

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه اصفهان
دانشکده فنی و مهندسی
گروه کامپیوتر

پایان نامه‌ی کارشناسی ارشد رشته‌ی مهندسی کامپیوتر گرایش نرم افزار

مدیریت برخورد در تطبیق آنتولوژی

استاد راهنما:
دکتر محمدرضا خیام باشی

پژوهشگر:
مهديه کارگرقوی

شهریورماه ۱۳۹۰

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات
و نوآوری‌های ناشی از تحقیق موضوع این پایان‌نامه
متعلق به دانشگاه اصفهان است.


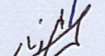
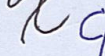


دانشگاه اصفهان
دانشکده فنی و مهندسی
گروه کامپیوتر

پایان نامه‌ی کارشناسی ارشد رشته‌ی مهندسی کامپیوتر گرایش نرم افزار
خانم مهدیه کارگرقوی تحت عنوان

مدیریت برخورد در تطبیق آنتولوژی

در تاریخ ۱۳۹۰/۰۶/۲۸ توسط هیأت داوران زیر بررسی و با درجه عالی به تصویب نهایی رسید.

- ۱- استاد/استادان راهنمای پایان نامه دکتر محمدرضا خیام باشی با مرتبه‌ی علمی استادیار امضا 
- ۲- استاد/استادان داور داخل گروه دکتر بهمن زمانی با مرتبه‌ی علمی استادیار امضا 
- ۳- استاد/استادان داور خارج از گروه دکتر سید رسول موسوی با مرتبه‌ی علمی استادیار امضا 

امضای مدیر گروه

سپاسگزاری

خدا را شاکرم که توفیق رفیق راه شد تا این پایان نامه تدوین گردد و بر خود لازم

می دانم تا مراتب تشکر خود را از استاد فرهیخته،

جناب آقای دکتر محمدرضا خیام باشی

به خاطر راهنمایی های بی دریغشان که اگر نمی بود این وظیفه تحقق پیدا نمی کرد، اعلام

نموده و تعالی مدارج علمی ایشان را از خداوند منان مسئلت می نمایم.

با تقدیم خالصانه به

پدر و مادر مهربانم

تکیه‌گاه سختی‌ها و سنگ‌صبور گرفتاری‌هایم.

همسر عزیزم

مشوق و حامی بی‌منت و پناهگاه امن همیشگی‌ام.

چکیده:

با پیشرفت فناوری‌های اطلاعاتی و مخابراتی، مقادیر فراوانی اطلاعات ناهمگن و توزیع‌شده فراهم شده است. وب معنایی سعی دارد تا این اسناد و اطلاعات را به‌فرم داده قابل فهم (دانش) برای ماشین نمایش دهد. یکی از راه‌های نمایش دانش در وب معنایی، آنتولوژی است که واژه‌ها و ارتباط بین آن‌ها را در یک دامنه خاص نشان می‌دهد. امروزه آنتولوژی‌ها به روش‌های مختلف و توسط توسعه‌دهندگان متفاوتی مدل می‌شوند، بنابراین همواره تفاوت‌هایی بین مفاهیم مدل‌شده در آنتولوژی‌ها وجود خواهد داشت که تعامل میان آن‌ها را مشکل می‌سازد.

تطبیق آنتولوژی‌ها، به‌عنوان زیرساختی مناسب برای انجام تعاملات در وب معنایی مطرح شده است و به فرآیند یافتن روابط و یا شباهت‌ها بین موجودیت‌های آنتولوژی‌های مختلف اشاره دارد. تاکنون تطبیق‌دهنده‌های بسیاری مطرح شده‌اند که در برخی موارد خوب عمل می‌کنند و در برخی موارد خوب عمل نمی‌کنند. این مساله، اهمیت خاصی به انتخاب تطبیق‌دهنده و ترکیب تطبیق‌دهنده‌ها می‌بخشد. از طرفی، تطبیق خودکار آنتولوژی‌ها به دلیل عدم تعامل با کاربر، نمی‌تواند نتایجی با کیفیت بالا به‌ویژه بر روی مجموعه داده‌های بزرگ حاصل نماید و این چالشی دیگر است که با عنوان عدم قطعیت در تطبیق آنتولوژی‌ها بیان شده است.

در این پایان‌نامه، یک متاسیستم تطبیق‌دهنده با نام MixAlign پیشنهاد شده است که از یک سو به ترکیب نتایج تطبیق‌دهنده‌های آنتولوژی‌ها می‌پردازد و از سوی دیگر به‌طور همزمان، مساله عدم قطعیت را رفع می‌کند تا خروجی تطبیق (هم‌ترازی)، بهبود یابد. معماری پیشنهادی این سیستم، متشکل از سه بخش با نام‌های تبدیل‌کننده فازی، تجمیع‌کننده فازی و استخراج‌کننده پویا می‌باشد. در اینجا با استفاده از تئوری مجموعه‌های فازی که یکی از روش‌های ارائه عدم قطعیت است، طبیعت غیرقطعی نگاشت‌ها مدل‌سازی می‌شود. در ادامه، این هم‌ترازی‌ها با تعریف درجه توافق مبتنی بر مرکز ثقل اعداد فازی ذوزنقه‌ای، ترکیب می‌شوند و سپس، نگاشت‌های مناسب جهت هم‌ترازی نهایی استخراج می‌شوند. بدین ترتیب با هم‌ترازی‌های ناقص، متناقض و مخرب، به‌نحوی قابل قبول رفتار می‌شود تا در نهایت بتوان از میان چندین هم‌ترازی ارائه شده برای دو آنتولوژی، به یک هم‌ترازی توافقی با دقت بیشتر رسید.

نتایج تجربی نشان داد که ترکیب هم‌ترازی‌های خروجی سیستم‌های تطبیق‌دهنده به‌کمک این روش، کیفیت و دقت واقعی هم‌ترازی نهایی را نسبت به سیستم‌های اولیه (قبل از ترکیب) و نسبت به سایر روش‌های تجمیع مشابهت بهبود می‌بخشد.

واژگان کلیدی: تجمیع مشابهت، تطبیق آنتولوژی‌ها، تئوری مجموعه‌های فازی، وب معنایی.

فهرست مطالب

صفحه

عنوان

فصل اول: کلیات

۱-۱- مقدمه	۱
۲-۱- ضرورت تطبیق آنتولوژی	۲
۳-۱- مثال	۶
۴-۱- بیان مساله پژوهشی	۷
۵-۱- اهداف تحقیق	۱۱
۶-۱- اهمیت و ارزش تحقیق	۱۱
۷-۱- کاربرد نتایج تحقیق	۱۲
۸-۱- ساختار پایان نامه	۱۵

فصل دوم: مروری بر ادبیات موضوع

۱-۲- مقدمه	۱۷
۲-۲- وب معنایی	۱۸
۱-۲-۲- مؤلفه‌های وب معنایی	۱۸
۳-۲- چهارچوب توصیف منابع (RDF)	۱۹
۴-۲- آنتولوژی	۲۰
۱-۴-۲- تاریخچه آنتولوژی	۲۰
۲-۴-۲- تعریف آنتولوژی	۲۱
۳-۴-۲- اجزاء آنتولوژی	۲۳
۴-۴-۲- زبان آنتولوژی	۲۵
۵-۲- تطبیق آنتولوژی‌ها	۲۶
۱-۵-۲- فرآیند تطبیق	۲۶
۲-۵-۲- نگاشت	۲۸
۳-۵-۲- هم‌ترازی	۲۸
۴-۵-۲- معماری یک سیستم تطبیق‌دهنده	۲۹
۶-۲- خلاصه فصل	۳۴

فصل سوم: تطبیق آنتولوژی

۳۶	۱-۳-۱- مقدمه
۳۷	۲-۳-۲- روش‌های تطبیق آنتولوژی
۳۷	۱-۲-۳- تطبیق نحوی در برابر تطبیق معنایی
۳۸	۲-۲-۳- دسته‌بندی رام و برنشتاین
۴۰	۳-۲-۳- دسته‌بندی شوایکو
۴۱	۴-۲-۳- دسته‌بندی‌های سه سطحی
۴۳	۱-۴-۲-۳- تکنیک‌های در سطح عنصر
۴۷	۲-۴-۲-۳- تکنیک‌های در سطح ساختار
۴۹	۳-۳- سیستم‌های تطبیق‌دهنده مشهور
۵۸	۴-۳- تکنیک‌های تجمیع مشابهت
۶۱	۱-۴-۳- نرم‌های مثلثی
۶۲	۲-۴-۳- فواصل چندبعدی و جمع وزنی
۶۸	۳-۴-۳- تجمیع فازی
۶۹	۵-۳- خلاصه فصل

فصل چهارم: ارائه راهکار پیشنهادی

۷۰	۱-۴-۱- مقدمه
۷۱	۲-۴-۲- عدم قطعیت و صحت نگاشت‌ها در تطبیق آنتولوژی
۷۲	۳-۴-۳- متاسیستم تطبیق‌دهنده
۷۳	۴-۴-۴- متاسیستم تطبیق‌دهنده پیشنهادی
۷۴	۱-۴-۴- درجه آزادی
۷۵	۲-۴-۴- عدم قطعیت
۷۶	۳-۴-۴- عدد فازی دوزنقه‌ای (TFN)
۷۷	۴-۴-۴- فرآیند پیشنهادی
۸۸	۵-۴-۵- مثال عددی
۹۷	۶-۴-۶- خلاصه فصل

فصل پنجم: ارزیابی کارایی و ارائه نتایج

۹۸	۱-۵-۱- مقدمه
----	--------------

۹۹	۲-۵- معیارهای ارزیابی کیفیت
۱۰۰	۱-۲-۵- دقت
۱۰۱	۲-۲-۵- فراخوانی
۱۰۱	۳-۲-۵- معیار - F
۱۰۲	۴-۲-۵- دلتا
۱۰۲	۵-۲-۵- دقت واقعی
۱۰۲	۳-۵- ارزیابی کیفیت تطبیق دهندگان
۱۰۳	۴-۵- مجموعه داده
۱۰۳	۱-۴-۵- مجموعه داده محک
۱۰۶	۲-۴-۵- بررسی آزمون‌های محک
۱۰۸	۳-۴-۵- مجموعه داده کنفرانس
۱۱۰	۵-۵- پیاده‌سازی
۱۱۱	۶-۵- ارزیابی راهکار پیشنهادی
۱۱۱	۱-۶-۵- نتایج تجربی حاصل شده بر روی مجموعه داده محک
۱۱۸	۲-۶-۵- نتایج تجربی حاصل شده بر روی مجموعه داده کنفرانس
۱۱۹	۳-۶-۵- تحلیل نتایج
۱۲۱	۷-۵- خلاصه فصل

فصل ششم: نتیجه‌گیری و پیشنهادها برای کارهای آینده

۱۲۲	۱-۶- مقدمه
۱۲۲	۲-۶- خلاصه مباحث و نتایج بدست آمده
۱۲۴	۳-۶- کارهای آینده
۱۲۶	منابع و مأخذ

فهرست شکل‌ها

عنوان	صفحه
شکل ۱-۱ دو شمای ساده XML	۷
شکل ۲-۱ بخش‌هایی از دو آنتولوژی	۱۰
شکل ۳-۱ نمایی گرافیکی از تکه‌های ساختاری کاربردهای تطبیق	۱۲
شکل ۴-۱ ادغام آنتولوژی	۱۳
شکل ۵-۱ ترجمه داده	۱۴
شکل ۶-۱ واسطه‌گری پرس‌وجو	۱۵
شکل ۱-۲ یک لایه‌ای تکنولوژی وب معنایی و استانداردهای تشکیل‌دهنده آن	۱۹
شکل ۲-۲ مراحل ایجاد Taxonomy، Ontology و Knowledge Base	۲۵
شکل ۳-۲ هم‌ترازسازی آنتولوژی‌ها	۲۷
شکل ۴-۲ فرآیند تطبیق آنتولوژی	۲۷
شکل ۵-۲ نمونه‌ای از هم‌ترازسازی دو آنتولوژی	۲۹
شکل ۶-۲ معماری یک سیستم هم‌ترازسازی	۳۰
شکل ۷-۲ ترکیب سریال تکنیک‌های تطبیق	۳۱
شکل ۸-۲ ترکیب موازی تکنیک‌های تطبیق	۳۲
شکل ۹-۲ مراحل تفسیر نگاشت‌ها	۳۳
شکل ۱-۳ تطبیق نحوی در برابر تطبیق معنایی	۳۷
شکل ۲-۳ دسته‌بندی رام و برنشتاین از روش‌های تطبیق	۳۸
شکل ۳-۳ دسته‌بندی شوایکو از روش‌های تطبیق مبتنی بر شما	۴۱
شکل ۴-۳ دسته‌بندی سه سطحی از روش‌های اصلی تطبیق	۴۳
شکل ۵-۳ یک فرآیند تطبیق معمول	۵۸
شکل ۶-۳ رفتار تابع سیگموید	۶۶
شکل ۱-۴ معماری پیشنهادی متاسیستم تطبیق‌دهنده	۷۴
شکل ۲-۴ تابع عضویت عدد فازی دوزنقه‌ای $A=(a,b,c,d)$	۷۷
شکل ۳-۴ مقادیر پیشنهادی آستانه، حاصل از اجرای الگوریتم جهت ترکیب سیستم‌هایی با تعداد متفاوت	۸۳
شکل ۴-۴ ترسیم نمونه‌هایی از هیستوگرام و منحنی نرمال مقادیر آستانه	۸۵

عنوان

صفحه

- شکل ۴-۵ ترسیم نمونه‌هایی از خطوط بدست آمده از اتصال مقادیر آستانه که به صورت نزولی مرتب شده‌اند
۸۶.....
- شکل ۴-۶ دو کلاس هم‌ترازی بین بخش‌هایی از دو آنتولوژی O_1 و O_2
۸۸.....
- شکل ۴-۷ ماتریس‌های مشابهت چهار نمونه هم‌ترازی.....
۸۹.....
- شکل ۴-۸ خروجی بخش تبدیل در معماری پیشنهادی برای چهار سیستم تطبیق‌دهنده.....
۹۰.....
- شکل ۴-۹ درجات توافق روی نگاشت‌ها بین هر زوج از سیستم‌های تطبیق‌دهنده.....
۹۰.....
- شکل ۴-۱۰ درجه توافق متوسط هر سیستم روی نگاشت‌ها.....
۹۱.....
- شکل ۴-۱۱ درجه توافق نسبی هر سیستم روی نگاشت‌ها.....
۹۱.....
- شکل ۴-۱۲ مولفه‌های عقیده توافقی $T_m = (a_m, b_m, c_m, d_m)$
۹۲.....
- شکل ۴-۱۳ ماتریس مشابهت یک نمونه هم‌ترازی.....
۹۳.....
- شکل ۴-۱۴ نتایج حاصل شده از ترکیب دو هم‌ترازی Nearmiss و farone با روش‌های تجمیع مختلف.....
۹۳.....
- شکل ۴-۱۵ نتایج حاصل شده از ترکیب سه هم‌ترازی Nearmiss, farone, A1 با روش‌های تجمیع مختلف.....
۹۴.....
- شکل ۴-۱۶ نتایج حاصل شده از ترکیب چهار هم‌ترازی Nearmiss, farone, A1, A2 با روش‌های تجمیع مختلف.....
۹۴.....
- شکل ۴-۱۷ نتایج حاصل شده از ترکیب پنج هم‌ترازی Nearmiss, farone, A1, A2, A3 با روش‌های تجمیع مختلف.....
۹۵.....
- شکل ۵-۱ دو هم‌ترازی به صورت یک مجموعه از شباهت‌ها و روابط بین آن‌ها.....
۱۰۰.....
- شکل ۵-۲ فرآیند ارزیابی پایه.....
۱۰۳.....
- شکل ۵-۳ ساختار سلسله مراتبی کلاس‌های آنتولوژی مرجع با نرم‌افزار protégé.....
۱۰۵.....
- شکل ۵-۴ مشخصات آزمون‌های محک.....
۱۰۷.....
- شکل ۵-۵ عملکرد روش پیشنهادی روی دو سیستم بالارته و با مجموعه داده محک 1xx.....
۱۱۲.....
- شکل ۵-۶ عملکرد روش پیشنهادی روی چهار سیستم با رتبه متوسط و با مجموعه داده محک 1xx.....
۱۱۳.....
- شکل ۵-۷ عملکرد روش پیشنهادی روی دو سیستم بالارته و با مجموعه داده محک 2xx.....
۱۱۴.....
- شکل ۵-۸ عملکرد روش پیشنهادی روی چهار سیستم با رتبه متوسط و با مجموعه داده محک 2xx.....
۱۱۴.....
- شکل ۵-۹ عملکرد روش پیشنهادی روی دو سیستم بالارته و با مجموعه داده محک 3xx.....
۱۱۵.....
- شکل ۵-۱۰ عملکرد روش پیشنهادی روی چهار سیستم با رتبه متوسط و با مجموعه داده محک 3xx.....
۱۱۶.....
- شکل ۵-۱۱ عملکرد روش پیشنهادی روی دو سیستم بالارته و کل مجموعه داده محک.....
۱۱۷.....

صفحه

عنوان

- شکل ۵-۱۲ عملکرد روش پیشنهادی روی چهار سیستم با رتبه متوسط و کل مجموعه داده محک ۱۱۷
- شکل ۵-۱۳ عملکرد روش پیشنهادی روی سه سیستم بالارترتبه در مجموعه داده کنفرانس ۱۱۸
- شکل ۵-۱۴ عملکرد روش پیشنهادی روی سه سیستم با رتبه متوسط در مجموعه داده کنفرانس ۱۱۹

فهرست جدول‌ها

عنوان	صفحه
جدول ۱-۲ جدول هم‌ترازی میان دو آنتولوژی	۲۹
جدول ۱-۳ دسته‌بندی تطبیق‌دهنده‌ها بر مبنای دسته‌بندی شوایکو	۵۷
جدول ۲-۳ دسته‌بندی تطبیق‌دهنده‌ها بر مبنای دسته‌بندی رام و برنشتاین	۵۸
جدول ۳-۳ ماتریس مشابهت حاوی فاصله لونشتاین نرمال شده	۵۹
جدول ۴-۳ ماتریس مشابهت حاوی فاصله مبتنی بر وردنت	۵۹
جدول ۵-۳ ماتریس مشابهت حاصل از تجمیع مینیمم	۵۹
جدول ۱-۴ مقادیر مجموعه سنجش زبانی	۷۹
جدول ۱-۵ توصیف مجموعه داده محک	۱۰۸
جدول ۲-۵ مجموعه آنتولوژی‌های کنفرانس	۱۰۹
جدول ۳-۵ مشخصه‌های دو گروه از آزمون‌های OAEI	۱۱۰

فصل اول

کلیات

۱-۱- مقدمه

با پیشرفت فناوری‌های اطلاعاتی و مخابراتی، مقادیر فراوانی اطلاعات ناهمگن^۱ و توزیع شده فراهم شده است. با توجه به آن که تعداد منابع جمع‌آوری‌کننده اطلاعات به سرعت رو به افزایش است، مساله مدیریت ناهمگنی اطلاعات در بین این منابع نیز، رو به رشد است.

افزایش قابل توجه در داده‌ها، دانش و رشد اینترنت، به خصوص سرویس‌های وب، نیاز به قابلیت تعامل^۲ را در میان منابع ناهمگن بوجود آورده است [۱]. در اجتماع وب معنایی^۳، قابلیت تعامل به امری بسیار مهم و حیاتی تبدیل شده است؛ وب معنایی در رویای یک وب از اطلاعات قابل فهم برای ماشین به سر می‌برد، که می‌تواند به طور خودکار توسط کاربردهای معنایی دستیابی و مبادله شوند. طراحان وب، هم‌چنان با مشکل قابلیت تعامل معنایی روبرو هستند، که در مسیر استفاده از کلیه امکانات بالقوه وب قرار دارد. به دنبال آن راه‌حل‌های مختلفی

¹ Heterogeneous

² Interoperability

³ Semantic web

برای پرداختن به این مساله و به خصوص در جهت خودکارسازی تجمیع^۱ منابع اطلاعاتی ناهمگن و توزیع شده، پیشنهاد شده‌اند.

آنتولوژی^۲، یکی از راه‌های نمایش دانش^۳ در وب معنایی است. آنتولوژی، مشخصه‌ای است که کلمات، مفاهیم و اصطلاحات یک حوزه را تعریف می‌نماید. دقت این مشخصه و نوع بیان آن، بسته به کاربرد می‌تواند متفاوت باشد. نقش آنتولوژی در وب معنایی آن است که واژه‌ها و ارتباط بین آن‌ها را در یک دامنه خاص نشان دهد. به عبارت دیگر، ارتباط بین مفاهیم در اسناد وب و دنیای واقعی را نشان می‌دهد تا ماشین‌ها قادر به پردازش و اشتراک‌گذاری این اطلاعات باشند. آنتولوژی‌ها در همه‌جا حضور دارند [۲]. امروزه ایجاد و بیان آنتولوژی‌ها در محیط، با استفاده از ابزارهای متفاوت و با دید و تجریده‌های متفاوتی دنبال می‌شود. در نتیجه آن، در هر حوزه، آنتولوژی خاص خود تولید می‌شود و امید به ایجاد یک آنتولوژی واحد و سراسری، نه تنها قابل دسترس بلکه خواستنی نیز به نظر نمی‌رسد [۲، ۳]. زمانی که آنتولوژی‌های رقابتی^۴ متعددی در برنامه‌های کاربردی متفاوت استفاده می‌شوند، نمی‌توانند آن‌گونه که لازم است، با هم تعامل داشته باشند. در این میان، تطبیق آنتولوژی^۵، در زمینه تکنولوژی‌های معنایی، توجه بسیاری را به خود جلب کرده است و یکی از محتمل‌ترین راهکارها جهت چیرگی بر مشکلات ناهمگنی محتوای آنتولوژیکی وب است.

بخش ۱-۲ به ضرورت تطبیق آنتولوژی می‌پردازد. در بخش ۱-۳ با یک مثال ساده، مساله تطبیق آنتولوژی شرح داده می‌شود. در ادامه در بخش ۱-۴ مساله پژوهشی مطرح می‌شود و سپس در بخش‌های ۱-۵، ۱-۶ و ۱-۷ به ترتیب اهداف تحقیق، اهمیت و ارزش آن و کاربرد نتایج تحقیق بیان می‌شود. در نهایت ساختار پایان‌نامه ارائه می‌گردد.

۱-۲- ضرورت تطبیق آنتولوژی

در دهه‌های اخیر، آنتولوژی‌های فراوانی در حوزه‌های متفاوتی از علوم کامپیوتر مانند مدیریت دانش، بازیابی اطلاعات، چند رسانه‌ای، مهندسی نرم‌افزار و سرویس‌های وب ایجاد شده‌اند. جنبه‌های مختلفی از یک آنتولوژی وجود دارد؛ آنتولوژی می‌تواند شامل داده‌های متعدد یا مدل‌های مفهومی باشد مانند طبقه‌بندی‌ها^۶ و شمای‌های

^۱ Aggregation

^۲ Ontology

^۳ Knowledge representation

^۴ Competitive

^۵ Ontology matching

^۶ Classification

پایگاه داده^۱، یا شامل تئوری‌های بیان شده در قالب مجموعه‌ای گزاره^۲ باشد. به‌طور طبیعی، توسعه‌ی نامتمرکز وب معنایی، رشد نمایی تعداد آنتولوژی‌ها و کاربردهای مبتنی بر آنتولوژی، امری قابل پیش‌بینی است. چون همیشه بیشتر از یک نمایش برای حوزه‌ی تحت بررسی وجود دارد، بسیاری از این آنتولوژی‌ها، حوزه‌های مشابهی را هر چند با استفاده از ادبیات مختلف توصیف می‌کنند و بقیه آنتولوژی‌ها ممکن است همپوشانی حوزه‌ای داشته باشند [۴]. بنابراین ناهمگنی در محیطی مانند وب معنایی اجتناب‌ناپذیر است. ناهمگنی آنتولوژی‌ها می‌تواند از موارد متعددی ناشی شود که مهم‌ترین آن‌ها عبارتند از [۲]:

- **ناهمگنی نحوی**^۳: به‌علت نداشتن زبان مشترک در مدل‌سازی دو آنتولوژی، یا تفاوت در نحوه نمایش دانش بوجود می‌آید؛ مثلاً یکی به زبان OWL^۴ و دیگری به زبان F-Logic بیان شده است.
- **ناهمگنی لغوی**^۵: ناشی از نام‌گذاری‌های مختلف برای یک مفهوم یکسان در یک زبان و یا ناهمگنی بین زبان‌های طبیعی مختلف و یا زبان‌های فنی مختلف می‌باشد.
- **ناهمگنی مفهومی**^۶: این نوع ناهمگنی مربوط به مدل‌سازی‌های متفاوت از یک دامنه و یا تفاوت‌هایی در بیان مفاهیم می‌باشد. این نوع ناهمگنی‌ها به‌عنوان عدم تطبیق‌های منطقی یا ناهمگنی‌های معنایی نیز مورد بررسی قرار می‌گیرند. تفاوت در مدل‌سازی یک دامنه، ناهمگنی معنایی نامیده می‌شود. این تفاوت می‌تواند در قواعد و اصول اولیه باشد که از آن برای تعریف دامنه استفاده می‌گردد. سه دلیل اصلی تفاوت در مدل‌سازی معنایی عبارتند از [۵]:
 - ✓ تفاوت در پوشش^۷: زمانی که دو آنتولوژی، دو ناحیه متفاوت، ولی احتمالاً دارای اشتراک از یک دامنه را با یک سطح از جزئیات و یک منظر مدل می‌کنند.
 - ✓ تفاوت در دانه‌بندی^۸: وقتی دو آنتولوژی، یک ناحیه را از یک منظر، اما در سطوح مختلفی از جزئیات، مدل می‌نمایند.
 - ✓ تفاوت در منظر^۹ یا تفاوت در حوزه: زمانی رخ می‌دهد که دو آنتولوژی، یک ناحیه را، در یک سطح از جزئیات، اما از منظرهای متفاوت مدل می‌نمایند.

¹ Database schema

² Axiom

³ Syntactic heterogeneity

⁴ Ontology Web Language

⁵ Terminological heterogeneity

⁶ Conceptual heterogeneity

⁷ Coverage

⁸ Granularity

⁹ Perspective

- **ناهمگنی ماورای منطق**^۱: در این نوع ناهمگنی که ناهمگنی عملی^۲ می‌باشد، دو موجودیت که به لحاظ معنایی کاملاً یکی می‌باشند، ممکن است توسط افراد مختلف با ذهنیت‌های مختلف، متفاوت تفسیر شوند و یا در زمینه‌های مختلف، متفاوت تفسیر شوند. چون این اطلاعات زمینه‌ای، خارج از دسترس کامپیوتر قرار دارند، تفسیر انطباق یا عدم انطباق دو آنتولوژی، مشکل به نظر می‌رسد. در تطبیق خودکار آنتولوژی‌ها، به این نوع ناهمگنی پرداخته نمی‌شود.

اکثر ابزارهای تطبیق امروزی بر روی ناهمگنی در سطح لغوی یا مفهومی یا هر دوی آن عمل می‌کنند [۲]. هدف تطبیق آنتولوژی‌ها آن است که بر ناهمگنی معنایی بین آن‌ها غلبه شود. تنوع ناهمگنی‌ها و این حقیقت که آن‌ها اغلب با هم اتفاق می‌افتند، شیوه‌های متعددی را برای تطبیق آنتولوژی‌ها پدید آورده است که در فصل سوم بیشتر به آن‌ها پرداخته شده است. به‌منظور تطبیق آنتولوژی‌ها باید مشابهت‌ها^۳ یا روابط^۴ موجود میان مولفه‌های دو آنتولوژی مشخص گردند. در ادامه، انواع روش‌های محاسبه مشابهت که بر اساس خصوصیات مختلفی از آنتولوژی‌ها پدید آمده‌اند، به‌طور خلاصه مورد بررسی قرار می‌گیرند [۲]:

۱. تکنیک‌های مبتنی بر شباهت لغوی: این تکنیک‌ها خود به دو دسته تقسیم می‌شوند که عبارتند از روش‌های مبتنی بر شباهت‌های رشته‌ای^۵ و روش‌های مبتنی بر شباهت‌های زبان‌شناختی^۶. این تکنیک‌ها وابسته به مجموعه اصطلاحات هستند. روش‌های مبتنی بر رشته، بر روی رشته‌ها بر اساس دنباله‌ای از حروف عمل می‌کنند. روش‌های مبتنی بر رشته معمول، روش‌های فاصله‌ویرایشی^۷ و شباهت رشته‌ای هستند. تکنیک‌های زبان‌شناختی با در نظر گرفتن نام‌ها به‌عنوان کلمات، روش‌های پردازش زبان طبیعی را به کار می‌گیرند. این روش‌ها قبل از تکنیک‌های مبتنی بر رشته، به‌عنوان یک فاز پیش‌پردازش به کار برده می‌شوند، برای نمونه می‌توان جداسازی نشانه‌ها^۸ و پیدا کردن یک کلمه در فرهنگ‌لغت یا لیستی از کلمات را نام برد.

۲. تکنیک‌های مبتنی بر شباهت ساختاری: اطلاعات ساختاری موجودیت‌های تشکیل‌دهنده یک آنتولوژی را می‌توان به دو دسته ساختاری داخلی و ساختاری ارتباطی تقسیم نمود. به‌عنوان مثال

¹ Semiotic

² Pragmatic

³ Similarity

⁴ Relation

⁵ String-based

⁶ Linguistic-based

⁷ Edit Distance

⁸ Lemmatization

مشابهت میان فیلدها و تعداد و نوع آنها میان دو موجودیت از آنتولوژی‌ها، در دسته شباهت‌های ساختاری داخلی قرار می‌گیرند. در تکنیک‌های ساختاری ارتباطی، روابط میان موجودیت‌ها در نظر گرفته می‌شود. مهم‌ترین حالت نمایشی تکنیک‌های ساختاری، تکنیک‌های مبتنی بر گراف و مبتنی بر طبقه‌بندی هستند. تکنیک‌های مبتنی بر گراف، آنتولوژی‌ها را به‌عنوان ساختارهای یک گراف برچسب‌گذاری شده در نظر می‌گیرند. در این شیوه، مشابهت دو موجودیت را بر اساس مشابهت همسایه‌های آنها بدست می‌آورند. تکنیک‌های مبتنی بر طبقه‌بندی، حالت خاصی از تکنیک‌های مبتنی بر گراف هستند؛ با این تفاوت که به تشخیص روابط بین نودها می‌پردازند.

۳. تکنیک‌های مبتنی بر نمونه‌ها: اگر برای هر موجودیت در آنتولوژی، مجموعه‌ای از نمونه‌ها وجود داشته باشد، می‌توان از این نمونه‌ها برای محاسبه مشابهت میان موجودیت‌ها استفاده کرد. در واقع نمونه‌های آنتولوژی را به‌منظور پیدا کردن ترتیب یا اختلافی در سطوح داده، مورد بررسی قرار می‌دهند.

۴. تکنیک‌های معنایی: ویژگی اصلی این روش، قدرت استنباطی آن در تولید روابط جدید و یافتن ناسازگاری‌ها پس از تطبیق آنتولوژی می‌باشد. تکنیک‌های معنایی بر اساس تفسیر معنایی دو موجودیت یکسان که توصیفات را به اشتراک دارند، در نظر گرفته می‌شوند. این روش‌ها بر اساس منطقی توصیفی یا گزاره‌ای^۱ هستند.

تکنیک‌های پایه ارائه شده در بالا، تکه‌های^۲ ساختاری هستند که یک راه‌حل تطبیق با آنها ساخته می‌شود. از آنجا که اندازه‌ی داده‌های دنیای واقعی بسیار زیاد است و منابع بسیار بزرگی با هزارها مفهوم و ویژگی وجود دارند، تشخیص دستی مشابهت‌ها میان این منابع، یک امر غیرممکن است. یک سیستم تطبیق‌دهنده^۳، سعی دارد تا دانشی را که به‌فرم صریح در آنتولوژی‌ها کدشده، مورد بهره‌برداری قرار دهد. جنبه‌های زیر برای ساخت یک سیستم تطبیق‌دهنده در نظر گرفته شده‌اند [۲]:

۱. محاسبه مشابهت بین موجودیت‌های دو آنتولوژی به‌کمک تکنیک‌های پایه

۲. سازمان‌دهی نحوه تجمیع نتایج برای محاسبه مشابهت مرکب بین موجودیت‌ها (به آن الگوریتم‌های تطبیق نیز می‌گویند)

۳. توسعه یک شیوه برای محاسبه این مشابهت‌ها

¹ Propositional

² Block

³ Matching system