

موتورهای جستجوی هوشمند

در مطالعاتی که با هدف دستیابی به راهکارهایی جدید برای طراحی نسل جدید جستجوگرهای اینترنتی در مرکز تحقیقات IBM در نیویورک صورت پذیرفت...



جام جم آنلاین: در مطالعاتی که با هدف دستیابی به راهکارهایی جدید برای طراحی نسل جدید جستجوگرهای اینترنتی در مرکز تحقیقات IBM در نیویورک صورت پذیرفت، توانایی 2 انسان و یک ابرکامپیوتر برای پاسخگویی به یک سوال مشخص به رقابت گذاشته شد. جالب این که در پایان ابرکامپیوتر طراحی شده توسط محققان IBM - که واتسون نام دارد - توانست 2 شرکت کننده دیگر این رقابت را مغلوب خود سازد.

در 2 دوره رقابت‌هایی که در ژانویه 2011 برگزار شد واتسون علاوه بر حل کردن جداول واژگان موفق شد به سؤالاتی که درباره سیاهچاله‌ها مطرح شده بود نیز در مدت زمانی کوتاه و بدرستی پاسخ گوید. اگرچه این ابرکامپیوتر گاهی در مسیری که برای یافتن پاسخ صحیح به سؤالات مطرح شده طی می‌کرد، مرتکب اشتباهات جالبی می‌شد، اما در نهایت توانست بر 2 رقیب خود که از قهرمانان مسابقات جنوپاردی (Jeopardy) سال گذشته بودند غلبه کند. 14 سال از زمانی که ابرکامپیوتر Deep Blue که از اجداد و نیاکان ابرکامپیوترهای امروزی نظیر واتسون بوده است توانست گری کاسپاروف، قهرمان و استاد بازی شطرنج در سطح دنیا را مغلوب خود سازد، می‌گذرد. اما در حقیقت بازی شطرنج یک رقابت ریاضیاتی است که می‌تواند در یک نبرد شانه به شانه و تا مرز توان محاسباتی کامپیوترها پیش برود. اما موفقیت واتسون در این رقابت آزمایشی نشان‌دهنده این است که سرانجام کامپیوترها موفق شده‌اند حتی در قلمرو دانش و حیطه فهم و ادراک زبان طبیعی و تقابل منطقی خواسته و پاسخ نیز از انسان‌ها پیشی گیرند.

نسل جدیدی از موتورهای جستجو در راهند

ورود سیستم‌های جدید کامپیوتری مانند واتسون که مثل انسان از قابلیت فهم زبان برخوردارند، نویدبخش این است که بزودی شاهد رونمایی از نسل جدیدی از سیستم‌های محاسباتی و پردازشگر کامپیوتری خواهیم بود. طراحی موتورهای جستجوی اینترنتی براساس عملکرد این سیستم‌ها، این امکان را در اختیار جستجوگرهای هوشمند قرار می‌دهد تا بتوانند علاوه بر دسترسی به واقعیت‌ها و حقایق موجود، مسوولیت بخشی از تفکرات انسانی را نیز به عهده گیرند. ابرکامپیوتر واتسون تنها یک نمونه از فناوری‌های جدیدی است که در طراحی و ساخت جستجوگرهای هوشمند مورد استفاده قرار گرفته است. این فناوری‌ها به جستجوگرها کمک می‌کنند تا بتوانند براساس تجزیه و تحلیل داده‌های موجود و ایجاد یک ارتباط منطقی بین این داده‌ها به مجموعه‌ای از سؤالات مختلف پاسخ گویند. پیش‌بینی می‌شود که با استمرار این روند، پس از گذشت 5 سال این سیستم‌ها حتی بتوانند به سؤالاتی که ممکن است تا به حال کمتر به آنها فکر کرده باشیم نیز پاسخگو باشند. محققان مرکز تحقیقاتی IBM بر این باورند که حجم اطلاعات تولیدی توسط ما انسان‌ها هر سال به بیش از 2 برابر افزایش می‌یابد که بخشی از آن مدیون تلاش‌ها و زحمات متخصصانی است که با استفاده از فناوری‌های دیجیتال امکان دسترسی به اطلاعات پایه و تولید دانسته‌های جدید را فراهم می‌کنند. اما با توجه به این موضوع، ابرکامپیوترهایی که قرار است مبنای طراحی و ساخت موتورهای جستجوگر آینده قرار گیرند، باید چه ویژگی‌هایی داشته باشند؟ بدیهی است که این جستجوگرها که شرکت‌های بزرگ و پیشگام در این عرصه همچون گوگل و IBM، هوشمندی را ویژگی منحصر به فرد آنها در مقایسه با جستجوگرهای امروزی می‌دانند، باید بتوانند به جای استفاده از کلمات و واژه‌های کلیدی برای یافتن پاسخ سوال مطرح شده، از توانایی فهم اطلاعات موجود در پایگاه داده‌ها برخوردار باشند و به عبارتی دیگر بتوانند به جای ما این اطلاعات را خوانده، بفهمند و پاسخ مناسب را بیابند. ابرکامپیوتر واتسون به آسانی از عهده انجام این کار برمی‌آید و پس از گذشت چنددهم ثانیه اطلاعات وارد شده را از نظر دستور زبان و واژه‌شناسی مورد تجزیه و تحلیل قرار می‌دهد و بر این اساس اطلاعات خواسته شده را که می‌تواند در محدوده وسیعی از نام یک کتاب، نام یک نویسنده یا نوعی وسیله خانگی قرار داشته باشد به راحتی پیدا می‌کند.

ورود سیستم‌های جدید کامپیوتری مانند واتسون که مثل انسان از قابلیت فهم زبان برخوردارند، نویدبخش این است که بزودی شاهد رونمایی از نسل جدیدی از سیستم‌های محاسباتی و پردازشگر کامپیوتری خواهیم بود اما این نوع جستجوگرها به‌رغم ویژگی‌های مثبتی که دارند با 2 محدودیت اساسی مواجه هستند. اول این که ابرکامپیوترهایی نظیر واتسون سوال و جواب‌ها را به عنوان مجموعه‌ای از متغیرهای صفر و یک پردازش می‌کنند و در نتیجه فاقد قدرت استدلال لازم برای پاسخگویی خواهند بود. برای مثال، این سیستم‌ها به آسانی درخواست خواهند یافت که آب در دمای صفر درجه سانتی‌گراد تبدیل به یخ خواهد شد، اما نمی‌توانند پیش‌بینی کنند که یخ‌زدگی معابر در یک شب سرد زمستانی چگونه می‌تواند برای عبور و مرور وسایل نقلیه محدودیت ایجاد کند و دوم این که این ابرکامپیوترها متشکل از مجموعه‌ای از کامپیوترها هستند و درست مانند این است که 2880 لپ‌تاپ را در کنار هم قرار داده باشید. پاسخگویی به هر سوال 2 تا 5 ثانیه از توان پردازشی این سیستم در مقیاس صنعتی را به خود

اختصاص خواهد داد و برای جلوگیری از ذوب شدن سیستم در نتیجه گرمای ایجاد شده در این مدت زمان کوتاه باید مجموعه‌ای از سیستم‌های خنک‌کننده در این مجموعه بزرگ فعال باشند و در این در حالی است که همزمان و در این چند ثانیه، موتور جستجوگر گوگل به طور متوسط به 68 هزار تا 170 هزار کاربر پاسخگو بوده است و برای هر یک از این جستجوها تنها یک بخش بسیار اندک از انرژی مصرفی واتسون را مصرف کرده است.

گوگل و نسل جدید موتور جستجوی گوگلی

با توجه به آنچه گفته شد می‌توان پیش‌بینی کرد که حداقل در آینده‌ای نزدیک، ابرکامپیوتر واتسون نمی‌تواند جایگزین جستجوگر کنونی گوگل شود. اما به نظر می‌رسد که این دو سیستم بتوانند با بهره‌گیری از ویژگی‌های مناسب همدیگر، نسل آینده جستجوگرها را بسازند. از آنجا که IBM قصد دارد در آینده‌ای نه‌چندان دور این فناوری را به شرکت‌ها و مراکز بزرگی مانند بانک‌ها و آزمایشگاه‌های داروسازی بفروشد، بنابراین دور از انتظار نخواهد بود که کاربران عادی نیز بتوانند در مدت زمان یک یا 2 سال آینده استفاده از این فناوری پاسخگویی به سوالات را به عنوان راهنمای پاسخگویی تلفنی یا راهنمای پاسخگویی به سوالات در صفحات وب تجربه کنند. اگرچه گوگل در تلاش است تا بتواند با ایجاد تغییراتی در ساختار موتور جستجوی خود به جای این که مجموعه‌ای از صفحات وب را برای یافتن سوال یا واژه کلیدی مطرح‌شده توسط کاربر ارائه نماید پاسخ این سوال را در اختیار جستجوکننده بگذارد، اما به گفته مدیران بخش موتور جستجوی گوگل، در ابتدا این خدمات جدید تنها محدود به پاسخگویی به سوالات ساده‌ای مانند یافتن بزرگ‌ترین شهر یا استان یک کشور خواهد بود، اما پیشرفت عملکرد این جستجوگرها، موتور جستجوی گوگل - که بهتر است از این پس آن را موتور پاسخگویی گوگل بنامیم - داده‌هایی از منابع مختلف اطلاعاتی را در کنار هم قرار می‌دهد تا بتواند پاسخ مناسبی را ارائه دهد. برای مثال اگر کاربری بخواهد درآمد سرانه ساکنان خاورمیانه را براساس میزان تحصیلات افراد در هر یک از کشورهای خاورمیانه مورد بررسی قرار دهد، موتور پاسخگو به کسب اطلاعات بیشتری از کاربر نیاز خواهد داشت، بنابراین به جای این که کاربر تنها 3 یا 4 واژه کلیدی را در بخش جستجو وارد کند لازم است که یک جمله یا حتی گاهی یک پاراگراف را در این بخش وارد کند. در چنین شرایطی گوگل باید از عملکردی مشابه ابرکامپیوتر واتسون برخوردار باشد تا بتواند به سوال مطرح شده پاسخ دهد، اما ویژگی مشترک گوگل و واتسون در این است که در پاسخ به موضوعات مطرح‌شده درباره گرمایش جهانی یا زندگی سیاستمداران بازتابی از اطلاعات و دانش عمومی موجود را منعکس خواهند ساخت که شاید چندان قابل اطمینان نباشد.

جستجوگرهای اختصاصی

2 سال پیش ریاضیدانی به نام استفان ولفرام از موتورهای پاسخگویی تحت عنوان موتورهای جستجوگر اطلاعاتی آلفا رونمایی کرد. این موتورهای جستجوگر از مسیر متفاوتی برای پاسخگویی به سوالات مطرح‌شده استفاده می‌کنند. ویژگی این جستجوگرها این است که در یافتن پاسخ سوالات از مجموعه اطلاعاتی که مورد تایید و بررسی انسان‌ها قرار گرفته است به عنوان یک منبع اطلاعاتی بهره می‌گیرند و بنابراین در مقایسه با جستجوگرهایی که در حال حاضر از آنها استفاده می‌کنیم پاسخ‌های محدودتری را در رابطه با سوال مطرح شده در اختیار کاربر قرار خواهند داد. استفاده از این روش درخصوص موضوعات خاص مانند موضوعات علمی از قابلیت اطمینان بیشتری برخوردار است.

برای مثال اگر «آب در دمای صفر درجه سانتی‌گراد» را به عنوان واژه‌های کلیدی در بخش جستجو وارد کنید، این جستجوگر بلافاصله درباره وضعیت، ساختار و مشخصات فیزیکی آب در این دما و همچنین چگالی یخ، نقطه جوش آب و همچنین دیگر اطلاعات تخصصی مرتبط با آن را در اختیار شما قرار خواهد داد. برخلاف گوگل و IBM، استفان ولفرام بر این باور است که افراد و موسسات باید مسئولیت تهیه اطلاعات برای کامپیوترها را برعهده گیرند و به جای اطلاعاتی که توسط کامپیوترها قابل خواندن باشند، اسناد و داده‌های محاسباتی را که قابل تجزیه و تحلیل آماری هستند به عنوان منبع اطلاعاتی در اختیار موتورهای جستجو قرار دهند. چنانچه در دهه‌های اخیر تبدیل اطلاعات چاپی به اطلاعات دیجیتالی یک ضرورت بوده است، ایجاد داده‌ها و اسنادی با قابلیت انجام محاسبات کامپیوتری نیز لازم و ضروری خواهد بود. اما بی‌شک ایجاد چنین تغییرات اساسی نیازمند تلاش گروه‌های متخصص انسانی، سرمایه‌گذاری و تعیین استانداردهایی است که می‌تواند مسیر دستیابی به نسل بعدی جستجوگرها را تسهیل سازد. ایده ساخت شبکه‌های ارتباطی با محتوای معنایی اولین بار در سال 1378/1999 توسط تیم برنر که شبکه ارتباطی جهانی (WWW) را برای نخستین بار راه‌اندازی کرد، مطرح شد. هدف از ساخت شبکه‌های ارتباطی معنایی ایجاد لایه‌های جدیدی در ساختارهای اطلاعاتی است که به کامپیوترها کمک می‌کند تا اطلاعاتی که از طریق آنلاین به روز می‌شوند را خوانده و براساس آنها عمل کنند. به این ترتیب با هوشمند شدن جستجوگرها، شبکه‌های ارتباطی و اطلاعاتی به منابع مفیدی برای یادگیری و ایجاد ارتباط بین انسان‌ها مبدل خواهند شد. بنابراین پیش‌بینی می‌شود که در آینده‌ای نزدیک به منابع عظیم اطلاعاتی دسترسی داشته باشیم که زمینه لازم برای توسعه و پیشرفت هوش مصنوعی در جنبه‌های مختلف زندگی انسان‌ها را امکان‌پذیر سازند. اما حال ممکن است این سوال مطرح شود که در دنیایی که سیستم‌های هوشمند اطراف بشر را فراگرفته است چه اطلاعات یا دانشی باید در ذهن انسان ذخیره شود؟

فرانک فراهانی‌جم / جام جم

