

مطالعه‌ی روش‌های تشخیص روابط معنایی غیرساختاری بین‌نودهای تشکیل دهنده‌ی پایگاه‌وجود

مر ترضی ذاکر^۱، محمد باقرمنه‌اج^۲

۱- دانشجوی دکتری مهندسی نرم افزار دانشگاه صنعتی امیرکبیر تهران و عضو هیئت علمی دانشگاه آزاد Smzaker@aut.ac.ir

۲- استاد تمام و عضو هیئت علمی دانشگاه صنعتی امیر کبیر تهران، Menhaj@aut.ac.ir

چکیده

در فرایند تشخیص و ساخت روابط، یکی از بنیادی‌ترین زمینه‌های علمی در سیستم‌های بازیابی اطلاعات اعم از آنتولوژی یا معنی فارسی آن " پایگاه وجود" است. یکی از مهمترین این روابط، روابط معنایی است که برای برنامه‌های کاربردی مختلف همچون پردازش متن، سیستم‌های مبتنی بر دانش و غیره مفیدند. یکی از منابعی که همه روابط معنایی را در خود جای داده است، پایگاه وجود است که برای ساختن آن نیز نیاز به تشخیص روابط معنایی مختلف من جمله روابط نان‌تاکسونومی، داریم. حال چه پایگاه وجود و چه منابع موجود دیگر که این روابط را دارند اغلب برای یک دامنه‌ی مشخص در دسترس نیستند. چاره این است که منابع مورد نیاز و روابط متشکله آن‌را به صورت دستی بسازیم. منتها ساختن به صورت دستی کار آسانی نیست و زمان‌بر بوده و از طرفی ایجاد روابط معنایی بین مفاهیم، پردازشی خطاپذیر از نظر ذهنی است. علاوه بر این، به روز رسانی این منابع هم فعالیت دستی زمان‌بری است. راه‌حلهایی برای استخراج روابط معنایی پایگاه وجود و یا سایر محصولات وابسته به آن تا کنون ارائه شده است. این پژوهش مروری نشان می‌دهد که علی‌رغم اینکه از پایگاه وجود جهت شناسایی مفاهیم و روابط بین این مفاهیم و خوشه‌بندی اطلاعات قلمروهای علمی به صورت خودکار استفاده شده است، لکن منابع آنها عمدتاً از منابع ساختاریافته بوده است و اکثراً روابط مفهومی نان‌تاکسونومی در آن بصورت دستی بوده است. ما ضمن تعریف پیشنهادی از یک پایگاه وجود، در صدد پاسخ به سوالاتی هستیم: آیا با استفاده از روابط می‌توان به حقیقت حقایق دست یافت؟ آیا ذهن انسان با کمک ماشین می‌تواند به سرعت به شناخت برسد؟ برای رسیدن به شناخت به روابطی متکی هستیم؟ نتایج مروری کامل روش‌های موجود در استخراج روابط معنایی و همچنین روابط نان‌تاکسونومی از منابع مختلف ورودی را ارائه کردیم که حاکی از آنست که چارچوبی مناسب که ترکیب و مقایسه متدهای استخراج مختلف را بتوانا سازد وجود ندارد و از طرف دیگر ورودی‌های ناسازگار و بررسی این ناسازگاری‌ها نیز یک چالش خاص بوده و پشتیبانی کلی در ابزارهای یادگیری پایگاه وجود از جنبه مهندسی، تکامل، استفاده مجدد، ادغام، تنظیم و تطبیق این پایگاه وجود ندارد. فلذا موارد یاد شده جنبه‌های مهمی برای تحقیقات در این زمینه است.

واژه‌های کلیدی: پایگاه وجود، روابط معنایی، رابطه نان‌تاکسونومی (غیر ساختاری)، رابطه تاکسونومی (ساختاری)،

معیارهای اندازه‌گیری شباهت و عدم شباهت (معنایی)

۱- مقدمه

در متون تحقیقی مختلف، آنتولوژی^۱ دارای معانی متنی متفاوتی است. یکی از معانی آن به اداراک رسمی اطلاعات وابسته به دامنه اشاره می‌کند [1] و یا یک مدل مفهومی زیربنایی یک دامنه‌ی خاص و توصیف مفهومی که وجود دارند به‌مراه خواص