

ارائه روشی برای تعیین موقعیت ماشینی آدرس‌های توصیفی

علی جاویدانه^{۱*}، فرید کریمی‌پور^۲

^۱دکتری سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی

- دانشگاه تهران

javidaneh.ali@gmail.com

^۲استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی - دانشگاه تهران

fkarimpour@ut.ac.ir

(تاریخ دریافت آذر ۱۳۹۵، تاریخ تصویب دی ۱۳۹۸)

چکیده

آدرس قدیمی‌ترین و پرکاربردترین شناسه مکانی محسوب می‌شود و در مقایسه با شناسه‌های مکانی دیگر مانند مختصات، مورد استفاده بیشتر قرار گرفته است. آدرس‌ها در دنیا به اعتبار وجود یا عدم وجود استاندارد، به دو دسته آدرس‌های استاندارد و آدرس‌های غیراستاندارد یا توصیفی تقسیم می‌شوند. از طرفی غالب روش‌های موجود برای تعیین موقعیت ماشینی آدرس‌ها، برای آدرس‌ها با ساختار استاندارد توسعه داده شده‌اند. اما از آنجا که آدرس‌های توصیفی دارای پیچیدگی ساختاری و معنایی بیشتری نسبت به آدرس‌های استاندارد بوده، لذا تعامل ماشین با این دسته از آدرس‌ها، دشوارتر می‌باشد. البته علی‌رغم دشواری تعامل ماشین با این گونه آدرس‌ها، آدرس‌های توصیفی مزایایی نسبت به آدرس‌های استاندارد دارند که باعث می‌شود به جای آن که به دنبال جایگزینی آن‌ها با آدرس‌های استاندارد باشیم، به توسعه روشی برای حل مسئله تعیین موقعیت ماشینی آدرس‌های توصیفی بپردازیم. از این رو، هدف اصلی این تحقیق، ارائه روشی برای تعیین موقعیت ماشینی آدرس‌های توصیفی می‌باشد تا اولاً چالش تعیین موقعیت ماشینی این آدرس‌ها برطرف شود ثانیاً امکان بهره‌مندی از مزایای آدرس‌های توصیفی با جایگزینی آن‌ها با آدرس‌های استاندارد، از بین نرود. رویکرد این تحقیق برای حل مسئله آدرس‌یابی ماشینی برای آدرس‌های توصیفی شامل دو بخش اصلی «ایجاد هستی‌شناسی گرامر رسمی زبان آدرس‌های توصیفی» برای تجزیه متن آدرس و «طراحی و توسعه یک مدل داده مکانی غنی شده با ادراک مکانی» برای تعیین موقعیت یا انطباق مکانی اجزای آدرس توصیفی می‌باشد. بر اساس رویکرد پیشنهادی تحقیق، پیاده‌سازی جهت نمایش قابلیت اجرا و ارزیابی روش انجام گرفت. در نهایت، رویکرد حل مسئله در بهره‌گیری از مزایای آدرس‌های توصیفی به جای جایگزینی این آدرس‌ها با آدرس‌های استاندارد، باعث انعطاف‌پذیری راه حل شده و دامنه کاربرد آدرس‌های توصیفی را گسترش داده است.

واژگان کلیدی: آدرس‌یابی، سیستم آدرس‌دهی، آدرس توصیفی، ادراک مکانی، پردازش زبان طبیعی

۱- مقدمه

آدرس‌ها متداول‌ترین منابع برای دستیابی افراد به مکان‌های جغرافیایی هستند و در مقایسه با شناسه‌های مکانی دیگر، مانند مختصات، بیشتر مورد استفاده قرار گرفته‌اند [۱]. آدرس، ابزار مناسبی برای مکان‌مند کردن پایگاه‌های اطلاعاتی گوناگون موجود محسوب می‌شود [۲]. اهمیت آدرس و نیاز به تعیین موقعیت آدرس‌ها منجر شده تا کشورهای توسعه یافته با صرف هزینه و زمان قابل توجه، به دنبال تولید و نگهداری پایگاه‌های داده‌ی با کیفیت و قابل اعتماد از مکان آدرس‌های موجود در کشور خود باشند. به عنوان نمونه، دفتر سرشماری وزارت تجارت^۱ ایالات متحده آمریکا از دهه ۹۰ میلادی، با طراحی و ایجاد TIGER^۲، اطلاعات مکانی موجود در پایگاه‌های داده خود را در قالب پایگاه داده مکانی استخراج و منتشر می‌کند [۳]. همچنین سازمان نقشه‌برداری انگلستان با ایجاد پایگاه داده Address-Point، مختصات چندین میلیون آدرس را در سراسر انگلیس، اسکاتلند و ولز تولید و نگهداری می‌کند [۴]. از طرفی دولت استرالیا نیز طرح GNAF^۳ را برای دستیابی به یک فایل آدرس-دهی دقیق و قابل اعتماد تصویب کرد، که از سال ۱۹۹۸ میلادی به مدت شش سال مطالعات آن انجام شده و از سال ۲۰۰۴ میلادی هم در مدت دو سال اجرا گردید [۵]. این گونه فعالیت‌های زانبر و پرهزینه دولت‌ها، نشان-دهنده اهمیت آدرس و مدیریت آن می‌باشد.

آدرس‌ها در دنیا به اعتبار وجود یا عدم وجود استاندارد، به دو دسته آدرس‌های استاندارد و آدرس‌های غیراستاندارد یا توصیفی تقسیم می‌شوند. آدرس‌های غیراستاندارد، غالباً ماهیت توصیفی دارند، به این معنا که از یک مکان آشنا یا به آسانی قابل دسترسی، برای مخاطب، شروع شده و با توصیف مسیر و مقصد، مرحله به مرحله به مخاطب کمک می‌کنند تا مکان مورد اشاره در آدرس را بیابد. لازم به یادآوری است که این آدرس‌های توصیفی براساس آنچه شخص از محیط پیرامون خود تا مکان مورد اشاره در آدرس درک می‌کند، بیان می‌شوند. به تعبیر دیگر، آدرس‌های توصیفی بر اساس مؤلفه‌های

ادراک مکانی^۴ و در قالب ساختار زبان طبیعی^۵ می‌باشند. بدین معنا که در تبادل آدرس، ادراک مکانی دو عامل (آدرس‌دهنده و گیرنده) دخیل است [۶].

از طرفی غالب روش‌های موجود برای تعیین موقعیت ماشینی آدرس‌ها، برای آدرس‌ها با ساختار استاندارد توسعه داده شده‌اند.

آدرس‌های توصیفی، در مقایسه با آدرس‌های استاندارد، نیاز به ابزارهای کمک ناوبری را کاهش می‌دهند. از طرفی این گونه آدرس‌ها، باعث توجه بیشتر شخص به محیط و در نتیجه، شکل‌گیری نقشه ادراکی و ارتقای سطح مهارت‌های مکانی کاربر می‌شوند [۷].

چنین مزایایی باعث می‌شود به جای آن که به دنبال جایگزینی آدرس‌های توصیفی با آدرس‌های استاندارد باشیم، به توسعه روشی برای حل مسئله تعیین موقعیت ماشینی آدرس‌های توصیفی بپردازیم. از این رو این تحقیق با توجه به خلاء موجود، در پی آن است که روشی را مبتنی بر ادراک مکانی و پردازش زبان‌های طبیعی^۶ برای تعیین موقعیت ماشینی آدرس‌های توصیفی توسعه دهد.

بر این اساس، در ادامه در بخش دوم به موضوع آدرس و آدرس‌یابی پرداخته شده است. در بخش سوم برخی تحقیقات مرتبط با تحقیق حاضر در حوزه‌های «آدرس‌یابی و کدگذاری مکانی» و «توصیفات مکانی» با هدف مطالعه تحقیقات پیشین مورد بررسی و مطالعه قرار گرفته‌اند. همانطور که ذکر شد، دو رکن اصلی برای حل مسئله تعیین موقعیت ماشینی آدرس‌های توصیفی، موضوعات «ادراک مکانی» و «پردازش زبان طبیعی» می‌باشند. این دو موضوع به عنوان مبانی نظری تحقیق در بخش چهارم معرفی شده‌اند.

در بخش پنجم رویکرد پیشنهادی تحقیق معرفی شده است. بر اساس مطالعات صورت گرفته در بخش‌های پیش از آن، دو مرحله اصلی «ایجاد هستی‌شناسی گرامر رسمی زبان آدرس‌های توصیفی» و «طراحی و توسعه یک مدل داده مکانی غنی شده با ادراک مکانی» به عنوان مراحل اصلی رویکرد حل مسئله تحقیق معرفی شده‌اند. در بخش ششم پیاده‌سازی راه حل پیشنهادی جهت تفسیر و

^۴ Spatial cognition

^۵ Natural language

^۶ Natural Language Processing (NLP)

^۱ U.S. department of commerce - Census Bureau

^۲ Topologically Integrated Geographic Encoding and Referencing

^۳ Geocoded National Address File (GNAF)

تفاوت سیستم‌های آدرس‌دهی منجر به تفاوت اجزای آدرس‌ها و نحوه ترکیب آن‌ها با یکدیگر می‌شود. این تفاوت‌ها هم غالباً ریشه‌های فرهنگی، اجتماعی و حتی تاریخی دارند.

به عنوان نمونه در غالب کشورهای غربی، سیستم آدرس‌دهی شامل نام بلوک ساختمانی یا معبر و شماره ساختمان می‌باشد. در این کشورها غالباً نام بلوک‌ها و معابر یکتا بوده و شماره ساختمان‌ها هم از نظر مکانی، مرتب می‌باشند. در چنین سیستم‌های آدرس‌دهی، قالب و اجزا و حتی علائم سجاوندی به کار رفته در آدرس، دقیقاً از پیش تعریف شده است.

بر این اساس، در چنین شرایطی، با معرفی علائم سجاوندی به کار رفته در این آدرس به ماشین، می‌توان به راحتی عمل تجزیه متن آدرس به اجزای آن را انجام داد. همچنین از آن‌جا که اجزا و تفسیر آن‌ها، کاملاً روشن است، لذا به راحتی عمل انطباق مکانی اجزای آدرس نیز انجام می‌شود.

از طرف دیگر، در برخی کشورها از مؤلفه‌های غیرمکانی برای تعریف سیستم آدرس‌دهی استفاده شده است. در این کشورها، از کدگذاری‌های عددی که گاه بر اساس زمان و حتی گاه به صورت تصادفی تولید می‌شوند، استفاده می‌گردد. این نوع کدگذاری اگرچه ساختاریافته بوده و برای ماشین قابل تعریف است اما به تنهایی و بدون ابزارهای کمکی برای انسان قابل استفاده نیست. سیستم آدرس‌دهی در ژاپن و کره جنوبی، نمونه‌ای از این حالت سیستم آدرس‌دهی هستند. [۱۷]

اما در کشورهایی مانند ایران، ترکیه، برزیل و هند، آدرس‌ها معمولاً در قالب توصیفی از مسیر دسترسی به نقطه مورد نظر بیان می‌شوند از این رو مؤلفه‌های گوناگونی مانند نشانه‌های زمینی^۲ یا روابط مکانی در چنین سیستم‌های آدرس‌دهی وجود دارد. در واقع در این سیستم‌های آدرس‌دهی، قالب از پیش تعریف شده وجود ندارد. حتی مؤلفه‌های به کار رفته در آدرس نیز از قبل تعیین نشده است. همچنین در این‌گونه سیستم‌ها معمولاً نام معابر و حتی محلات در سطح یک شهر، یکتا نمی‌باشد. از این رو حتی اگر بتوان متن چنین آدرسی را به اجزای آن تجزیه نمود، باز هم با یک تطابق ساده با نام اماکن و

ارزیابی نتایج عملی ارائه شده است. در این بخش، ابتدا محیط پیاده‌سازی و منطقه مورد مطالعه معرفی گردیده است. سپس نحوه عملکرد نرم‌افزار در هر دو مرحله آدرس‌یابی (یعنی تجزیه متن آدرس و انطباق مکانی اجزا) با استفاده از چند نمونه آدرس توصیفی، بیان شده است. در انتهای این بخش، نتایج عملی به دست آمده، تفسیر و ارزیابی شده‌اند. در بخش هفتم، یافته‌های حاصل از تحقیق، جمع‌بندی شده است.

۲- آدرس‌یابی: تعاریف اولیه و مفاهیم پایه

آدرس را می‌توان به عنوان گزاره متنی توصیفی حاوی اطلاعات مکانی تعریف نمود که به موقعیتی منحصر به فرد ارجاع می‌دهد. آدرس‌ها به طور معمول در قالب سیستم‌های آدرس‌دهی و به صورت ترکیبی از اجزا همچون عوارض مکانی، روابط مکانی، کدهای پستی و غیره بیان می‌شوند [۸].

آدرس‌ها در دنیا به اعتبار پیروی یا عدم پیروی از یک ساختار استاندارد، به دو دسته آدرس‌های استاندارد و آدرس‌های غیراستاندارد یا توصیفی تقسیم می‌شوند. توضیح آن که کشورهای توسعه یافته مانند ایالات متحده آمریکا، انگلستان، استرالیا و غیره و یا حتی دستگاه‌های اجرایی گوناگون برای این که اجزای آدرس‌ها را برای ماشین قابل شناسایی نمایند، اقدام به تدوین استانداردهای آدرس‌دهی نموده‌اند.

با این وجود در بسیاری از کشورهای در حال توسعه مانند برزیل، هند و ایران، استاندارد برای آدرس‌دهی تعریف نشده و یا استفاده از استانداردهای تعریف شده، در کاربردهای روزانه، معمول نیست. بنابراین آدرس موقعیت‌ها معمولاً به صورت توصیف مسیر [۶، ۹-۱۲] یا توصیف مقصد [۱۳-۱۵] ارائه می‌شوند. البته حتی در کشورهایی که استانداردهای آدرس‌دهی وجود دارد، ممکن است افراد جامعه به منظور راهنمایی برای رسیدن به مقصد، از ساختارهای توصیفی آدرس‌دهی نیز استفاده کنند. به چنین آدرس‌هایی که ماهیت توصیفی داشته و از ساختار استاندارد پیروی نمی‌کنند، آدرس‌های توصیفی^۱ گفته می‌شود.

^۲ Landmark(s)

^۱ Descriptive address(es)

معابر نمی‌توان آدرس تجزیه شده را تعیین موقعیت کرد بلکه لازم است اجزا بر اساس ترکیبی که در آدرس قرار گرفته‌اند مورد جستجو قرار بگیرند تا امکان تعیین موقعیت آن‌ها فراهم آید. در این روش، آدرس بر اساس دانش مکانی آدرس‌گیرنده و ادراک مکانی آدرس‌دهنده از محیط، در قالب توصیف مسیر دسترسی به موقعیت مورد نظر از یک نقطه آشنا یا قابل دسترس برای آدرس‌گیرنده در چارچوب ساختارهای زبان طبیعی، ارائه می‌شود. با این شرایط، هر مؤلفه‌ای که به آدرس‌گیرنده برای دسترسی به موقعیت کمک می‌کند، امکان دارد در آدرس به کار رود. مثال بارز چنین وضعیتی، نحوه آدرس‌دهی در ایران می‌باشد. به عنوان نمونه برای نقطه مشخص شده در شکل ۱ آدرس‌های زیر را می‌توان ارائه داد:

- مسیر ۱: تهران، خیابان شریعتی، قلهک، خیابان پابرجا، بلوار آینه، سمت غربی کوچه گل یخ پلاک ۲ واحد ۹
- مسیر ۲: تهران، دروس، بلوار شهرزاد، خیابان پابرجا، بلوار آینه، سمت غربی کوچه گل یخ پلاک ۲ واحد ۹



شکل ۱- مسیرهای گوناگون برای دستیابی به یک نقطه براساس آدرس‌های توصیفی متفاوت

گرچه چنین روش آدرس‌دهی در ابتدای امر، مخصوصاً برای تعیین موقعیت ماشینی، ناکارآمد به نظر می‌رسد اما کارایی آن هم قابل توجه می‌باشد: اولاً این روش آدرس‌دهی نه تنها به موقعیت ارجاع داده بلکه نحوه دستیابی به آن را نیز بیان می‌کند. به بیان دیگر، در مقایسه با روش‌های کاملاً ساختاریافته، نیاز کاربر به ابزارهای کمک ناوبری در چنین شرایطی کاهش می‌یابد. ثانیاً این روش باعث توجه بیشتر به محیط شده و به دنبال آن موجب شکل‌گیری نقشه ادراکی افراد می‌شود که این امر به نوبه خود منجر به ارتقای سطح مهارت‌های مکانی خواهد شد.

در صورتی که تحقیقات انجام شده بیانگر آن است که وابستگی زیاد به ابزارهای کمک ناوبری (که در مورد آدرس‌های ساختاریافته رخ می‌دهد)، این مهارت‌ها را به شکل محسوسی کاهش داده است [۱۶].

بر اساس توضیحات فوق، مزایا و معایب هر کدام از این سیستم آدرس‌دهی را می‌توان این گونه جمع‌بندی نمود که آدرس‌های استاندارد به راحتی برای ماشین قابل تعریف هستند از این رو امکان تجزیه ماشینی آن‌ها به راحتی فراهم است در حالی که آدرس‌های غیراستاندارد یا توصیفی به خاطر پیروی از قالب‌های زبان‌های طبیعی، برای تجزیه ماشینی با چالش مواجه هستند. اما در طرف مقابل، آدرس‌های استاندارد به دلیل فقدان ارائه اطلاعات کافی درباره روابط مکانی اجزای آدرس و وابستگی شدید به ابزارهای کمک ناوبری، معمولاً کمکی به ارتقای دانش مکانی افراد نمی‌کنند در صورتی که آدرس‌های غیراستاندارد با توجه به ماهیت توصیفی که دارند، در ارتقای توانمندی‌های مکانی افراد و شکل‌گیری نقشه ادراکی آن‌ها از محیط مؤثر هستند. به بیان دیگر این گونه آدرس‌ها نه تنها فرایند راه‌یابی را تسهیل می‌کنند بلکه به افزایش دانش مکانی افراد هم کمک می‌کنند لذا حتی برای افراد ناآشنا با محیط هم قابل درک و استفاده هستند. در صورتی که آدرس‌های استاندارد بدون بهره‌مندی افراد از دانش قبلی از ساختار شهر قابل درک و استفاده نیستند و حتی در صورت آشنایی افراد با محیط، باز هم اطلاعات کافی در اختیار ایشان قرار نمی‌دهند از این رو به ابزارهای کمکی برای تعیین موقعیت وابسته هستند.

معمولاً آدرس‌یابی شامل دو مرحله اصلی تجزیه^۱ متن آدرس و انطباق مکانی^۲ اجزا می‌باشد [۱۷]. طرحواره مراحل کلی این فرایند در شکل ۲ نشان داده شده است. اگر چه آدرس‌یابی، ظاهراً یک مسئله سراسر است به نظر می‌رسد اما هنوز سئوالات اساسی درباره آن بی‌پاسخ مانده است. به عنوان نمونه برای یک موقعیت، آدرس‌های مختلفی می‌تواند وجود داشته باشد؛ یا ممکن است آدرس‌ها در قالب ساختارهای گوناگون بیان شوند. چنین مشکلاتی را می‌توان با طراحی یک تجزیه‌گر جامع، تا حد

^۱ Parsing

^۲ Spatial matching

زیادی کاهش داد. با این حال، طبق مطالعات صورت گرفته حتی در چینی حالتی، انطباق حدود ۳۰٪ از آدرس‌ها به دخالت کاربر نیاز داشته و به صورت کاملاً خودکار، تعیین موقعیت نمی‌شوند [۱۷].



شکل ۲- طرحواره مراحل اصلی آدرس‌یابی

پیچیدگی این مراحل، برای آدرس‌های توصیفی (غیراستاندارد) به علت ساختارنیافته بودن، بیشتر می‌باشد. حتی Davis و همکاران که به دنبال ارائه روشی برای آدرس‌یابی در کشور برزیل با آدرس‌های غیرساختاریافته بودند نیز صرفاً آدرس‌های غیراستاندارد مستقیم مطلق را مورد مطالعه قرار داده و درباره آدرس‌های غیراستانداردی که شامل روابط مکانی نیز بوده، سکوت کرده‌اند [۱۸]. برای توضیح بیشتر، آن‌ها آدرس‌ها را به دو دسته کلی مستقیم و غیرمستقیم تقسیم کرده‌اند. آدرس‌های مستقیم آدرس‌هایی است که توصیف ساختاریافته‌ای از موقعیت را ارائه می‌دهند، در حالی که آدرس‌های غیرمستقیم، از کدهای از پیش تعیین شده برای موقعیت استفاده می‌کنند. آدرس‌های مستقیم نیز حسب این که شامل روابط مکانی نیز باشند یا خیر، به دو دسته مطلق (بدون رابطه مکانی) و نسبی (با رابطه مکانی) تقسیم می‌شوند. به تعبیری نقطه انتهایی تحقیق مذکور، نقطه آغازین تحقیق حاضر محسوب می‌شود، بدین معنا که تحقیق حاضر به دنبال ارائه راه‌حلی برای تعیین موقعیت آدرس‌های غیرساختاریافته که شامل اطلاعاتی در قالب گروه‌های مکانی (که شامل روابط مکانی نیز هستند) می‌باشد.

۳- تحقیقات مرتبط پیشین

در این بخش، تحقیقات صورت گرفته پیشین در حوزه مرتبط با موضوع تحقیق یعنی «آدرس‌یابی و کدگذاری مکانی» مورد بررسی قرار گرفته‌اند. جمع‌بندی مطالعات صورت گرفته در زمینه تحقیقات مرتبط پیشین نشان می‌دهد؛ اولاً موضوع آدرس‌یابی و کدگذاری مکانی برای آدرس‌های غیراستاندارد، کم‌تر مورد توجه قرار گرفته است. ثانیاً ادراک مکانی و زبان طبیعی نقش کلیدی در توصیفات مکانی ایفا می‌کنند و به تعبیری این دو مفهوم، زیربنای توصیفات مکانی محسوب می‌شوند. ثالثاً با توجه

به میزان اطلاعات موجود در آدرس‌های توصیفی و بر اساس مفاهیم ارتباطات مکانی، لازم است نحوه انطباق اجزای آن‌ها، برای تعامل با ماشین، بهینه شود.

تا کنون غالباً تحقیقات درباره آدرس‌ها با ساختار استاندارد انجام شده است. در آدرس‌یابی ماشینی^۱ یا کدگذاری مکانی، آدرس یا مجموعه آدرس‌ها به عنوان ورودی دریافت می‌شود و سپس با یک الگوریتم مشخص، در یک پایگاه داده مرجع برای تطبیق آدرس با مکان مورد جستجو قرار می‌گیرد [۱۹]. به بیان دیگر، آدرس‌یابی ماشینی، همانند تحلیل‌های مکانی دیگر، از سه بخش ورودی، فرایند و خروجی تشکیل شده است. هر یک از بخش‌های آدرس‌یابی در گذر زمان دچار تحولات زیادی شده‌است [۲۰-۲۵]. هدف اولیه برای آدرس‌یابی در دهه ۱۹۶۰، صرفاً انتساب یک کد مکانی به آدرس‌های شهری بود. از این رو O'Reagan از داده‌های پستی ساده برای آدرس‌یابی استفاده نموده است [۲۶]. در حالیکه Hutchinson و همکاران با فرض این که هنوز آدرس‌های زیادی وجود دارد که با روش‌های موجود قابل تعیین موقعیت نیستند، پیشنهاد نمودند تا از عامل‌ها برای نمایش اطلاعات موجود در آدرس‌ها استفاده شود

خروجی‌ها نیز از کدهای مکانی ساده [۲۶-۲۸] به عوارض مکانی سه بعدی تغییر یافته است [۲۹]. در نهایت، فرایند آدرس‌یابی نیز از انطباق ساده عوارض [۲۶] به الگوریتم‌های پیچیده متحول شده است [۲۳].

با بررسی تحقیقات صورت گرفته در حوزه GIS برای کدگذاری مکانی یا آدرس‌یابی ماشینی ملاحظه می‌شود، غالب این تحقیقات، برای آدرس‌های استاندارد صورت گرفته است. به بیان دیگر، نکته قابل توجه درباره تحقیقات صورت گرفته در حوزه کدگذاری مکانی و آدرس‌یابی، این است که این تحقیقات، غالباً امکان تجزیه و فهم اجزای آدرس طبق یک ساختار مشخص را، پیش‌فرض گرفته‌اند، به نحوی که غالب تحقیقاتی که به دنبال ارزیابی و یا بهبود دقت بوده‌اند، با پیش‌فرض امکان تجزیه، تفسیر و انطباق ماشینی آدرس‌ها بر اساس یک ساختار مشخص، پی‌گرفته شده‌اند. به نحوی که، غالب تحقیقاتی که به دنبال ارزیابی [۲۲، ۲۷، ۲۸، ۳۰-۳۴] و یا بهبود دقت [۱۷، ۲۵، ۳۵-۳۸] بوده‌اند، با پیش‌فرض امکان تجزیه، تفسیر و انطباق ماشینی آدرس‌ها براساس یک ساختار مشخص، پی‌گرفته شده‌اند.

^۱ Machine-based addressing

۴- مبانی نظری تحقیق

در این بخش به بررسی ادراک مکانی و پردازش زبان طبیعی به عنوان مبانی مورد نیاز برای توسعه روش پیشنهادی تحقیق به منظور فراهم نمودن زمینه تعیین موقعیت ماشینی آدرس‌های توصیفی می‌پردازیم.

۴-۱- ادراک مکانی

ادراک مکانی انسانی، بخشی از یک محدوده تحقیقاتی وسیع و بین‌رشته‌ای^۱ علوم ادراکی می‌باشد. محققان زیادی از رشته‌هایی مانند روانشناسی، زبانشناسی، انسان‌شناسی، فلسفه و علوم رایانه درباره ذهن، استدلال، تجربه و مفهوم-سازی مردم از دنیایی که در آن زندگی کرده، تحقیق می-کنند [۶، ۱۵، ۳۹-۴۸]. به طور مشخص، علوم ادراکی با مطالعه هوش انسانی در تمام اشکال آن از شناخت^۲ و عمل^۳ تا زبان و استدلال^۴، سروکار دارد [۴۳]. اهمیت ادراک مکانی برای مدلسازی و پیش‌بینی رفتار مردم در مکان، توسط محققان متعددی مورد تاکید قرار گرفته است [۶، ۴۹-۵۲].

آدرس‌یابی همانند مسیریابی، در فضاهای بزرگ مقیاس صورت می‌گیرد [۵۳، ۵۴]. چنین فضاهایی قابل ادراک با منظر منفرد^۵ نبوده و عامل باید از طریق حرکت در محیط، فضا را درک کند. Lynch در سال ۱۹۶۰ برای نقشه‌های ادراکی پنج مؤلفه پیشنهاد داده [۵۵] که می-توان در این مرحله از این مؤلفه‌ها به عنوان مؤلفه‌های ادراک مکانی مؤثر بر آدرس‌یابی استفاده کرد. نقشه ادراکی تصویری در ذهن از محیط و ارتباطات مکانی است. این نقشه بستر فعالیت انسان بوده و برای ساماندهی فعالیت‌های انسان کاربرد دارد. این پنج مؤلفه به شرح زیر می-باشند:

۱. معبر^۶: محلی برای گذر که خیلی عادی و ساده راهبر در آن حرکت می‌کند.
۲. مرز^۷: حد فاصل بین دو مؤلفه مانند مرز دو محله می‌باشد.

۳. محله^۸: یک منطقه دو بعدی در وسعت بزرگ یا متوسط درون یک شهر بوده که در آن برخی ویژگی-های خاص تکرار می‌شود.

۴. رأس^۹: نقاطی پر کاربرد هستند که راهبر به صورت ذهنی وارد آنها می‌شود. این نقاط اهمیت بالایی در تصمیم‌گیری مسیر دارند.

۵. نشانه زمینی: یک عارضه نقطه‌ای پر کاربرد بوده که بر خلاف گره، فرد وارد آن نمی‌شود. این عارضه برای توجیه کردن و به دست آوردن موقعیت و انتخاب مسیر کاربرد دارد.

این مؤلفه‌ها در مدل پیشنهادی Kuiper بهبود داده شده‌اند [۵۳]. مدل مذکور می‌تواند مبنایی برای ارائه راه حل آدرس‌یابی در این تحقیق باشد. تمرکز اصلی توجه این مدل نقشه ادراکی بوده که اجزای آن شامل پنج مؤلفه زیر می‌باشند:

۱. مسیر^{۱۰}: شامل زنجیره‌ای از اقدامات که صورت می-پذیرد تا شخص از یک نقطه به نقطه دیگر جابجا شود.
 ۲. ساختار توپولوژیک شبکه خیابان‌ها^{۱۱}: توصیف خیابان‌ها و اماکن که شامل دانش جزئی از ترتیب اماکن یک خیابان و هندسه تقاطع دو خیابان می‌باشد.
 ۳. موقعیت نسبی دو مکان^{۱۲}: برداری در یک سیستم مختصات فرضی است.
 ۴. مرزهای جداکننده^{۱۳}: مناطق را تعریف می‌کنند.
 ۵. مناطق^{۱۴}: بیان‌گر رابطه شمول بوده که سطحی از انتزاعی‌سازی ارتباطات را نمایش می‌دهد.
- در بخش پنجم (رویکرد پیشنهادی تحقیق)، نحوه به کارگیری این مؤلفه‌ها در تعریف گرامر و در تعریف مدل داده بیان خواهد شد.

۴-۲- پردازش زبان طبیعی

پردازش زبان طبیعی یکی از زیربناهای نظری حل مسئله تحقیق حاضر می‌باشد. از این رو در این بخش، به تعریف و بیان اقسام این پردازش، پرداخته شده و بعد از بیان برخی کارکردهای مهم آن، درباره موضوع «تجزیه» به

^۸ District

^۹ Node

^{۱۰} Route

^{۱۱} Topological structure of a street network

^{۱۲} Relative position of two places

^{۱۳} Dividing boundaries

^{۱۴} Regions

^۱ Interdisciplinary

^۲ Perception

^۳ Action

^۴ Reasoning

^۵ Single viewpoint

^۶ Path

^۷ Edge

گرفته هم در حوزه هوش مصنوعی و محاسبات نرم و هم در حوزه زبان‌شناسی محاسباتی، کارکردهای پردازش زبان طبیعی هم در عناوین و هم در راه‌حل‌ها توسعه یافته است از این رو امروزه، اغلب تحقیقات با استفاده از الگوریتم‌های تحت نظارت یا نیمه تحت نظارت^۵، به شناسایی و بهره‌گیری از ساختارهای زبان طبیعی پرداخته‌اند [۶۸-۷۰].

تجزیه متن آدرس، یکی از مراحل اصلی در آدرس‌یابی و کدگذاری مکانی چه در حالت استاندارد و چه در حالت توصیفی می‌باشد. در حالت استاندارد، با توجه به ساختار دقیق آدرس‌ها، تجزیه آن‌ها برای ماشین، دشواری خاصی ندارد. اما در حالت آدرس‌های توصیفی، از آن‌جا که موقعیت مورد اشاره در آدرس در قالب ساختارهای زبان طبیعی توصیف می‌شود، تجزیه متن آدرس، با چالش‌هایی مواجه می‌شود. در واقع تجزیه آدرس‌های توصیفی، حالت خاصی از تجزیه در پردازش زبان‌های طبیعی به حساب می‌آید. از این رو می‌توان گفت تجزیه، عمده‌ترین کارکرد پردازش زبان طبیعی در آدرس‌یابی توصیفی می‌باشد. تجزیه یا تحلیل دستوری^۶، فرایند پردازش یک جمله (زبان طبیعی) یا رشته‌ای از علائم (زبان ماشین) بر اساس تطبیق آن با مجموعه‌ای از قواعد یا گرامر برای تعیین قالب، کارکرد و ارتباطات نحوی اجزای آن می‌باشد [۷۱]. به بیان دیگر، منظور از تجزیه یا فرایند ساختاردهی یک ارائه خطی^۷ از اطلاعات با توجه به گرامر مشخص است [۷۲]. این تعریف یک تعریف کاملاً انتزاعی می‌باشد. ارائه خطی می‌تواند یک جمله، یک برنامه کامپیوتری، یک قطعه صوتی یا هر رشته خطیباشد که در آن، مؤلفه قبلی، مؤلفه بعدی را تحدید کند. فرایند تجزیه جمله، عکس فرایند تشکیل جمله می‌باشد [۷۳]. این مفهوم در حوزه‌هایی از قبیل زبان‌شناسی، زبان‌شناسی محاسباتی و علوم رایانه با همین معنای کلی به کاربرده می‌شود [۷۴]. همان‌طور که در تعریف تجزیه اشاره شد، برای انجام تجزیه لازم است ابتدا گرامر مورد نظر به دقت تعریف شود.

نکته قابل توجه این است که غالب روش‌های موجود تجزیه ماشینی، بر یک دامنه مشخص از داده‌های آموزشی استوار هستند [۷۵-۷۷]. هر زبان یک مجموعه نامتناهی از ترکیبات نحوی است از این رو برای تعریف آن، از گرامر به

عنوان کارکرد اصلی پردازش زبان طبیعی در آدرس‌یابی توصیفی، توضیحات بیشتری ارائه شده است. زبان، ابزار تفکر است [۵۶] و زبان طبیعی همان چیزی است که انسان‌ها در محاورات روزانه خود، در قالب آن با یکدیگر صحبت می‌کنند [۵۷]. پردازش زبان طبیعی به عنوان حوزه مشترکی میان علوم رایانه، هوش مصنوعی و زبان‌شناسی به مطالعه استفاده از ماشین برای پردازش زبان گفتاری یا نوشتاری می‌پردازد [۵۸]. اهمیت این موضوع در آن است که همواره اطلاعات زیادی در قالب ساختارهای زبان طبیعی وجود دارد که می‌تواند به وسیله ماشین استخراج شود [۵۷]، مانند آن‌چه در آدرس‌های توصیفی رخ می‌دهد.

از طرفی، پردازش زبان طبیعی به دو حوزه کلی «فهم زبان طبیعی»^۱ و «تولید زبان طبیعی»^۲ تقسیم می‌شود [۵۹]. در حالت «فهم»، زبان طبیعی به عنوان ورودی و در حالت «تولید»، زبان طبیعی به عنوان خروجی ماشین محسوب می‌شود. در حالت اول، هدف اصلی، قابل فهم کردن معانی زبان طبیعی استفاده شده به وسیله انسان برای ماشین می‌باشد در حالیکه، در حالت دوم هدف تولید اطلاعات در قالب زبان طبیعی توسط ماشین و برای استفاده انسان است. درباره آدرس‌های توصیفی می‌توان گفت مسئله تعیین موقعیت این آدرس‌ها که موضوع تحقیق حاضر می‌باشد، نیاز به فهم زبان طبیعی دارد اما تولید آدرس توصیفی مناسب یک مکان برای مخاطب معلوم، که خارج از دامنه این تحقیق می‌باشد، مربوط به تولید زبان طبیعی می‌باشد.

یکی از مهم‌ترین انقلاب‌ها در زبان‌شناسی و مخصوصاً زبان‌شناسی محاسباتی، «گرامر جهانی» بود که Chomsky آن را پیشنهاد داد [۶۰]. با این حال، غالب اقدامات عملی صورت گرفته در این زمینه تا اواخر دهه ۸۰ میلادی، به موفقیت اساسی دست نیافتند [۶۱]. اکثر تحقیقات انجام شده در آن بازه زمانی به تدوین هستی‌شناسی‌های مفهومی جهت قابل فهم کردن داده‌ها برای ماشین‌ها پرداخته بودند [۶۲-۶۶].

بعد از دهه ۸۰ میلادی با توسعه الگوریتم‌های یادگیری ماشینی^۳، برای مسئله پردازش زبان طبیعی، راه‌حل‌های نرم^۴ پیشنهاد شد که انقلاب بعدی در این حوزه محسوب می‌شود [۶۷]. امروزه نیز با توجه به پیشرفت‌های صورت

^۵ Supervised / semi-supervised algorithms

^۶ Syntactical analysis

^۷ Linear representation

^۱ Natural language understanding

^۲ Natural language generation

^۳ Machine learning

^۴ Soft solutions

عنوان یک مجموعه متناهی استفاده می‌شود. در واقع گرامر، زبان را توصیف می‌کند. این موضوع، اهمیت جایگاه گرامر را در پردازش زبان طبیعی و مخصوصاً در تجزیه نشان می‌دهد. نمونه‌هایی از تعریف گرامر برای آدرس‌های توصیفی و تجزیه متن آدرس بر مبنای آن‌ها در بخش پنجم ارائه شده است.

۵- رویکرد پیشنهادی تحقیق

در این بخش، به رویکرد پیشنهادی تحقیق برای هر دو مرحله اصلی تعیین موقعیت آدرس‌های توصیفی یعنی «تجزیه متن آدرس» و «انطباق مکانی اجزا» پرداخته شده است. بر اساس راه حل پیشنهادی، مرحله اول با «ایجاد هستی‌شناسی گرامر رسمی زبان آدرس‌های توصیفی» و مرحله دوم با «طراحی و توسعه یک مدل داده مکانی غنی شده با ادراک مکانی» حل می‌شود. هر کدام از این دو اقدام، در این بخش به تفصیل بررسی و نحوه پیاده‌سازی آن‌ها در بخش بعد، ارائه شده است.

۵-۱- ایجاد هستی‌شناسی گرامر رسمی زبان آدرس‌های توصیفی

همانطور که توضیح داده شد، مرحله اول راه حل پیشنهادی، ارائه تعریف رسمی از گرامر آدرس‌های توصیفی می‌باشد. به دلیل گستردگی و پیچیدگی زیاد ساختارهای زبان طبیعی، داده‌های مرتبط با زبان صرفاً برای دامنه کاربردی^۱ مشخص و با استفاده از استقرای ناقص قریب تام جمع‌آوری شده است.

یک زبان رسمی، از مجموعه‌ای از نشانه‌ها (معادل کلمات در زبان طبیعی) و مجموعه‌ای از قواعد برای ترکیب این نشانه‌ها (معادل گرامر در زبان طبیعی) تشکیل یافته است. نشانه‌ها به سه دسته تقسیم می‌شوند:

- نشانه‌های انتهایی که ساده‌تر نمی‌شوند. این دسته نشانه‌ها می‌توانند ثابت یا متغیر باشند.
- نشانه‌های غیرانتهایی که باید به نشانه‌های انتها ساده شوند.
- نوع خاصی از نشانه‌های غیرانتهایی که نشانه‌های آغازین نامیده می‌شود.

منظور از گرامر، مجموعه از قواعد برای ترکیب نشانه‌ها (لغات) می‌باشد. اگر چه زبان طبیعی از چنین ساختار استواری پیروی نمی‌کند، با این حال می‌توان به سطحی از قاعده‌مندی^۲ در آن دست یافت و بالتبع امکان تجزیه ترکیبات ارائه شده در آن با استفاده از روش‌های موجود در زبان رسمی را فراهم آورد [۷۸]. از این رو می‌توان با تعریف یک گرامر، تجزیه آدرس توصیفی را به اجزای آن، ممکن نمود.

قواعد ترکیب زبان‌های رسمی، معمولاً در قالب Backus-Naur Form (BNF) بیان می‌شوند. BNF در واقع یک فرازبان برای بیان قواعد زبان‌های دیگر می‌باشد [۷۹].

بر اساس تعریفی ارائه شده، در آدرس‌های توصیفی باید به دنبال اولاً نشانه‌ها (اعم از انتهایی و غیرانتهایی) و ثانیاً قواعد ترکیب آن‌ها باشیم. می‌توان دو دسته علائم انتهایی را به شرح زیر در غالب آدرس‌های توصیفی شناسایی کرد:

۱- مکان نام^۳ (GN): شامل دوبخش «مکان نام ثابت» و «مکان نام متغیر» می‌باشد.

۱-۱- مکان نام ثابت (cGN): مانند الفاظ «خیابان»، «کوچه»، «پلاک» و ...

۲-۱- مکان نام متغیر (vGN): مانند نام محله یا نام کوچه و شماره پلاک و ...

۲- روابط مکانی (SR): مانند بالاتر از یا نرسیده به و ...
براین اساس، می‌توان قواعدی را برای ترکیب این علائم به شرح زیر در قالب BNF، ارائه نمود:

- S := (SG)
- SG := [SR] GN
- GN := cGN vGN | vGN cGN
- cGN := "avenue" | "ave." | "street" | "alley" | "number" | "unit" | ...
- SR := "after" | "before" | "in front of" | "left of" | "right of" | ...

قاعده اول بیان می‌کند که یک آدرس، شامل تعدادی «گروه مکانی» می‌باشد. قاعده دوم، ساختار این گروه‌های مکانی را نشان می‌دهد. طبق این قاعده، هر گروه مکانی حتماً شامل یک مکان نام و احتمالاً شامل یک رابطه مکانی می‌باشد. بر اساس قاعده سوم، هر مکان نام ترکیبی از مکان نام

^۲ Regularity
^۳ Geo-name

^۱ Corpus

شرایط عدم وجود استاندارد پرداخته نشده و راهبرد غالب در تحقیقات و ابزارهای کدگذاری مکانی موجود نیز محدود به آدرس‌های استاندارد (مانند ساختارهای آدرس آمریکایی و اروپایی) می‌باشد [۱۸]. حتی در تحقیقاتی هم که به موضوع آدرس‌یابی در شرایط غیراستاندارد پرداخته شده، آدرس‌یابی در شرایط مطلق و بدون در نظر گرفتن رابطه مکانی مطالعه شده است [۱۸]. از این رو لازم است برای حل مسئله آدرس‌یابی برای آدرس‌های توصیفی که ساختار مشخصی نداشته، بر مبنای مؤلفه‌های ادراک مکانی بوده و شامل روابط مکانی نیز می‌باشند، مدل داده جدیدی با در نظر گرفتن نیازمندی‌های مختص به این حالت، تعریف شود. در این بخش، مدل داده مورد نظر برای تعیین موقعیت آدرس‌های توصیفی تعریف می‌شود. یک مدل داده، ترکیبی از حداقل سه جزء به شرح زیر می‌باشد [۸۰]:

۱. مجموعه‌ای از انواع ساختار داده^۳
 ۲. مجموعه‌ای از عملگرها یا قواعد استنتاج
 ۳. مجموعه‌ای از قواعد جامعیت عمومی^۴
- یکی از نکات مهم در تعریف مدل داده، این است که ویژگی‌هایی نگه داشته شوند که مابقی ویژگی‌ها از آن‌ها قابل استخراج باشد. روابطی که تغییر نمی‌کنند از این جهت، مناسب‌ترین روابط هستند [۷۹].
- برای توسعه مدل داده پیشنهادی، لازم است اقداماتی به شرح زیر صورت پذیرد:
- تعریف روابط مکانی موجود در آدرس‌های توصیفی، برای مؤلفه‌های ادراک مکانی مؤثر بر آدرس‌یابی
 - شناسایی روابطی که قابل استخراج از روابط دیگر هستند و چگونگی این ارتباط
 - شناسایی مابقی روابطی که قابل استخراج از روابط دیگر نیستند

با توجه به جایگاه ادراک مکانی در آدرس‌های توصیفی که در بخش چهارم به آن پرداخته شد، در این مرحله لازم است به دنبال تعیین مؤلفه‌های ادراک مکانی موجود در آدرس‌های توصیفی باشیم. همانطور که پیش از این توضیح داده شد، Lynch و درپی آن، Kuiper به تعیین مؤلفه‌های نقشه‌های ادراکی پرداخته‌اند [۵۳، ۵۵]. در

ثابت و متغیر می‌باشد. قواعد چهارم و پنجم نیز مکان‌نام‌های ثابت و روابط مکانی را برمی‌شمرد. قواعد چهارم و پنجم در واقع لیست‌های شمارش^۱ هستند. باید توجه کرد این قبیل اطلاعات از سنخ داده‌های مرتبط با زبان^۲ می‌باشد. این گونه اطلاعات صرفاً از طریق استقرا و جمع‌سپاری در یک دامنه کاربردی مشخص، قابل جمع‌آوری هستند. به بیان دیگر به راحتی نمی‌توان ادعای کامل بودن این لیست‌ها را داشت همان‌طور که هیچ فرهنگ لغتی ادعای کامل بودن ندارد! نکته دیگر آن که در این گرامر تعریف شده، مرزهای گروه-های مکانی در آدرس، مستقل از علائم سجاوندی، تعیین می‌شود. لیست مربوط به قاعده چهارم شامل «میدان/م، چهارراه/۴راه، سهراب/۳راه، فلکه/ف، بزرگراه/اتوبان/ا، بلوار/بولوار/ل، خیابان/خ، جاده/ج، کوچه/ک/فرعی، بن‌بست/بن‌بست/بن، آزادراه/ازادراه، آزادراه، جاده‌خاکی، کنارگذر/کندرو، تونل، لوپ/دوربرگردان، رمپ/ورودی/خروجی/بریدگی، پل/روگذر، زیرگذر/زیرگذر، محله، منطقه، محدوده» و لیست مربوط به قاعده پنجم شامل «دستچپ/سمتچپ، دست راست/سمت راست، بعداز/بعد/بالاتراز/ازسمت، قبل از/قبل/پایین‌تر از/به سمت/مانده/نرسیده به/نرسیده، ابتدای/ابتدا/اول/اوایل/سر، انتهای/انتها/آخر/اواخر/ته، به شمال/به سمت شمال/جنوب به شمال، به شمال‌شرق/به سمت شمال‌شرق، به شرق/به سمت شرق/غرب به شرق، به جنوب‌شرق/به سمت جنوب‌شرق، به جنوب/به سمت جنوب/شمال به جنوب، به جنوب‌غرب/به سمت جنوب‌غرب، به غرب/به سمت غرب/شرق به غرب، به شمال‌غرب/به سمت شمال‌غرب، نبش/کنار/جنب، حوالی/نزدیک، اواسط/وسط، تقاطع، پشت $\neq$ روبرو/مقابل/جلوی» می‌باشد.

خروجی این تجزیه، تعیین گروه‌های مکانی موجود در آدرس، نوع این گروه‌ها و ترتیب وقوع آن‌ها در متن آدرس می‌باشد.

۵-۲- طراحی و توسعه یک مدل داده مکانی غنی شده با ادراک مکانی

همانطور که پیشتر ذکر شد، غالباً در تحقیقات حوزه کدگذاری مکانی و آدرس‌یابی، به موضوع آدرس‌یابی در

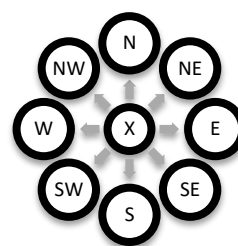
^۳ Data structure types
^۴ General integrity rules

^۱ Enumeration lists
^۲ Language-related data

تحقیق حاضر با توجه به توضیحات ارائه شده و بر مبنای اجزای آدرس‌های توصیفی و مدل‌های مذکور، در مؤلفه-های ادراک مکانی آدرس‌های توصیفی به عنوان اجزای مدل داده مورد استفاده در مرحله انطباق مکانی اجزای آدرس، به شرح زیر تعریف می‌شوند:

۱. مختصات نشانه‌های زمینی و رؤوس
۲. رؤوس ابتدایی، میانی و پایانی هر مسیر
۳. مسیرهای تشکیل‌دهنده هر محله
۴. روابط مکانی مستقل

روابطی از قبیل «دست چپ، بالاتر از، نرسیده به، اواخر» وابسته به توجیه عامل هستند، لذا هر رابطه همراه با رابطه مزدوج خود مانند «دست راست، پایین‌تر از، بعد از، اوایل»، نگهداری می‌شود. در مورد جهات اصلی و فرعی، نیازی نیست که به صورت صریح نگه داشته شوند، بلکه همانطور که در شکل ۳ نشان داده شده، در صورت نیاز می‌توان این دسته از روابط را با استفاده از توصیفگر X استخراج نمود. این توصیفگر می‌تواند این روابط را برای هر دو مؤلفه‌ای در نقشه استخراج کند. به این ترتیب که این توصیفگر با استفاده از مختصات دو مؤلفه مورد نظر پس از محاسبه ژیزمان امتداد بین دو مختصات، نوع جهت اصلی بین دو مؤلفه مذکور را تعیین می‌کند. همچنین این توصیفگر می‌تواند با استفاده از مختصات یک مؤلفه، محدوده‌ای را که در جهت مشخصی نسبت به مؤلفه مذکور قرار دارند، مشخص نماید. این توصیفگر در واقع مسئله را با استفاده از ژیزمان، حل می‌کند.



شکل ۳- طرحواره توصیفگر X

روابط «ابتدا و انتها» برای رأس و محله، این همانی^۱ هستند. رابطه «واسط»، نقیض رابطه «ابتدا و انتها» می‌باشد. اما برای مسیر در واقع شامل رؤوس ابتدایی و انتهایی مسیر می‌شوند. البته این که تا کجا «واسط» و تا کجا «اوایل» یا «اواخر» مسیر به حساب می‌آید، کاملاً وابسته به

بافت است، لذا اگر جواب نهایی منوط به این رابطه باشد، نمی‌توان به آن اعتماد کرد. رابطه «جنب»، این همانی است. رابطه «حوالی»، اجتماعی از روابط «چپ و راست»، «قبل و بعد»، «جنب» است. روابط «روبرو» و «پشت» هم همان رابطه «حوالی» تعریف می‌شود. تا اینجا، مجموعه ساختار داده و قواعد استنتاج مدل داده تعریف شده‌اند.

دو رویکرد برای جستجو در مدل داده برای انطباق مکانی اجزای آدرس، وجود دارد. در رویکرد اول، ترتیب جستجو به همان ترتیبی است که انسان طبق آن، آدرس را تعیین موقعیت می‌کند، یعنی از ابتدای آدرس، مرحله به مرحله بین تمامی اجزا به ترتیب جستجو انجام گیرد. این نوع جستجو تفاوتی بین مکان‌نام‌های ثابت و روابط مکانی قائل نیست. مشکل این رویکرد برای ماشین، امکان انباشت اطلاعات اضافی در متن آدرس و تفاسیر مشتبه از مفاهیم اجزای آدرس، می‌باشد. از آن‌جا که اساساً اطلاعات ارائه شده در یک آدرس توصیفی برای انسان بوده و غالباً شامل اطلاعات اضافی (که معمولاً مبتنی بر بافت نیز هستند) برای راهنمایی بهتر مخاطب انسانی می‌باشد، در اکثر موارد، ماشین به همه این اطلاعات، برای تعیین موقعیت آدرس نیاز ندارد، لذا این نحوه جستجو در مدل داده برای تعیین موقعیت ماشینی آدرس، مناسب به نظر نمی‌رسد.

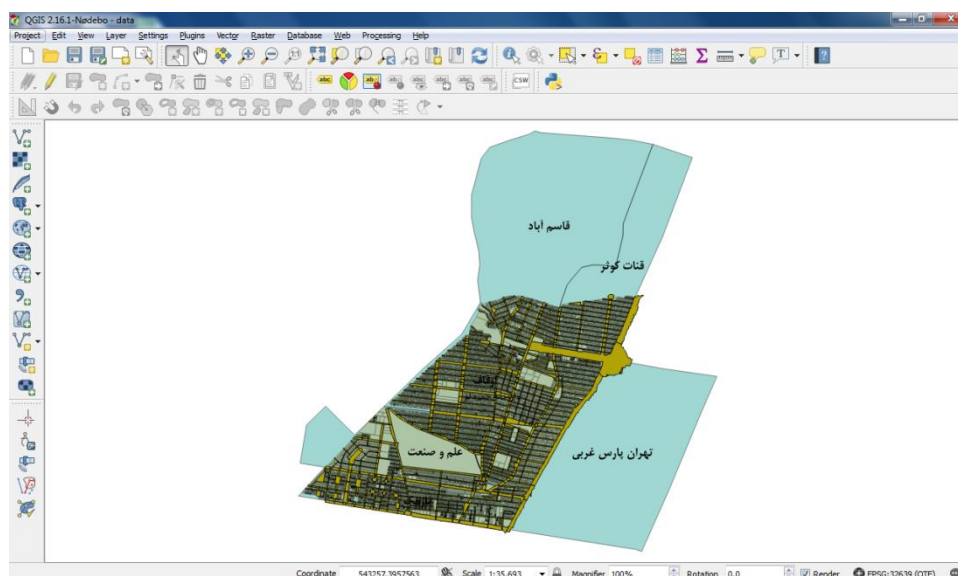
در رویکرد دوم، برخلاف آنچه عامل انسانی برای یافتن موقعیت آدرس انجام می‌دهد، جستجو از انتهای آدرس و در بدو امر صرفاً بین مکان‌نام‌های ثابت صورت می‌گیرد یعنی در ابتدا، از جستجوی روابط مکانی موجود در متن آدرس، صرف‌نظر می‌شود. در این حالت، به محض این که به موقعیت یکتا دست یافته شود، جستجو پایان می‌گیرد. مزیت اصلی این رویکرد، کاهش دو مشکل مذکور در حالت قبل می‌باشد. لذا امکان اشتباه ناشی از تفسیر مشتبه این روابط، کاهش می‌یابد و اساساً با این رویکرد، نقش روابط مکانی در تعیین موقعیت آدرس توصیفی، کمرنگ شده و بنابراین امکان وقوع اشتباه ناشی از ابتدای تفسیر این روابط بر بافت، کم می‌شود. به دلیل مزایای مذکور، در تحقیق حاضر، رویکرد دوم جهت جستجو در مدل داده برای انطباق مکانی اجزای آدرس، انتخاب شده است.

۶- پیاده‌سازی

در این بخش پس از ارائه توضیحاتی درباره محیط پیاده‌سازی منطقه مورد مطالعه، به نحوه پیاده‌سازی تجزیه ماشینی آدرس توصیفی بر مبنای گرامر رسمی تعریف شده و مدل داده مکانی غنی شده با ادراک مکانی، برای انطباق مکانی اجزای آدرس توصیفی، پرداخته شده است و در

^۱ Identity

محلّه با استفاده از معابر اطراف آن نیز معلوم فرض می-شود. محدوده داده به کار رفته در این مرحله، بخشی از منطقه ۴ شهرداری تهران می‌باشد که به دلیل موجود، آماده و قابل دسترس بودن، انتخاب شده است. موقعیت این محدوده در تهران، در شکل ۴ نشان داده شده است.



شکل ۴- موقعیت داده مورد استفاده در بین محله‌های شهر

نهایت نتایج عملی حاصل از تعیین موقعیت چند آدرس نمونه ارائه شده و مورد ارزیابی قرار گرفته است.

آنچه در مرحله انطباق مکانی اجزای آدرس، دانسته فرض می‌شود، یک نقشه GIS-ready بوده که نام معابر و اماکن در آن مشخص است. همچنین تعریف محدوده هر

شناسایی شده قبلی، به عنوان مقدار گروه مکانی قبلی یا همان مکان نام متغیر، در نظر گرفته می‌شود.

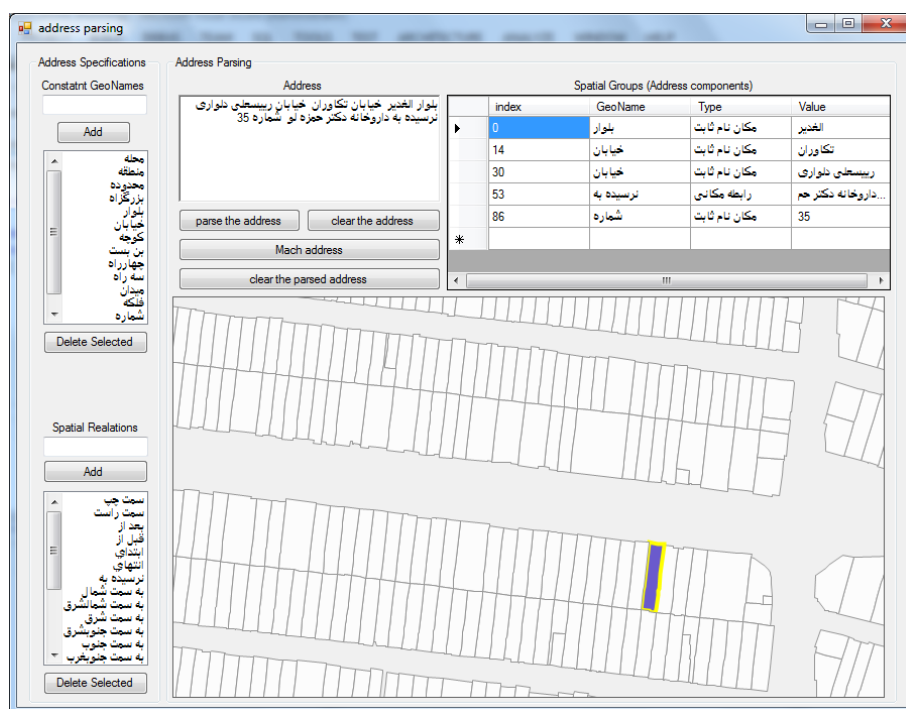
با توجه به نیاز مرحله انطباق مکانی، نوع هر گروه مکانی و ترتیب آن در آدرس و همچنین نوع مکان نام ثابت با توجه به مؤلفه‌های ادراک مکانی، در مرحله تجزیه آدرس تعیین می‌گردد. در این مرحله، پس از تعیین محتوا و ترتیب اجزای آدرس توصیفی در مرحله تجزیه متن آدرس، جستجو در مدل داده، صورت می‌گیرد تا پاسخ یکتا حاصل شود. همانطور که بیان شد از میان دو رویکرد ممکن برای ترتیب جستجو در میان اجزای آدرس، رویکرد دوم برای ماشین مناسب‌تر است. یعنی جستجو از انتهای آدرس و در ابتدا فقط بین مکان نام‌های ثابت انجام می‌پذیرد که موجب کاهش تأثیر عوامل خطا می‌شود.

نمایی از تعیین موقعیت ماشینی آدرس‌های توصیفی ۱ و ۲ در شکل‌های ۵ و ۶، نشان داده شده‌اند. ملاحظه می‌شود که در این شکل، آدرس توصیفی، ابتدا تجزیه شده و پس از تعیین اجزای آن و ترتیب و نوع هر کدام، تعیین موقعیت شده و موقعیت یافت شده در نقشه نشان داده شده است.

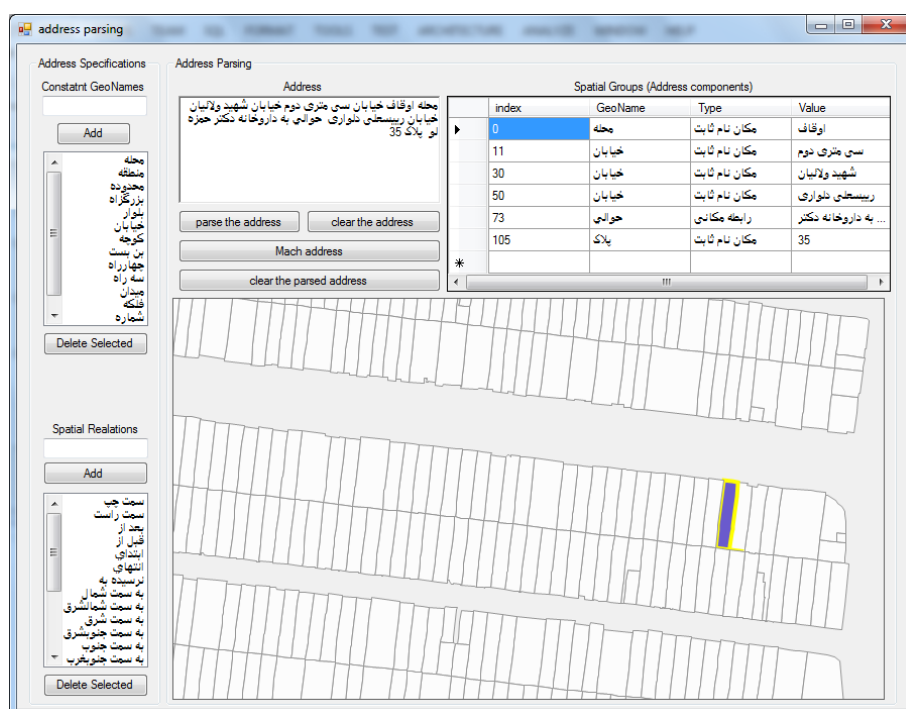
دو آدرس توصیفی ممکن مربوط به موقعیت موردنظر که بر مبنای دو مسیر متفاوت ایجاد شده، به شرح زیر می‌باشد:

- آدرس ۱: بلوار الغدير خیابان تکاوران خیابان رییسعلی دلواری نرسیده به داروخانه دکتر حمزه لو شماره ۳۵
- آدرس ۲: محله اوقاف، خیابان سی‌متری دوم، خیابان شهید ولایتیان، خیابان رئیسعلی دلواری، حوالی داروخانه دکتر حمزه‌لو، پلاک ۳۵

برای تعیین موقعیت ماشینی این آدرس‌ها لازم است ابتدا متن آن‌ها تجزیه شود. این آدرس‌ها بر اساس مکان نام‌های ثابت و روابط مکانی که از پیش به نرم‌افزار معرفی شده‌اند، تجزیه می‌شود. برای تجزیه ماشینی متن آدرس فوق، برنامه کاربردی بر اساس گرامر تعریف شده در بخش پنجم تهیه شده است. پیاده‌سازی صورت گرفته برای تجزیه متن آدرس، مستقل از علائم سجاوندی بین اجزای آدرس بوده و گروه‌های مکانی را صرفاً بر اساس مکان نام‌های ثابت و روابط مکانی، مشخص می‌کند. به این ترتیب که با شناسایی هر مکان نام ثابت یا رابطه مکانی، محتویات بین گروه مکانی شناسایی شده فعلی و گروه مکانی



شکل ۵- تعیین موقعیت ماشینی آدرس توصیفی ۱



شکل ۶- تعیین موقعیت ماشینی آدرس توصیفی ۲

این صورت آدرس یکتای به دست آمده، معتبر یا نامعتبر خواهد بود. در حالت دوم ممکن است هیچ مکانی برای آدرس، یافت نشود. در این حالت، یافت نشدن آدرس به اطلاع کاربر نرم افزار رسانده می‌شود. در حالت سوم نیز ممکن است بیش از یک مکان با توصیفات ارائه شده در آدرس منطبق باشد که تمام موقعیت‌های یافت شده به عنوان جواب، به کاربر ارائه می‌شود.

در این روش، سه حالت برای خروجی فرایند آدرس-یابی ممکن است. حالت اول این است که با جستجو در بین گروه‌های مکانی و از انتهای آدرس، مکان یکتا یافت شود. در این حالت ممکن است، جواب یافت‌شده معتبر یا نامعتبر باشد. چرا که سایر اجزای باقی مانده از آدرس که مورد جستجو برای تعیین موقعیت قرار نگرفته‌اند، ممکن است با مکان تعیین شده، سازگار یا ناسازگار باشند که در

برای حالت‌های اول و سوم، خروجی آدرس‌یابی، مختصات موقعیت یا موقعیت‌هایی می‌باشد که آدرس در پی بیان آن بوده است. جواب نهایی اعم از رأس، مسیر یا محله همراه با مختصات و به صورت نقشه‌ای نمایش داده می‌شود. همانطور که مشاهده شد، می‌توان با استفاده از راه حل ارائه شده، آدرس‌های توصیفی را که در قالب زبان طبیعی بیان شده‌اند تعیین موقعیت نمود. نکات زیر درباره ویژگی‌های خروجی این روش قابل ذکر می‌باشد:

از آنجا که جستجو از انتهای آدرس، انجام می‌گیرد، آدرس‌های توصیفی مختلف مربوط به یک موقعیت، به پاسخ واحد منجر می‌شوند.

از آنجا که به محض رسیدن به پاسخ یکتا، فرایند جستجو پایان می‌گیرد، لذا امکان تعیین موقعیت نشدن آدرس به دلیل انباشت اطلاعات اضافی که به جهت ناکاملی، در پایگاه داده وجود ندارد، کاهش می‌یابد.

چنانچه آدرسی اطلاعات کافی برای رسیدن به رأس یا نشانه زمینی را نداشته باشد، موقعیت نهایی در سطح ممکن، از قبیل قالب مسیر یا محله تعیین خواهد شد.

از آنجا که روابط مکانی، مبتنی بر بافت بوده و بسته به ادراک کاربر، امکان تفاسیر گوناگونی از آن‌ها وجود دارد، لذا در فرایند انطباق، جستجو در این روابط، فقط در صورتی انجام می‌گیرد که جستجو در مکان‌نام‌های ثابت به پاسخ یکتا نرسد. در غیر این صورت، از روابط مکانی صرفاً برای اعتبارسنجی موقعیت یافت شده، استفاده می‌شود.

دلیل اصلی در مواردی که این سیستم قادر به تعیین موقعیت آدرس نشده، عدم وجود مکان‌نام‌های متغیر آدرس در مدل داده می‌باشد. گاهی هم ترکیب به کار رفته در آدرس، ترکیب معتبری نیست یعنی علی‌رغم وجود تک تک گروه‌های مکانی در مدل داده، ترکیب آن‌ها در دنیای واقعی ممکن نیست. مثلاً به نام مسیری در یک محله اشاره شده در حالیکه در آن محله، چنین مسیری وجود ندارد.

۷- جمع‌بندی

همانطور که بیان شد، علی‌رغم دشواری تعامل ماشین با آدرس‌های توصیفی، این آدرس‌ها مزایایی نسبت به آدرس‌های استاندارد دارند که باعث می‌شود به جای آن که به دنبال جایگزینی آن‌ها با آدرس‌های استاندارد باشیم، ترجیح داده می‌شود، به توسعه روشی برای حل مسئله

تعیین موقعیت ماشینی آدرس‌های توصیفی بپردازیم. از جمله مزایای مذکور، این است که آدرس‌های توصیفی، علاوه بر ارجاع به موقعیت مورد نظر، نحوه دستیابی به آن را نیز تا حدی بیان می‌کنند از این رو در مقایسه با آدرس‌های استاندارد، نیاز به ابزارهای کمک ناوبری را کاهش می‌دهند. از طرفی این گونه آدرس‌ها، باعث توجه بیشتر شخص به محیط و در نتیجه، شکل‌گیری نقشه ادراکی و ارتقای سطح مهارت‌های مکانی کاربر می‌شوند. در صورتی که آدرس‌های استاندارد بدون بهره‌مندی افراد از دانش قبلی از ساختار شهر قابل درک و استفاده نیستند و حتی در صورت آشنایی افراد با محیط، باز هم اطلاعات کافی در اختیار ایشان قرار نمی‌دهند از این رو به ابزارهای کمکی برای تعیین موقعیت وابسته هستند.

از طرفی همان‌گونه که توضیح داده شد، ادراک مکانی و زبان طبیعی نقش مهمی در مطالعه آدرس‌ها، مخصوصاً آدرس‌های توصیفی دارد. بنابراین در این تحقیق سعی شد تا با دخیل نمودن مفاهیم ادراک مکانی انسانی و پردازش زبان‌های طبیعی در آدرس‌یابی ماشینی، پاسخی برای چگونگی تعیین موقعیت آدرس‌های توصیفی ارائه داده شود. در واقع در حالتی که استاندارد آدرس‌دهی وجود ندارد، آدرس‌های توصیفی در قالب ساختارهای زبان طبیعی بر مبنای ترکیب‌باز مؤلفه‌های ادراک مکانی انسانی بیان می‌شود. به بیان دیگر آدرس‌دهی، روشی برای تبادل ادراک مکانی در مورد موقعیت یک پدیده است.

با توجه به مطالب بالا، هدف اصلی این تحقیق، ارائه روشی برای تعیین موقعیت ماشینی آدرس‌های توصیفی می‌باشد تا اولاً چالش تعیین موقعیت ماشینی این آدرس‌ها برطرف شود ثانیاً امکان بهره‌مندی از مزایای آدرس‌های توصیفی با جایگزینی آن‌ها با آدرس‌های استاندارد، از بین نرود.

برای تجزیه متن یک آدرس توصیفی به یک زبان رسمی متشکل از کلمات و گرامر معتبر و البته منعطف نیاز است. این گرامر در قالب BNF تعریف شد. بر اساس گرامر تعریف شده، هر آدرس مجموعه‌ای از گروه‌های مکانی می‌باشد. این گروه‌های مکانی مشتمل بر «مکان نام ثابت» و «مکان نام متغیر» می‌باشند که می‌توانند با «روابط مکانی» نیز ترکیب شوند. مکان‌نام‌های ثابت و روابط مکانی موجود در آدرس‌های توصیفی از قبل، تعیین و فهرست شده‌اند. از این رو، در میان سه دسته نشانه‌های انتهایی آدرس (یعنی مکان‌نام‌های متغیر، مکان‌نام‌های ثابت و روابط مکانی)، دو دسته آخر از طریق

لیست‌های شمارش، شناسایی شده و مابقی اجزا، مکان‌نام متغیر محسوب می‌شوند. با این روش، مرزهای گروه‌های مکانی در آدرس، مستقل از علائم سجاوندی، تعیین می‌شوند. به این ترتیب، آدرس‌های توصیفی و غیراستاندارد، بر اساس گرامر رسمی تعریف شده، قابل تجزیه شدند. خروجی حاصل از تجزیه متن آدرس با روش مذکور، تعیین نوع، محتوا و ترتیب هر کدام از گروه‌های مکانی در متن آدرس بود که به عنوان ورودی مرحله انطباق مکانی محسوب می‌شود.

برای مرحله انطباق مکانی اجزا، می‌بایست برای آدرس‌های توصیفی که ساختار مشخصی نداشته، یعنی بر مبنای مؤلفه‌های ادراک مکانی و در قالب ساختارهای زبان طبیعی بوده، مدل داده جدیدی با در نظر گرفتن نیازمندی‌های مختص به این حالت، تعریف شود.

یک مدل داده، ترکیبی از حداقل سه جزء «مجموعه‌ای از انواع ساختار داده»، «مجموعه‌ای از عملگرها یا قواعد استنتاج» و «مجموعه‌ای از قواعد جامعیت عمومی» می‌باشد. از این رو برای توسعه مدل داده پیشنهادی، لازم بود روابط مکانی موجود در آدرس‌های توصیفی، برای مؤلفه‌های ادراک مکانی مؤثر بر آدرس‌یابی، تعیین؛ روابطی که قابل استخراج از روابط دیگر هستند (روابط وابسته) و مابقی روابطی که قابل استخراج از روابط دیگر نیستند (روابط مستقل) شناسایی و نهایتاً نحوه استخراج روابط وابسته از روابط مستقل برای تبیین نحوه تعیین موقعیت روابط مکانی موجود در آدرس‌های توصیفی تعریف شوند. در ادامه، مؤلفه‌های ادراک مکانی آدرس‌های توصیفی به عنوان اجزای مدل داده مورد استفاده در مرحله انطباق مکانی تعریف و مدل پیشنهادی به مدل‌های موجودنگاشت شد. همچنین رویکرد جستجو که در این مرحله برای ماشین انتخاب شد، برعکس روش انسان برای تعیین موقعیت آدرس‌های توصیفی، از انتهای آدرس و آن هم

در ابتدا صرفاً میان مکان‌نام‌های ثابت، انجام می‌گیرد. دلایل انتخاب این رویکرد جلوگیری از اشتباه ناشی از انباشت اطلاعات اضافی در متن آدرس و تفاسیر مشتبه از مفاهیم اجزای آدرس بود. مزیت اصلی این رویکرد، کاهش دو مشکل مذکور می‌باشد. به این صورت که از آنجا که با حداقل اطلاعات مورد نیاز ماشین، تعیین موقعیت انجام می‌شود، امکان اشتباه ناشی از انباشت اطلاعات اضافی در متن آدرس، کاهش می‌یابد. چرا که تعیین موقعیت با حداقل اطلاعات مورد نیاز، انجام می‌شود و نیازی به استفاده از همه اجزای آدرس نیست. همچنین از آنجا که جستجو در میان روابط مکانی، صرفاً پس از عدم تعیین موقعیت با استفاده از مکان-نام‌های ثابت انجام می‌گیرد، لذا امکان اشتباه ناشی از تفسیر مشتبه این روابط، کاهش می‌یابد و اساساً با این رویکرد، نقش روابط مکانی در تعیین موقعیت آدرس توصیفی، کمرنگ شده و بنابراین امکان وقوع اشتباه ناشی از ابتدائی تفسیر این روابط بر بافت، کم می‌شود.

درباره نقش ادراک مکانی در تعریف مدل داده و زبان رسمی، می‌توان گفت: آدرس‌های توصیفی، ارتباط وثیقی با زبان‌های طبیعی دارند، چرا که توصیف، یکی از کارکردهای زبان‌های طبیعی بوده و آدرس توصیفی نیز از زیرمجموعه‌های توصیف محسوب می‌شود. به بیان دیگر، توصیف مکان یکی از کارکردهای اولیه زبان طبیعی محسوب می‌شود. همچنین ادراک به صورت عام و ادراک مکانی به صورت خاص با زبان طبیعی در هم تنیده است. از طرفی همانطور که بیان شد، اجزای مدل داده تعریف شده بر مبنای مؤلفه‌های ادراکی آدرس‌های توصیفی می‌باشد. از این رو ملاحظه می‌شود ادراک مکانی در تعریف مدل داده و زبان رسمی در تحقیق حاضر، نقش زیربنایی ایفا کرده است.

مراجع

- [1] Proposal to Create Street Address Data Standards Submitted to the Federal Geographic Data Committee (FGDC), 2007, Urban and Regional Information Systems Association (URISA)
- [2] PSMA, A Geocoded National Address File for Australia: The G-NAF What, Why, Who and When, 2003.
- [3] TIGER. 2014; Available from: <http://www.census.gov/geography.html>.
- [4] Address-Point database. 2014; Available from: <http://www.ordnancesurvey.co.uk/business-and-government/products/address-point.html>.
- [5] PSMA, G-NAF – yesterday, today and tomorrow, 2014.
- [6] Weiser, P. and A.U. Frank, Cognitive Transactions – A Communication Model, in Spatial Information Theory, T. Tenbrink, et al., Editors. 2013, Springer International Publishing. p. 129-148.

- [7] Karimipour, F., et al., Spatial Knowledge Acquisition from Addresses, in COSIT2015: Santa Fe, New Mexico, USA.
- [8] Longley, P.A., et al., Geographic information science and systems. 2015: John Wiley & Sons.
- [9] Wakabayashi, Y., S. Itoh, and Y. Nagami, The Use of Geospatial Information and Spatial Cognition of Taxi Drivers in Tokyo. *Procedia - Social and Behavioral Sciences*, 2011. 21(0): p. 353-361.
- [10] A, B. and J.-O. P, Is Route Learning More Than Serial Learning? *Spatial Cognition & Computation*, 2008. 8(4): p. 289-305.
- [11] Schuldes, S., et al., Modeling spatial knowledge for generating verbal and visual route directions, in *Knowledge-Based and Intelligent Information and Engineering Systems*. 2011, Springer. p. 366-377.
- [12] Jacob, R., et al., Pedestrian navigation using the sense of touch. *Computers, Environment and Urban Systems*, 2012. 36(6): p. 513-525.
- [13] Tomko, M., *Destination descriptions in urban environments*, 2007, Citeseer.
- [14] Tomko, M. and S. Winter, Pragmatic construction of destination descriptions for urban environments. *SpatialCognition & Computation*, 2009. 9(1): p. 1-29.
- [15] Karimipour, F. and A. Niroo. Agent-based modelling of cognitive navigation: How spatial knowledge is constructed and used. in *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*. 2013. Budapest, Hungary.
- [16] Hirtle, S.C., *Geographical design spatial cognition and geographical information science*. 2011: Morgan & Claypool.
- [17] Yang, D.H., et al., Improving geocoding practices: evaluation of geocoding tools. *Journal of medicalsystems*, 2004. 28(4): p. 361-70.
- [18] Davis Jr, C.A. and F.T. Fonseca, Assessing the certainty of locations produced by an address geocoding system. *Geoinformatica*, 2007. 11(1): p. 103-129.
- [19] Davis, C.A., F.T. Fonseca, and K.A. Borges. A Flexible Addressing System for Approximate Geocoding. in *GeoInfo*. 2003.
- [20] Drummond, W.J., Address Matching: GIS Technology for Mapping Human Activity Patterns. *Journal of the American Planning Association*, 1995. 61(2): p. 240-251.
- [21] J, B., et al., Geocoding police collision report data from California: a comprehensive approach. *International Journal of Health Geographics*, 2009. 8(1): p. 72.
- [22] Roongpiboonsopit, D. and H. Karimi, Quality Assessment of Online Street and Rooftop Geocoding Services. *Cartography and Geographic Information Science*, 2010: p. 301-318.
- [23] Goldberg, D.W., Advances in Geocoding Research and Practice. *Transactions in GIS*, 2011. 15(6): p. 727-733.
- [24] DW, G. and G.M. Jacquez, Advances in geocoding for the health sciences. *Spatial and Spatio-temporal Epidemiology*, 2012. 3(1): p. 1-5.
- [25] Qin, X., et al., Intelligent geocoding system to locate traffic crashes. *Accident Analysis & Prevention*, 2013. 50(0): p. 1034-1041.
- [26] O'Reagan, R.T. and D. United States. Bureau of the Census. Statistical Research, Geocoding theory and practice at the Bureau of the Census. 1987, Washington, D.C.: Bureau of the Census.
- [27] SE, C., et al., Errors in postcode to enumeration district mapping and their effect on small area analyses of health data. *Journal of public health medicine*, 1998. 20(3): p. 325-30.
- [28] Yu, L., Development and evaluation of a framework for assessing the efficiency and accuracy of street address geocoding strategies, 1996, State University of New York at Albany: NY, USA. p. 164.
- [29] Lee, J ,GIS-based geocoding methods for area-based addresses and 3D addresses in urban areas. *Environment and Planning B: Planning and Design*, 2009. 36(1): p. 86-106.
- [30] MR, B., et al., Positional accuracy of geocoded addresses in epidemiologic research. *Epidemiology*, 2003. 14(4): p. 408-412.
- [31] M, C. and T. T, Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics*, 2003. 2(1): p. 10.
- [32] Jacquez, G.M. and R. Rommel, Local indicators of geocoding accuracy) LIGA): theory and application. *International Journal of Health Geographics*, 2009. 8.

- [33] Kumar, S., M. Liu, and S.-A. Hwang, A multifaceted comparison of ArcGIS and MapMarker for automated geocoding. *Geospatial Health*, 2012. 7: p. 145-151.
- [34] Davis Jr, C.A. and R.O. de Alencar, Evaluation of the quality of an online geocoding resource in the context of a large Brazilian city. *Transactions in GIS*, 2011. 15(6): p. 851-868.
- [35] Goldberg, D.W., J.P. Wilson, and C.A. Knoblock, From text to geographic coordinates: the current state of geocoding. *URISA Journal*, 2007. 19(1): p. 33--47.
- [36] Hutchinson, M.J. and B. Veenendaal, An agent-based framework to enable intelligent geocoding, 2011, ISPRS.
- [37] Karimi, H.A., M.H. Sharker, and D. Roongpiboonsopit, Geocoding Recommender: An Algorithm to Recommend Optimal Online Geocoding Services for Applications. *Transactions in GIS*, 2011. 15(6): p. 869-886.
- [38] DW, G., Improving Geocoding Match Rates with Spatially-Varying Block Metrics. *Transactions in GIS*, 2011. 15(6): p. 829-850.
- [39] Tolman, E.C., Cognitive maps in rats and men. *Psychological review*, 1948. 55(4): p. 189-208.
- [40] Siegel, A. and S. White, The Development of Spatial Representations of Large-Scale Environments, in *Advances in Child Development and Behaviour*, H.W. Reese, Editor. 1975, Academic Press. p. 9-55.
- [41] Neisser, U., *Cognition and reality: Principles and implications of cognitive psychology*. 1976: WH Freeman/Times Books/Henry Holt & Co.
- [42] Lakoff, G., *Cognitive semantics. Meaning and mental representations*, 1988. 119: p. 154.
- [43] Raubal, M., Agent-based simulation of human wayfinding, in *Faculty of Science and Informatics 2001*, Vienna University of Technology.
- [44] Campbell, J., *Handbook of mathematical cognition*. 2005, New York: Psychology Press.
- [45] Montello, D.R., Cognitive Research in GIScience: Recent Achievements and Future Prospects. *Geography Compass*, 2009. 3: p. 1824 - 1840.
- [46] Raubal, M., Cognitive Engineering for Geographic Information Science. *Geography Compass*, 2009. 3(3): p. 1087-1104.
- [47] Winter, S., Spatial Intelligence: Ready for a Challenge? *Spatial Cognition & Computation*, 2009. 9(2): p. 138-151.
- [48] Morash, V., et al., A Review of Haptic Spatial Abilities in the Blind. *Spatial Cognition & Computation*, 2011. 12(2-3): p. 83-95.
- [49] Kuipers, B., The "map in the head" metaphor. *Environment and Behavior*, 1982. 14(2): p. 202-220.
- [50] Tversky, B., Cognitive maps, cognitive collages, and spatial mental models, in *Spatial Information Theory A Theoretical Basis for GIS*. 1993, Springer. p. 14. r z-
- [51] Mark, D., Cognitive perspectives on spatial and spatio-temporal reasoning. *Geographic information research bridging the Atlantic*, 1997: p. 308-319.
- [52] Hirtle, S.C., The cognitive atlas: using GIS as a metaphor for memory. *Spatial and temporal reasoning in geographic information systems*, 1998: p. 267-276.
- [53] Kuipers, B., Modeling Spatial Knowledge*. *Cognitive science*, 1978. 2(2): p. 129-153.
- [54] Downs, R.M. and D. Stea, *Maps in minds: Reflections on cognitive mapping*. 1977: Harper & Row New York.
- [55] Lynch, K., *The image of the city*. Vol. 11. 1960: the MIT Press.
- [56] Allen, J., *Natural language understanding*. 1995, Redwood City, California: Benjamin/Cummings Publishing Company.
- [57] Chopra, A., A. Prashar, and C. Sain, Natural Language Processing. *International Journal of Technology Enhancements and Emerging Engineering Research*, 2013. 1(4).
- [58] L, B., *Natural language communication with computers*. 1978, Berlin; New York: Springer-Verlag.
- [59] Horacek, H., *New concepts in natural language generation: planning, realization and systems*. 2015: Bloomsbury Publishing.

- [60] Lees, R.B. and N. Chomsky, Syntactic Structures. Language, 1957. 33(3 Part 1): p. 375-408.
- [61] Dostert, L.E., The georgetown-ibm experiment. 1955). Machine translation of languages. John Wiley & Sons, New York, 1955: p. 124-135.
- [62] Schank, R.C., et al. MARGIE: Memory Analysis Response Generation, and Inference on English. in IJCAI. 1973.
- [63] Cullingford, R.E., Script application: computer understanding of newspaper stories, 1978, DTIC Document.
- [64] Lehnert, W.G. A conceptual theory of question answering. in Proceedings of the 5th international joint conference on Artificial intelligence-Volume 1. 1977. Morgan Kaufmann Publishers Inc.
- [65] Lehnert, W.G., Plot units and narrative summarization. Cognitive Science, 1981. 5(4): p. 293-331.
- [66] Wilensky, R., Pam. Inside computer understanding: Five programs plus miniatures, 1981: p. 136-179.
- [67] Hutchins, W.J., Machine translation: A brief history. Concise history of the language sciences .from the Sumerians to the cognitivists, 1995: p. 431-445.
- [68] Tyers, F.M., F. Sánchez-Martínez, and M.L. Forcada. Unsupervised training of maximum-entropy models for lexical selection in rule-based machine translation. in Proceedings of the 18th Annual Conference of the European Association for Machine Translation. 2015.
- [69] Evans, J.A. and P. Aceves, Machine Translation: Mining Text for Social Theory. Annual Review of Sociology, 2016(0.)
- [70] Luong, M.-T., et al., Multi-task sequence to sequence learning .arXiv preprint arXiv:1511.06114, 2015.
- [71] HC, B. and T. M, Recent advances in parsing technology. 1996, Dordrecht; Boston: Kluwer Academic Publishers.
- [72] Grune, D. and C.J.H. Jacobs, Parsing techniques. 2005, New York; London: Springer.
- [73] Greibach, S ,Inverses of Phrase Structure Grammars, 1963, Ph. D. Thesis, Harvard University.
- [74] Bolc, L. and J.G. Carbonell, Natural language parsing systems. 1987, Berlin; New York: Springer-Verlag.
- [75] Johnson, M., PCFG models of linguistic tree representations .Computational Linguistics, 1998. 24(4): p. 613-632.
- [76] Charniak, E. A maximum-entropy-inspired parser. in Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference. 2000. Association for Computational Linguistics.
- [77] Socher, R., et al. Parsing natural scenes and natural language with recursive neural networks. in Proceedings of the 28th international conference on machine learning (ICML-11). 2011.
- [78] Chomsky, N., Rules and Representations. The Behavioraland Brain Sciences, 1980. 3: p. 1-61.
- [79] Frank, A.U., PRACTICAL GEOMETRY, 2006.
- [80] Codd, E.F., Relational database: a practical foundation for productivity. Communications of the ACM, 1982. 25(2): p. 109-117