

پایگاههای معرفتی در سیستم های اطلاع رسانی

صفحه: ۶۶-۶۸

فصلنامه اطلاع رسانی، دوره ۱۱، شماره ۱، زمستان ۱۳۷۳

دکتر مرتضی انواری

دانشگاه کالیفرنیا برکلی و مرکز پژوهشهای مجلس شورای اسلامی
ملك آفاق فتحیان پور

کلید واژه ها :

• پایگاههای معرفتی * اطلاع رسانی * بانکهای اطلاعاتی

چکیده :

امروزه ، در بسیاری از رشته های فعالیتهای بشری مانند پزشکی ، مهندسی و مدیریت سیستمهایی به وجود آمده است که در آنها ، اطلاعات و معرفت بشری به صورت الکترونیک ضبط می گردد . این سیستمها ، قادرند به همان گونه که یک فرد متخصص ، مانند پزشک از اطلاعات خود استفاده کرده و استدلال منطقی و نهایتاً نتیجه گیری می نماید ، از اطلاعات ضبطشده درون سیستم بهره گرفته ، استدلال منطقی و نتیجه گیری کند . در این مقاله ، مطالب زیر در شش بخش بررسی خواهد شد .

بخش ۱) تعریف پایگاههای معرفتی

بخش ۲) ویژگیهای پایگاههای معرفتی

بخش ۳) طراحی پایگاههای معرفتی

بخش ۴) زمینه های مناسب برای پایگاههای معرفتی

بخش ۵) استفاده از منطقیهای غیر ارسطویی در پایگاههای معرفتی

بخش ۶) تکنولوژی پایگاههای اطلاعاتی شیء گرا (Object oriented) .

بخش ۱ - تعریف پایگاههای معرفتی

در اینجا ، بهتر است به چند روش نمایش " معرفت " اشاره نمود . معرفت انسان به انحاء مختلف نمایش داده می شود . یکی از متداولترین روشها ، آن است که معرفت انسان را به صورت یک رشته قاعده در می آورند . به عنوان مثال ، در رشته پزشکی می توان گفت : " اگر کسی تب داشت و این تب توام با عطسه و گلو درد نیز بود ، این شخص به احتمال ۸۰٪ به بیماری آنفولانزا - در مرحله اولیه آن - مبتلاست " .

مثال دیگر : " اگر ماشینی روشن نشود . دلیل آن را باید در عوامل زیر جستجو نمود : ضعیف بودن باتری ، نداشتن بنزین ، و یا خراب بودن Starter " . بنابراین در اینجا نیز ، می توان عوامل فوق را به صورت یک قاعده بیان نمود که مثلاً : اگر ماشین روشن نشد ، ابتدا باتری را چک کنید ، اگر باتری نیرو دارد ، سپس بنزین ماشین را کنترل نمایید و اگر بنزین نیز به اندازه کافی وجود داشت ، باید Starter را چک کرد و غیره .

رهیافت دیگری که برای نمایش معرفت به کار می برند ، رهیافت " فریم " (Fram) است . " فریم " در واقع ، سازه ای (Construct) است جامعتر از " رکورد " که در پایگاههای اطلاعاتی به کار برده می شود . به طور مثال : در این روش می توان ، یک فیل را توسط " فریم " توصیف نمود . " فیل حیوانی است چهارپا ، خرطوم و عاج دارد ، رنگش معمولاً خاکستری است ، و وزنش حدوداً بین ۵۰۰ تا ۱۰۰۰ کیلوگرم است و غیره " .

لازم به تاکید است که اکثر پایگاههای معرفتی ، در واقع از " قاعده " استفاده می کنند . در این زمینه ، می توان در رشته های مختلف پزشکی و مهندسی ، سیستمهای زیادی را نام برد مانند mycin , intemist , explorer و نظایر آن .

بخش ۲ - ویژگیهای پایگاههای معرفتی

یک سیستم پایگاه معرفتی ، دارای چند خصوصیت به شرح زیر است :

۱-۲ - نگهداری اطلاعات

۲-۲ - استدلال و نتیجه گیری از اطلاعات ذخیره شده در آن

۳-۲ - استفاده از روش interpolation ، به آن معنا که با توجه به تجاربی که اندوخته است می تواند حتی در زمینه هایی که مورد آنها در داخل پایگاه معرفتی وجود ندارد ، قضاوت نماید .

۴-۲ - پایگاههای معرفتی قادرند چگونگی حصول به هر نتیجه گیری را توضیح دهند .

۵-۲ - این سیستمها ، صرفاً از روشهای الگو ریمتی استفاده نمی کنند ؛ ولی البته استفاده از الگو ریمتها نیز جزو

سیستم است. روش دیگری که این سیستمها، به کار می‌گیرند روش heuristic است. در روش heuristic نتیجه گیری بر مبنای الگو و یا نمونه ای است که در وضعیت موجود، به چشم می‌خورد - بعلاوه استفاده از تجارب قبلی، به عنوان مثال: می‌توان گفت که یک شطرنج باز ماهر، چگونگی حرکت خود را بر مبنای الگوی صفحه شطرنج و تجارب شخصی خودش انجام می‌دهد که همان استفاده از روش heuristic است.

بخش ۳ - طراحی پایگاههای معرفتی

یک پایگاه معرفتی از چند مولفه تشکیل می‌شود:

۱-۲ - مولفه اول، آن قسمتی است که معرفت بشری در آن نگهداری می‌شود. این معرفت، ممکن است توسط متخصصان و یا از محتوای کتابها و مجلات و ... گرفته شده و به صورت قواعد در آمده باشد.

۲-۲ - inference engine و یا موتور استنتاج که توسط آن، می‌توان بر مبنای اطلاعات ذخیره شده به استدلال پرداخت و به نتیجه های تازه دست یافت. موتور استنتاج، در هر سیستمی قادر است که قواعد موجود در سیستم را با هم ترکیب نماید. به طور مثال: اگر قاعده ای به صورت:

if * is a then y is b

داشته باشیم. و اگر قاعده دیگری نیز به این صورت که:

If y is b then z is c

موتور استنتاج نتیجه می‌گیرد که:

if * is a then z is c

این نتیجه گیرها را قسمت موتور استنتاج پایگاه معرفتی انجام می‌دهد.

۳-۳ - مولفه دیگر پایگاههای معرفتی user interface است و آن برنامه ای است که بین "کاربر" و "پایگاه معرفتی" واسطه است.

۴-۳ - در پاره ای از پایگاههای معرفتی، مولفه ای وجود دارد که می‌توان آن را module natural language understanding و یا فهم زبان طبیعی نامید. این قسمت در واقع، مربوط به درک و فهم ماشینی متون است. مثلاً اگر متنی به سیستم داده شود، آن متن توسط این مولفه بررسی می‌شود و به صورت تعدادی قاعده در می‌آید. مولفه مذکور که یکی از قسمتهای بسیار مهم یک پایگاه معرفتی است. متاسفانه هنوز آن چنان که باید و شاید پیشرفت نموده است. برای این عدم پیشرفت، چند علت وجود دارد که مهمترین آن را می‌توان، در ابهام، ابهام و پیچیدگی زبان طبیعی جستجو کرد. به عبارت دیگر، برای آنکه یک سیستم کامپیوتری بتواند یک متن را بفهمد، باید همان قدر در زمینه آن متن اطلاع داشته باشد که، یک انسان در آن مورد اطلاع دارد. در واقع، به همین علت است که اکثر پایگاههای معرفتی موجود، فاقد یک چنین مولفه ای هستند.

بخش ۴- زمینه های مناسب برای پایگاههای معرفتی

ضوابط موجود برای به وجود آمدن زمینه های مناسب، جهت ایجاد پایگاه معرفتی به شرح زیر است:

۱ - ۴ - وجود متخصصان برجسته ای که بتوانند اطلاعات لازم را در اختیار مهندسان معرفتی (Knowledge engineer) قرار دهند.

۲ - ۴ - وجود نداشتن مسائل پیچیده و مشکل اجتماعی و اقتصادی و یا مسائل پیچیده روانی و روحی. مثلاً به جرئت می‌توان گفت که زمینه هایی همچون روانکاو، چون با روح و ذهن انسان سرو کار دارند، برای ایجاد پایگاههای معرفتی مناسب نیستند.

۳ - ۴ - به طور کلی زمینه هایی که درای آیین نامه و قاعده و قانون هستند، برای ایجاد پایگاههای معرفتی مناسب اند. مثلاً در نیروگاههای اتمی، برای امور ایمنی دستورالعملهای مفصلي وجود دارد که در آنها اقدامات مجاز، ممنوع، ممکن و غیر ممکن در شرایط اضطراری، توضیح داده شده است. این دستورالعملها، از منطق خاصی پیروی می‌نمایند که به آن منطق مدال (modal) یا منطق deontic میگویند. درباره این منطق در بخش بعدی سخن خواهیم گفت.

بخش ۵ - استفاده از منطقهای غیر ارسطویی در پایگاههای معرفتی

در این بخش درباره منطقهای سخن خواهیم گفت که، در پایگاههای معرفتی کاربردهای بسیاری دارند. این منطقها، یا منطق ارسویی و یا دوتایی تفاوت اساسی دارند.

۱ - ۵ - منطق نادقیق، در مواردی به کار می‌رود که معرفت انسان، به صورت قاعده هایی توأم با عدم دقت مطرح می‌شود. مثلاً گفته می‌شود: اگر درجه حرارت یک دیگ بخار به سرعت بالا رفت، شیر اطمینان را آسته باز کنید. " به سرعت بالا رفتن" و "آسته باز کردن"، دو عمل نادقیق هستند؛ اما در بین متخصصان متداول است که معرفت

خود را بدان گونه بیان نمایند .

منطق نادقیق ، فرمالیسمی است که می توان توسط آن ، با این نوع گزاره ها استدلال کرد . مثلاً اگرچند قاعده وجود داشته باشد که در آنها خبر يك قاعده مبتدای قاعده دیگر است و یا آنکه مبتدای دو جزء مختلف داشته باشد و یا بخواهیم گزاره ای را نفی نماییم ، فرمالیسم منطق نادقیق این عملیات را انجام می دهد و از آن نتیجه گیری میکند . هم اکنون در دنیا ، سیستم های زیادی بر پایه این منطق وجود دارد . به عنوان مثال : در زمینه کنترل ، می توان از قطار زیرزمینی sendei (شمال ژاپن) و دوربین فیلمبرداری ویدئویی پاناسونیک و ماشین لباسشویی پاناسونیک ، و جعبه دنده اتوماتیک اتومبیل هوندا نام برد .

۵ - ۲ - گزاره هایی که حتمی نیستند (uncertain) و وقوع آنها احتمالی است ، در غالب فرمالیسم منطق احتمال می گنجد . این با منطق نادقیق تفاوت دارد . تفاوت آنها در این است که در منطق نادقیق ، ما با گزاره هایی سر و کار داریم که با عدم دقت توأم هستند . و این عدم دقت ، به شکل حرفی (linguistic) بیان می گردد ، حال آنکه در گزاره های احتمالی ، عنصر عدم حتمیت وجود دارد .

۵ - ۳ - منطق مدال (modal logic) که به آن deontic logic نیز اطلاق می شود . در این منطق مبتدا و خبر به تعبیر متداول آن وجود ندارد . در منطق دوگانه ارسطویی قیاس مادی (implication material) در تمام موارد درست است - جز در موردی که در آن مبتدا درست و خبر نادرست باشد . در حالی که در منطق مدال ، شرایط وضع موجود ، به عمل خاصی رهنمون می گردد . به طور : مثال به اپراتور يك نیروگاه اتمی تعلیم داده می شود که در هر شرایط اضطراری خاص ، متناسب با آن شرایط اقدامی را به عمل آورد .

۵ - ۴ - منطقهای دیگر نیز ، در پایگاههای معرفتی به کار برده می شود ، من جمله : now monotonic logic (منطق نایکناخت) ، inductive logic (منطق استقراری) و cace based reasoning (استدلال موردگرا) که تفصیل آنها از حوصله این مقاله خارج است .

بخش ۶ - تکنولوژی پایگاههای اطلاعاتی شیء گرا (oriented object)

در اوایل سالهای ۱۹۷۰ ، سیستم پایگاههای اطلاعاتی رابطه ای relational به وجود آمدند . در این سیستمها ، اطلاعات به صورت پرونده هایی در داخل حافظه مغناطیسی نگهداری می شوند . هر پرونده از تعدادی رکورد تشکیل شده و هر يك از رکوردها ، نمایش دهنده اطلاعات مربوط به يك واقعیت (entity) است .

در اواسط دهه ۱۹۸۰ ، سیستم پایگاههای اطلاعاتی ، به این نتیجه رسیدند که " سیستمهای رابطه ای " ، به اندازه کافی قابلیت نمایش اطلاعات پیچیده را ندارند . مصلأً نمایش اشیاء هندسی سه تعدی و ارتباط فضایی آنها با یکدیگر ، مستلزم استفاده از ساختارهایی است که رابطه ها (relations) قادر به انجام آن نیستند . این عدم قابلیت ، به ایجاد " سیستمهای اطلاعاتی شیء گرا " (object oriented) منجر شد که در آن ، يك سیستم فیزیکی و یا اداری به وسیله تعدادی شیء (object) نمایش داده می شود و ضمناً ارتباط بین مولفه های يك سازمان و یا يك سیستم فیزیکی نیز ، به وسیله رابطه بین اشیاء نشان داده می شود . این تکنولوژی ، نسبت به تکنولوژی رابطه ای ، محاسن و برتریهایی دارد که ، مهمترین آنها عبارت است از : تلقین " داده ها " با اپراتورهایی که روی " داده ها " عمل می کنند .

یکی از این سیستمها ، Postgres است که ، در دانشگاه کالیفرنیا طرح و اجرا شد . در این سیستم ، علاوه بر " داده ها " ، " قواعد " نیز ذخیره می شوند . ضمناً اشکال هندسی ای که هر يك ، از اشکال ساده تري تشکیل می شوند نیز ، در این سیستم به آسانی قابل نمایش ، ذخیره و بازیابی اند . بحث تفصیلی در این زمینه ، به فرصت دیگری موکول می شود .