

همان‌طور که آرتور گلدهر، مترجم و پژوهشگر سرشناس فرانسوی می‌گوید بعضی پیچیدگی‌های زبانی این روند را دشوار کرده است.

مصائب ترجمه ماشینی؛



ایان — سال‌ها پیش در پروازی از آمستردام به بوستون، دو راهب آمریکایی سمت راست من نشسته بودند و به حرف‌های مرد جوان خوش‌صحبتی از اهالی هلند گوش می‌دادند. مرد جوان قصد داشت ایالات متحده را از نو کشف کند. او از راهبه‌ها پرسید اهل کجا هستند؟ افسوس که شهر فرامینگهام در ماساچوست جزو برنامه سفرش نبود؛ اما اعلام کرد که «یه خروار ۱ وقت داره و قراره از یه خروار جای مختلف دیدن کنه».

جوان سرخوش هلندی به‌وضوح این را دریافته بود که «خروار» مترادف جذابی برای «زیاد» بی‌رنگ‌وبوست. او علم نحو انگلیسی را به‌خوبی فرا گرفته بود و دایره لغات نسبتاً گسترده‌ای هم داشت؛ اما فاقد تجربه تناسب کلمات با بافت‌های اجتماعی بود.

خبر اخیر، که موتور ترجمه گوگل قرار است از سیستم عبارت‌محور به شبکه عصبی ۲ تغییر کند، این خاطره را به ذهنم آورد. تفاوت‌های فنی این دو در این لینک تشریح شده است. هر دو روش متکی به آموزش ماشین با یک «پیکره متنی» شامل جمله‌های جفتی‌اند: جمله اصلی و ترجمه آن. سپس کامپیوتر، براساس توالی کلمات در متن اصلی، قواعدی برای استنباط محتمل‌ترین توالی کلمات در زبان مقصد استخراج می‌کند.

این روال عملی در تطابق الگوهاست. زمانی که خطاب به گوشی هوشمندتان می‌گویید «مرا به بروکلین هدایت کن»، برای تفسیر هجاهای گفته‌شده، یا، هنگامی که یک برنامه عکس چهار دوستان را شناسایی می‌کند و به آن برچسب می‌زند، از الگوریتم‌های تطابق الگوی مشابهی استفاده می‌شود. ماشین چهره‌ها یا مقصدها را «نمی‌فهمد»؛ آن‌ها را به بردارهایی از اعداد تقلیل داده و پردازش می‌کند.

من مترجمی حرفه‌ای هستم و تاکنون حدود ۱۲۵ کتاب از زبان فرانسه ترجمه کرده‌ام. بنابراین ممکن است از من انتظار برود که در برابر ادعای گوگل از کوره در بروم که موتور ترجمه جدیدش تقریباً به‌خوبی مترجمی انسانی است، زیرا از صفر تا شش، نمره پنج را کسب می‌کند، درحالی‌که میانگین انسان‌ها 5.1 بوده است. اما علاوه‌برآن دکترای ریاضی هم دارم و نرم‌افزاری ساختم که روزنامه‌های اروپایی را به چهار زبان مختلف «می‌خواند» و نتایج را براساس موضوع دسته‌بندی می‌کند. پس به‌جای اینکه در برابر احتمال جایگزین‌شدن بامترجم ماشینی موضع دفاعی بگیرم، از قابلیت‌های چشمگیر ماشین‌ها آگاه و نسبت به پیچیدگی و مهارت تکنیکی کار گوگل سراپا ستایشم.

اما این ستایش چشم‌هایم را بر نواقص ترجمه ماشینی نمی‌بندد. مسافر هلندی را به یاد بیاورید که «یه خروار» انگلیسی می‌دانست. فصاحت مرد جوان نشان می‌داد که «خیس‌افزار» ۳ یا به عبارتی شبکه عصبی زند او به‌خوبی برای درک شهودی قواعد ظریف (و استثنایها)، که زبان را طبیعی می‌کنند، آموزش دیده بود. در مقابل، زبان‌های کامپیوتری دستورزبانی بدون بافت دارند. اما هلندی جوان فاقد تجربه اجتماعی از زبان انگلیسی بود تا قواعد ظریف‌تری را درک کند که به انتخاب واژه، لحن و ساختار زبانی یک بومی شکل می‌دهند. گوینده‌های بومی نیز ممکن است، برای دست‌یافتن به نتیجه‌ای خاص، شکستن این قواعد را انتخاب کنند. اگر من به دو نفر راهبه به‌جای «جاهای زیادی» می‌گفتم «یه خروار جا»، حتماً از آن منظوری می‌داشتیم؛ اما مرد هلندی به ورطه طنز ناخواسته افتاد.

موتور ترجمه گوگل با استفاده از پیکره‌های متنی متفاوت، از منابع خبری گرفته تا ویکیپدیا، «آموزش‌دیده» است. یگانه سرخنی که موجب پی بردن به بافت یک پیکره می‌شود، توصیف ساده آن پیکره است. استنباط تناسب یا عدم‌تناسب واژه‌ای مثل «خروار» از چنین اطلاعات اندکی سخت خواهد بود. اگر بنا به ترجمه این واژه به فرانسه باشد، ماشین می‌تواند معادلی مناسب را برای آن در beaucoup یا plusieurs حدس بزند. این معادل‌ها ممکن است معنای گفته را ارائه کنند، اما طنز آن را نه. زیرا این طنز وابسته به واژه «خروار» است که از لحاظ اجتماعی نشانه‌دار است، درحالی‌که plusieurs خنثی است. هرچقدر هم که الگوریتم پیچیده باشد، باید به اطلاعات فراهم‌شده تکیه کند و سرخ‌ها دربار بافت، به‌خصوص بافت اجتماعی،

برای انتقال از طریق گد بسیار مشکل‌اند.

مارسل وقتی .دهد معنا «عبارت» یا «جمله» تواند می انگلیسی به phrase .بگیریم مثال را فرانسه زبان در petite phrase پروست در رمان در جست‌وجوی زمان از دست‌رفته (۱۹۱۳)، در پار [Vinteuil de phrase petite la]، از آن در بافتی موسیقایی استفاده می‌کند، باید «عبارت» ترجمه شود، زیرا «جمله» بی‌معناست. ترجم [گوگل در برگردان آن بسیار خوب عمل می‌کند (منظور نسخ [قدیمی عبارت‌محور آن است؛ نسخ [جدید شبک [عصبی تا الان تنها برای چینی ماندن قابل‌استفاده است). اگر فقط phrase petite را وارد کنید، به شما «جمله کوتاه» را تحویل می‌دهد. اگر Vinteuil de phrase petite la را وارد کنید (وانتوی نام شخصیتی است که از قضا آهنگ‌ساز است)، به شما «عبارت کوتاه وانتوی» را تحویل می‌دهد که پژوهی از ترجمه‌های منتشرشده از اثر پروست است. کمیابی نام «وانتوی» بافت لازم را فراهم می‌آورد و الگوریتم آماری از آن بهره می‌گیرد. اما اگر Sarkozy de phrase petite le را وارد کنید، به جای «متلک سارکوزی» که صحیح است، «عبارت کوتاه سارکوزی» را می‌پراند، زیرا در بافت سیاسی‌ای که نام رئیس‌جمهور سابق آن را می‌نماید phrase petite une گفت [تدوینزی است که به‌سوی رقیبی سیاسی نشانه رفته است: یک متلک، نه عبارت موسیقایی. اما نام سارکوزی در چنان تنوعی از جملات پدیدار می‌شود که موتور آماری از ثبت درست آن در می‌ماند و آنگاه این خطاها را با یک اشتباه اصطلاحی تأسف‌آور همنشین می‌کند.

درست همانند تلاش‌های پیشین برای خلق هوش مصنوعی که سابق [آن به روزهای دانشجویی من در ام.آی.تی برمی‌گردد، مسئله این است که هوش به‌نحوی باورنکردنی پیچیده است. هوشمندبودن صرفاً توانایی استنباط منطقی از قواعد، یا استنباط آماری از قاعده‌مندی‌ها نیست. پیش‌ازان، شخص باید بداند که کدام قواعد کاربردی‌اند، هنری که نیازمند آگاهی‌داشتن از حساسیت به موقعیت است. برنامه‌نویس‌ها بسیار باهوش‌اند، اما هنوز آن‌قدر باهوش نیستند که تنوع گسترده بافت‌های حامل معنا را پیش‌بینی کنند. در نتیجه حتی بهترین الگوریتم‌ها نیز نکاتی را از دست می‌دهند و، همان‌طور که هنری جیمز گفته، مترجم ایدئال کسی است که «هیچ نکته‌ای از دستش در نرود».

نمی‌گویم که ترجم [مکانیکی سودمند نیست. مقدار زیادی از کار ترجم تکراری است. گاهی اوقات ماشین‌ها از عهد [انجام کاری بسنده برمی‌آیند. اما نباید انتظار معجزه، ترجمه‌های ادبی شایسته یا ارائه مناسب متلک‌های سیاسی را داشت. پژوهش‌های مرتبط با هوش مصنوعی، از همان روزهای آغازین، زیر سای [ادعاهای مغرورانه بوده‌اند. این را از ترس شغلم نمی‌گویم: از ترجمه بازنشسته شده‌ام و این روزها بخشی از وقتم را صرف کدنویسی می‌کنم.

پی‌نوشت‌ها:

* این مطلب در تاریخ ۲۱ نوامبر ۲۰۱۶ با عنوان zingers and Shitloads on the translation machine of perils the می‌تواند وب‌سایت ایان منتشر شده است و وب‌سایت ترجمان در تاریخ ۱۱ دی ۱۳۹۵ این مطلب را با عنوان آیا گوگل بالاخره می‌تواند کاروبار مترجمان را کساد کند؟ ترجمه و منتشر کرده است.

* آرتور گلدهر نویسنده، مترجم، پژوهشگر و نیز نویسنده وبلاگی درباره سیاست فرانسه است. وی بیش از ۱۲۰ کتاب را از فرانسه ترجمه کرده که از میان آن‌ها می‌توان به دموکراسی در آمریکا نوشت [توکویل و سرمایه در قرن بیست‌ویکم توماس پیکتی اشاره کرد. گلدهر یکی از اعضای مرکز مطالعات اروپایی هاروارد است و یادداشت‌های او در نشریات چون د امریکن پراسپکت و د نیشن منتشر می‌شوند.

[۱] Shitload

این کلمه کاربرد عامیانه و بار معنایی توهین آمیز دارد و به خاطر همین استفاده اش توسط آن هلندی در آنجا مناسب نبوده است. شده ریزی‌پایه عصبی دستگاه و انسان مغز براساس که کامپیوتری سیستمی [۲] Neural network

[۳] Wetware

[۴] هر دوی این‌ها در زبان فرانسه به‌معنای بسیار هستند.

آن یادگیری حال در که کسانی که است زبان یک شد [پذیرفته قواعد از انحرافی خطا، کاربردی شناسی‌زبان در [۵] Error به‌عنوان زبان دوم هستند مرتکب می‌شوند.

* منبع/ ترجمان

مولف: آرتور گلدهر

مترجم: علی امیری

برچسب‌ها: فضای مجازی، [1]

ادبیات [2]