیجه رسیدند که اگر دو اسم در یک جمله باشد آن دو اسم در رابطه شمولیت با یکدیگر در وردنت قرار دارند و در غیر اینصورت دارای مفهوم استعاری خواهند بود. آنها همچنین مدل دو نگاشتی را بر روی افعال و صفت‌هایی که به صورت استعاری به کار می‌روند نیز آزمایش کردند. استنتاج آنها بدین شکل بود که اگر ترکیب‌های فعل-اسم و صفت-اسم دارای بسامد کافی در مجموعه دادگان نباشند جمله برچسب استعاری به خود می‌گیرد. به طور خلاصه روش کلی آنها مقایسه بین مفاهیمی بوده که در وردنت وجود دارند و یا ندارند.

## ۲-۵- روش مبتنی بر نشانه‌های زبانی

این روش بر اساس کلیدواژگان، حضور عبارت استعاری را مشخص می‌سازد. مانند "به طور استعاری، به معنای واقعی کلمه و غیره". این روش را اولین بار گوتلی<sup>۶</sup> (۱۹۹۷) به کار برد و شوتووا<sup>۷</sup> بر اساس مطالعه ترکیبی روی پیکره دقت آن را اندازه گرفت. البته دقت حاصل شده کمتر از ۵۰ درصد گزارش شده است.

## ۲-۶- تشخیص منبع و ماخذ عبارات استعاری در حوزه‌های خاص

منبع و ماخذ اصطلاحات کلی برای مشبه-مشبهه و مجاز-مجاز از می‌باشد. تشخیص منبع و ماخذ نیز نوعی مدل مبتنی بر پیکره است، اما چون ویژگی متفاوتی با بقیه کارهای انجام شده بر روی پیکره است به صورت جدا در این بخش مورد بررسی قرار می‌دهیم.

<sup>1</sup> Krishnakumaran

<sup>2</sup> Zhu

<sup>3</sup> Trope Finder (TroFi)

<sup>4</sup> Birek

<sup>5</sup> Sarkar

<sup>6</sup> Goatly

<sup>7</sup> Shutova

نخستین شخصی که با استفاده از این روش استعاره‌ای بودن مفاهیم را شناسایی کرد، میسون<sup>۱</sup> با سیستم کور مت<sup>۲</sup> خود در سال ۲۰۰۴ بوده است. وی نگاشت خودکار را در بین حوزه‌های خاص تحت وب مورد بررسی قرار داد. در این متد نخست محدودیت‌های گزینشی کلمات خاص مشخص می‌شوند و سپس پس از محاسبه تقارن بین حوزه‌ها، نگاشت مفاهیم منبع و ماخذ مشخص می‌شود. به عنوان مثال، فعل ریختن در دو حوزه آزمایشگاه و دارایی دارای محدودیت گزینشی مایعات برای آزمایشگاه و پول در دارایی است. سیستم کورمت چنین استنتاج کرده که نگاشت حوزه‌ای به صورت دارایی-آزمایشگاه و نگاشت مفهومی به صورت پول نوعی مایع است.

### ۳- روش‌های مختلف تفسیر زبان استعاری

روش‌های تفسیر زبان استعاری همانند روش‌های موجود برای فرایند تشخیص هستند، با این تفاوت که پدیدآورندگان بخش تفسیر از خلاقیت‌ها و مهارت‌های منحصر به فرد خود بهره جسته‌اند که در ادامه بر آنها مروری انجام می‌دهیم.

تقریباً همزمان با تشخیص زبان استعاری توسط فاس، مارتین<sup>۳</sup> (۱۹۹۲) سیستم تفسیر، دلالت<sup>۴</sup> و یادگیری استعاره را با نام میداس<sup>۵</sup> طراحی کرد. ایده به کار رفته در این سیستم بر این اساس است که استعاره متداول و قدیمی از استعاره‌های عام سرچشمه می‌گیرند. با دریافت عبارت استعاره‌ای، میداس پایگاه دانش خود را برای استعاره متناظری جستجو می‌کرد که بی‌قاعدگی را توضیح دهد که در صورت نیافتن آن، سیستم به جستجوی استعاره در مفاهیم عام‌تر به کار خود ادامه می‌داد. در نهایت در صورت یافتن استعاره عام‌تر، سیستم نگاشتی را برای زیرشاخه‌های آن یا استعاره خاص‌تر ایجاد می‌کرد. میداس امروزه در سیستم مشاوره یونیکس<sup>۶</sup> به کار می‌رود.

---

<sup>۱</sup> Mason

<sup>۲</sup> CorMet

<sup>۳</sup> Martin

<sup>۴</sup> denotation

<sup>۵</sup> MIDAS

<sup>۶</sup> سیستم مشاوره‌ی یونیکس یک واسط زبان طبیعی هوشمند است که به کاربران جدید این امکان را می‌دهد که بتوانند راجع به نحوه عملکرد سیستم یونیکس اطلاعات لازم را کسب نمایند

دسته‌ای از دیدگاه‌ها وجود دارد که بر اساس یادگیری و استنتاج مشبه و مشبه‌به در درک زبان استعاری می‌باشد؛ مانند سیستم کارما<sup>۱</sup> و پروژه<sup>۲</sup> *ای‌تی‌تی-متا*. مجریان هر دو سیستم، چارچوبی استنتاجی طبق نظریه<sup>۳</sup> استعاره ادراکی برای استعاره تشکیل داده‌اند. فرایند استنتاج بر اساس دانشی سراسری است که از قبل در سیستم تعبیه شده است و بیشتر در حوزه مشبه عمل می‌کند. سپس نتایج با استفاده از نمود نگاشتی ادراکی به حوزه مشبه‌به منتقل می‌شود. اساس کار پروژه<sup>۴</sup> *ای‌تی‌تی-متا* توصیف استعاری و مجازی حالت‌های ذهنی و همچنین استنتاج آنها با استفاده از منطق مرتبه اول است. ورودی این سیستم زبان طبیعی نیست، بلکه عبارات منطقی است که نمود چارچوب‌های کلامی می‌باشند. کارما نیز به نوبه خود با اعمال و پیشامدهای انتزاعی سروکار دارد و متن تجزیه‌شده را به عنوان ورودی دریافت می‌کند.

وئال و هائو<sup>۳</sup> (۲۰۰۸) از "بازنمود دانش سیال برای تفسیر و تولید استعاره" با نام نشانه‌های ناطق<sup>۴</sup> یاد کرده‌اند. نشانه‌های ناطق، مجموعه‌ای از ویژگی‌های مفهومی هستند که به حوزه‌ی مشبه و مشبه‌به و همچنین حقایق مربوط به جهان مربوط می‌شوند که نویسندگان به طور اتوماتیک از وردنت و وب اقتباس می‌کنند. این نشانه‌ها، سپس در سلیپنت<sup>۵</sup> سازمان‌دهی می‌شوند. سلیپنت چارچوبی است که برای تعدادی از الحاقات، حذفیات و جایگذاری‌های موجود در تعاریف این‌گونه ویژگی‌ها، این امکان را می‌دهد که رابطه‌ای میان مشبه و مشبه‌به ایجاد کنند. این ایده بر اساس بازنمود دانش سلیپاژ<sup>۶</sup> برای درک قرینه‌ها<sup>۷</sup> در حوزه‌های انتزاعی بنا نهاده شده است.

اکاترینا شتووا<sup>۸</sup> مقالات متعددی را در رابطه با تشخیص و تفسیر مفاهیم استعاری و دارای مجاز در سال‌های ۲۰۰۹ و ۲۰۱۰ نوشته و به مدل‌های منطقی و زبان‌شناختی آن پرداخته است. وی نهایتاً در سال ۲۰۱۱ در رساله<sup>۹</sup> دکتری خود به استخراج روش‌های رایانشی در مورد زبان استعاری پرداخته است. وی با استفاده از دیدگاه کاملاً آماری و داده محور، کار مدل‌سازی را انجام داده و برای اولین بار سیستم حوزه باز تشخیص و تفسیر استعاره در متون نامحدود را ایجاد کرده است و سپس این روش را به حوزه‌های تفسیر مجاز نیز گسترش داده است.

---

<sup>۱</sup> KARMA

<sup>۲</sup> Att-Meta

<sup>۳</sup> Veale and Hao

<sup>۴</sup> Talking Points

<sup>۵</sup> Slipnet

<sup>۶</sup> slippage

<sup>۷</sup> analogies

<sup>۸</sup> Ekaterina VShutova



#### ۴- ارزیابی روش‌های تشخیص و تفسیر

معیارهای دقت<sup>۱</sup> و بازخوانی<sup>۲</sup> عمده‌ترین معیارهای ارزیابی تشخیص و تفسیر مفاهیم زبان استعاری هستند. این معیارها به صورت زیر تعریف می‌شوند:

$$\text{دقت} = \frac{a}{a+b} \quad (۱)$$

$$\text{بازخوانی} = \frac{a}{a+c} \quad (۲)$$

که  $a$  تعداد جملاتی است که به طور صحیح شناسایی و تفسیر شده‌اند،  $b$  تعداد جملات غیر استعاری است که به اشتباه برچسب جمله استعاری خورده‌اند و  $c$  تعداد جملات استعاری است که به اشتباه غیر استعاری شناسایی شده و تفسیر نشده‌اند.

#### ۵- جمع‌بندی و بحث

در این مقاله گلچینی از روش‌های رایانشی مهم در زمینه تشخیص و رفع ابهام معنایی مفاهیم استعاری و دارای مجاز در زبان انگلیسی به تفصیل تشریح شد. اما همانطور که قبلاً نیز بدان اشاره شد، اهمیت این زمینه تحقیقاتی برای علم پردازش زبان طبیعی و زبان‌شناسی رایانشی در ایران ناشناخته بوده و بدان خاطر مطالعاتی در این رابطه تا سال ۱۳۹۰ صورت نپذیرفته است. تا اینکه برای اولین بار پس از شکل‌گیری رشته زبان‌شناسی رایانشی در دانشگاه صنعتی شریف تحقیقات در این زمینه در مرکز زبان‌ها و زبان‌شناسی دانشگاه صنعتی شریف کلید خورد. در این تحقیقات سعی بر این است که برای زبان فارسی نیز بتوان انواع استعاره و مجاز را در جملات موجود در متون غیرادبی مشخص کرده و محتمل‌ترین و نزدیک‌ترین معنا را برای آنها استخراج کرد. در پژوهش مذکور دیدگاه‌ها و نظریه‌های مختلف زبان‌شناسی، روان‌شناسی، منطق، هوش مصنوعی در تشخیص و درک مفاهیم استعاری و مجازی به صورت دقیق مورد مطالعه و بررسی قرار می‌گیرد و امیدواریم که با تلفیقی از روش‌های معرفی شده در این رهگذر دستاوردی بهینه برای زبان فارسی حاصل شود.

نتیجه این تحقیق در بسیاری از مواردی که سامانه‌های مختلف پردازش زبان طبیعی مانند بازیابی اطلاعات، ترجمه ماشینی، استخراج اطلاعات، پاسخگوی خودکار، دسته‌بندی متون، درک گفتار، خوشه‌بندی متون، خلاصه‌سازی و تفسیر متون و همچنین عقیده‌کاوی یا تحلیل مقصود در زبان فارسی با ابهام روبرو می‌شوند،

<sup>۱</sup> precision

<sup>۲</sup> recall

کاربرد دارد. از این رو با تزریق نتایج حاصل از این پژوهش، قدرت و دقت عملکرد این سامانه‌ها به میزان قابل توجهی افزایش می‌یابد.

برداشتن هر گام اول در تولید یک ایده هیچگاه عاری از نقص نبوده و همواره دارای خطاهایی می‌باشد. در این تحقیق تلاش بر این است که تا حد امکان روش‌های دقیق و بدون نقصی در تشخیص و رفع ابهام استعاره و مجاز ارائه شود و حتی‌الامکان نتایج نهایی دارای کمترین خطا و ایراد باشد.

## فهرست منابع

عبدی قویدل، هادی. (مهرماه ۱۳۹۰). تشخیص خودکار و رفع ابهام معنایی مفاهیم استعاری و دارای مجاز. پروپوزال رساله کارشناسی ارشد، مرکز زبان‌ها و زبان‌شناسی، دانشگاه صنعتی شریف.

Barnden, J.A. (2012). *ATT-Meta Project Databank*.  
<http://www.cs.bham.ac.uk/~jab/ATT-Meta/Databank>.

Fass, D. (1991). met\*: A method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, 17(1):49–90.

Gedigian, M. and et.al.(2006) Catching metaphors. *In Proceedings of the 3rd Workshop on Scalable Natural Language Understanding*, pages 41–48, New York.

Goatly, A.(1997) *The Language of Metaphors*. Routledge, London.

Gruber, T.R. (1993). "A translation approach to portable ontology specifications". *Knowledge Acquisition* 5 (2): 199–220.

Krishnakumaran, S. and Zhu,X.(2007) Hunting elusive metaphors using lexical resources. *In Proceedings of the Workshop on Computational Approaches to Figurative Language*, pages 13–20, Rochester, NY.

Martin, H. (1990). *A Computational Model of Metaphor Interpretation*. Academic Press Professional, Inc., San Diego, CA, USA.

Mason, Z. C.( 2004) a computational, corpus-based conventional metaphor extraction system. *Computational Linguistics*, 30(1):23–44.

McCallum-Bayliss, H.(2011) *Metaphor Program Proposer*. INTELLIGENCE ADVANCED RESEARCH PROJECTS ACTIVITY (IARPA)  
<http://www.iarpa.gov>

- Narayanan, S. (1997). *Knowledge-based action representations for metaphor and aspect (KARMA)*. Technical report, PhD thesis, University of California at Berkeley.
- Peters, W. (2004). *Detection and Characterization of Figurative Language Use in WordNet*. MA Thesis, NLP Group, Department of Computer Science University of Sheffield, U.K.
- Shutova, E. (2010). Automatic Metaphor Interpretation as a Paraphrasing Task. In *Proceedings of NAACL 2010*, Los Angeles, USA
- Shutova, E. (2011). *Computational approaches to figurative language*. PhD thesis, Computer Laboratory, University of Cambridge, UK.
- Shutova, E. (2010). Models of Metaphor in NLP. In *Proceedings of ACL 2010*, Uppsala, Sweden.
- Shutova, E. (2009). Sense-based Interpretation of Logical Metonymy Using a Statistical Method. In *Proceedings of ACL 2009 Student Research Workshop*, Singapore.
- Shutova, E. and S. Teufel. (2009). Logical Metonymy: Discovering Classes of Meanings. In *Proceedings of the CogSci 2009 Workshop on Semantic Space Models*. Amsterdam, Holland.
- Shutova, E., L. Sun and A. Korhonen. (2010). Metaphor Identification Using Verb and Noun Clustering. In *Proceedings of COLING 2010*, Beijing, China.
- Shutova, E. and S. Teufel. (2010). Metaphor Corpus Annotated for Source – Target Domain Mappings. In *Proceedings of LREC 2010*, Valetta, Malta.
- Veale, T. and Hao, Y. (2008). A fluid knowledge representation for understanding and generating creative metaphors. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pages 945–952, Manchester, UK.

شماره تماس نویسندگان مقاله

تلفن همراه هادی عبدی قويدل: ۰۹۳۷۰۲۵۸۶۴۱

تلفن همراه محمد بحرانی: ۰۹۱۲۵۲۶۲۳۹۸

تلفن همراه بهرام وزیرنژاد: ۰۹۱۲۳۰۹۴۶۹۵