۱۰. فقه هوش مصنوعی (۱۴۰۲/۰۷/۰۶)

سال تحصیلی (۱۴۰۳-۱۴۰۲) - پنجشنبه، ۶ مهر ماه ۱۴۰۲

- پیشگفتار (چکیده)
- طرح بحث
  - امکان یا عدم امکان استناد ضمان به هوش مصنوعی و محوریت استناد در ضمان
  - ربات هوش مصنوعی عمومی بارزترین مثال برای طرح بحث
- عدم امکان تحقق هوش مصنوعی عمومی خودآگاه در کامپیوترهای سیلیکون پایه
- لزوم آموختن زبان‌های برنامه‌نویسی و نقص ندانستن آن در مسیر طلبگی
- عبور از کامپیوترهای سیلیکونی به زیستی و کوانتومی، و طرح سؤالی در مقدمه کتاب «جادوی هوص»
  - پیشنهاد اسم «خودرایان» برای هوش مصنوعی
  - فقره‌ای از کتاب «فهم اصول اخلاقی و امنیتی حاکم بر هوش مصنوعی»

پیش‌گفتار (چکیده)

طرح بحث استنادپذیری در هوش مصنوعی، تحریر محل نزاع در مسؤولیت‌پذیری هوش مصنوعی، عدم امکان تحقق هوش مصنوعی عمومی خودآگاه در کامپیوترهای سیلیکونی، امکان تحقق هوش مصنوعی رفتارگرا، پیشنهاد اسم «خودرایان» برای هوش مصنوعی

طرح بحث

جلسه اول فقه هوش مصنوعی پنجشنبه ۶ مهرماه ۱۴۰۲ (۱۲ ربیع الاول ۱۴۴۵)

طرح بحث

امکان ای عدم امکان استنادضمان به هوش مصنوعی و محوریت استناد در ضمان

فرمودند بحث فقهیای باشد راجع به مسئولیت پذیری در هوش مصنوعی که امروزه پیش آمده و محل ابتلاء است. صورت ساده مسأله این است: ماشینی هست که خودران است، یعنی راننده ندارد و با هوش مصنوعی اداره میشود. [اگر] کسی را کشت، چه کسی باید خسارت بدهد؟ اتومبیلی که رانندگی آن با هوش مصنوعی اداره میشود، [اگر] کسی را زد و کشت، از نظر فقهی چه کسی باید دیه بدهد؟ این اصل صورت مسأله است.

من مقدماتی را عرض میکنم. متنی را همه در ایتا[1] گذاشتند. سعی کردم آن متن ذهن شریف شما را آماده کند تا اگر خواستید به صورت منظم و کد نویسی، الگوریتم آن را در بیاورید راحتتر باشد. فعلاً بنا نداشتم. برای زمینهسازی بود که عرض کردم.

تقریباً از ضروریات فقه است که ضمان در فقه منوط به قصد نیست. کسی میخواهد ضامن چیزی شود –به معنای وسیعش- محوریتش بر استیلا است. در تلفِ یک چیزی –مال، نفس، عضو و...- چه زمانی ضامن است؟ وقتی مستند به او باشد. متلفِ کسی است که تلف تالف مستند به او است. این استیلا خیلی مهم است. کسی خواب بوده؛ در خواب که چیزی نمیفهمد؛ در خواب پایش را دراز کرده و به لیوانی که مال کسی بوده برخورد کرده که افتاده و شکسته، بگوییم خُب چه تکلیفی دارد؟! خواب بوده! درست است؛ یعنی کسی گناه نکرده است. اما جزء واضحات فقه است که ولو خواب بوده چون شکستن لیوان به او مستند است، ضامن است. یک مالی از بین رفته و تلف شده، چون تلف آن به این آقا مستند است ولو خواب بوده، ضمان ثابت است. پس محوریت ثبوت ضمان بر استیلا تلف به آن متلف است. نه به اینکه قاصد باشد یا هوشیار و آگاه باشد. این خیلی مهم است و بعدها همه مرتب با آن کار داریم. محوریت با استیلا است. البته بعداً قصد به کار میآید. ما با قصد هم حساب داریم؛ عمد و شبه عمد و ... را داریم. اما غیر از اصل مسأله ضمان است. اصل ضمان بر استیلا است.

خُب وقتی استیلا محور شد، ما باید روشی را پیش بگیریم که این استیلا برای ما روشن و واضح شود. بفهمیم که [تلف] مستند به چیست. مثلاً اگر یک اتومبیلی تصادف کرده، ماموری که میآید تا بفهمد علت تصادف چه بوده، اول سراغ این میرود که ببیند چراغش روشن بوده یا نبوده؟! چراغش سالم است یا چراغش سوخته بود؟! چرا نمیرود؟ بهخاطر اینکه راه عقلائی آن این نیست. همین الگوریتمی است که الآن گفتم؛ اول میبیند تصادف در روز بوده یا شب. اگر روز بوده که سفیهانه است او سراغ این برود که ببیند چراغش سالم بوده یا نبوده. برای پیجویی

مستند و اینکه ببینیم رمز وقوع یک حادثه چیست، باید مراحل عقلائی طی شود. اول بگوییم چه زمانی تصادف شده؛ اگر گفتیم شب بوده، سراغ این میرویم که ببینیم چراغش سالم بوده ای نبوده و مستند به خرابی چراغ هست ای نیست. اما اگر گفتند روز بوده باید سراغ چیزهای دیگری برویم. این مسأله استاندارد و روندی که در استاندارد طی میشود خیلی مهم است.

تحریر محل نزاع

در مانحن فیه، به اصطلاح طلبگی خودمان، باید تحریر محل نزاع را به‌صورت دقیق انجام بدهیم. چرا؟ به‌خاطر اینکه در مسأله استاندارد و نوشتن کدهای کامپیوتری و برنامه‌های هوش مصنوعی، آن چه که بزنگاه بحث ما است، این امر رایج نیست. یعنی ما که از هوش مصنوعی بحث میکنیم که مسئولیت‌پذیری در آن به چه صورت است، مصبّ استاندارد در هوش مصنوعی بسیار دقیق است. باید مصبّ آن را تشخیص بدهیم و الا همینطور بصورت کلی نیست.

مثال‌هایی که در ذهنم هست را عرض میکنم؛ مثلاً در خانه‌های کسی بوده که گاهی حواسش به هم میریخته و حالت روانی درستی نداشته، یک دفعه در خانه حالش به هم ریخته و چاقویی برداشته و کسی از اعضاء خانواده‌ای نفس محترم دیگری را کشته. الآن شما میخواهید مقصر را تعیین کنید و میخواهید ببینید کشته شدن این شخص توسط آقایی که گاهی حالش به هم میریخته و به آشپزخانه رفته و چاقو آورده و او را کشته؛ میخواهید ببینید این قتل مستند به چه کسی است. چه کسی مقصر است. او که متوجه نبوده. خُب مباشر هم او بوده.

شاگرد: عاقله چه کسی است؟

استاد: نمیخواهم بگویم طفل ای مجنون بوده تا سراغ عاقله برویم. در آن حال گاهی مشکوک میشود که آیا جنون بوده ای نبوده. در این مثال‌ها خیلی به این امور که عاقله کی بوده و ... کاری ندارید، با مقصودهای اصلی کار دارید. از این عبور میکنیم تا به مقصود اصلی مثال برسیم.

فعلاً عرض من این است: در اینجا شما دنبال این بلند میشوید تا ببینید مقصر چه کسی است. ذهنتان سراغ این میرود و میگویید در این خانه میدانستند که حال او خراب میشود لذا چاقو را در کمدی گذاشته‌اند که کلید داشته. اما میبینید میگویند فلانی آمده قفل کمد را باز کرده و از این طریق چاقو به دست او آمده است. ببینید در اینجا ذهنتان جلو میرود. میگویید چه کسی تقصیر کار است؟ مثلاً کسی که رفته این قفل را باز کرده و چاقویی که پنهان کرده بودند و نمیخواستند به دست او بیافتد در دست او افتاده. ذهن شما در پیجویی روند کار که مقصر را پیدا کند، خیلی از این اباء نمیکند. جای سراغ عاقله میروید که خود مباشر ضامن است و عاقله پرداخت میکند. اما اگر قبل از اینکه او را مقصر بدانیم، اگر مقصری داشته باشیم که از او اقوی است و این شخص نسبت به او کالآلة بوده و مباشر نسبت به او ضعیف‌تر باشد، سراغ او میروید. این بحث‌های فقهی جای خودش باشد. فقط میخواهم مثال بگویم. پس ذهن شما تا اینجا مشکلی نداشت.

حال عرض من این است: [آیا] به ذهن احدی از عالم و جاهل و عرف عام و خاص میآید که بپرسند با چه چیزی او را کشت؟! با یک چاقو کشت. بعد بروید ببینید این چاقو را از کدام مغازه خریده‌اند! ذهن کسی سراغ این میرود؟! خلاصه او فروخته است. بعد ببینید این چاقو را کدام کارخانه تولید کرده تا دیه را از او بگیرید. این به ذهن کسی میآید یا نمیآید؟! نمیآید. اینکه چرا نمیآید را بعداً عرض میکنم. حال بعداً با این مثال کار دارید.

در این برنامه‌هایی که نوشته میشود، زبان‌های برنامه‌نویسی‌ای هست که برنامه‌نویس طبق یک کدی که مینویسد، یک محصولی را ارائه میدهد. شما میتوانید تمام مراحل تقصیر، قصد، خطا را

برای این برنامه‌نویس بیاوری. یعنی یک برنامه‌نویس می‌تواند یک برنامه‌ای بنویسد که قصد او از برنامه این باشد که زید را بکشد. این خیلی امر روشنی است. ما که در اینجا مشکلی نداریم. ولو یک ماشینی باشد که هوش مصنوعی هم اجراء می‌کند، اگر این برنامه‌نویس کار این رابات را طوری تنظیم کرده که به وسیله آن برنامه و مستند به او و با قصد قتل او را کشته است. از نظر فقهی که در اینجا بحثی نداریم. روشن است. قتل به رابات مستند نیست. قتل به برنامه‌نویس مستند است و باید هم قصاص شود. چون قاتل عمد حساب می‌شود. حالا اگر قاصد فعل بود و قاصد نتیجه نبود، شبه عمد می‌شود. اگر نه قاصد فعل و نه نتیجه بود، برنامه‌نویس قاتل می‌شود اما قاتل به خطای محض. چون نه قاصد فعل بوده و نه قاصد نتیجه. در اینجا انواع و اقسامی هست.

آیا بحث ما در مسئولیت‌پذیری هوش مصنوعی اینها است؟ نه. استیناد اینها روشن است. بگوییم خُب هوش مصنوعی دارد اجراء می‌کند! خُب استینادی که سر آن بحث می‌کنیم جا دارد. صرف اینکه یک هوش مصنوعی در یک وسیله‌ای فعال باشد، این بحث ما نیست. بحث‌های روشنی دارد؛ مستندات روشنی دارد؛ جواب های فقهی آن هم روشن است. پس کجا است که مبهم است و دقیق است که ما پیرامون آن بحث می‌کنیم؟ مسأله خاص هوش مصنوعی است. یعنی همان مسأله‌ای است که امروزه می‌گوییم یادگیری ماشین.

در همان مثال که گفتم کسی سراغ مغازه ای کارخانه چاقو نمی‌رود، شبیه آن مثال در اینجا موجود است. در مسئولیت هوش مصنوعی از آن جهتی که به خود ماشین، بدون اینکه برنامه‌نویس خبر داشته باشد، بدون اینکه قصد کرده باشد، خود ماشین در روند یادگیری خودش چیزهایی را یاد می‌گیرد که حتی خود برنامه‌نویس از آن‌ها خبر نداشته است و باورش نمی‌شده که این هوشی که برای این رابات به کار برده‌ام این کار را می‌کند.

---

[1] مربوط به جلسه پنجشنبه (فقه هوش مصنوعی) (آمادگی پیش از جلسه):

۱- مثلث [متلف] [تالف] [تالفی]، تقسیمات گسترده دارد که در الگوریتم ضمان دخالت دارند.

۲- [متلف] آن است که تلف {استناد} به او دارد.

۳- [متلف] می‌تواند (مالک) یا (ممولک) و یا (نه مالک و نه ممولک) باشد.

۴- [متلف] می‌تواند با ملاحظه قصد او نسبت به (کار) و (نتیجه)، به سه قسم {عمد} و {شبه عمد} و {خطأ محض} تقسیم شود.

۵- [تالف] می‌تواند (نفس) یا (عضو) یا (عرض) یا (مال) باشد.

۶- [تالفی] می‌تواند (جریمه مالی) یا (کیفری) یا (رقبه) یا (هدر) باشد.

۷- {استناد} می‌تواند از نوع (مباشرت) یا (تسبیب) باشد که {اقوائیت} ملاحظه می‌شود.

۸- {استناد} می‌تواند مشتمل بر مؤلفه‌های متعدد باشد که به تساوی یا تفاضل بر آن‌ها {توزیع} می‌شود.

۹- {استناد} می‌تواند طولي باشد که {قرار ضمان} بر یکی از حلقه‌های زنجیره است.

۱۰- {قرار ضمان} می‌تواند بر نزدیک‌ترین محل [تلافی] باشد، هر چند تعاقب ادی ای اسباب طولیه، سبب رجوع شود.

ربات هوش مصنوعی عمومی بارزترین مثال برای طرح بحث

واضحترین مثالی که میخواهم عرض کنم که تا آخر مباحثه این مثال در ذهنتان بماند تا روی آن رفتوبرگشت کنیم، این است: فرض میگیریم یک رباتی هست که دارای هوش عمومی است. البته هنوز هوش عمومی بهصورت کامل پیادہ نشدہ ولی ما فرض میگیریم. این مثالی که الآن میخواهم عرض کنم تا سرلوحه مباحثه ما باشد، بهترین مثال است از نظر فناوری و علم، و بدترین و سختترین مثال است از حیث لوازم حقوقی.

عدم امکان تحقق هوش مصنوعی عمومی خودآگاه در کامپیوترهای سیلیکون پایه

ربات‌ی را فرض بگیرید که با هوش مصنوعی عمومی (General Artificial intelligence) کار می‌کند. البته مباحث فلسفی آن بماند. قبلاً عرض کرده‌ام و بعداً هم ان شاء الله بیشتر صحبت می‌کنیم؛ آیا هوش عمومی مصنوعی ممکن هست یا نه؟ جلوترها عرض کرده بودم که دو نگاه است: نگاه فلسفی به هوش عمومی؛ قبلاً مباحثه کرده‌ایم و با ضمانت طلبگی عرض می‌کنم که نسبت به مباحث فلسفی، بشر نمی‌تواند با این کامپیوترهای دیجیتال به آن برسد. حالا در کامپیوترهایی که بعداً با کوانتوم و با چیزهای دیگر می‌آید که چند فرض هم دارد، فعلاً با آن‌ها کار ندارم. فعلاً با این عصر دیجیتال و سیلیکون پایه، واضح و آشکار است که در سؤال فلسفی، بشر نمی‌تواند با این کامپیوترهایی که روی آن نرم‌افزار اجراء می‌شود به آن برسد. یعنی یک زبان سطح بالا دارید که چراهای ندارید آن زبان را به زبان سطح پایین ماشین ببرید و بعد پردازش کنید. غیر از این چراهای ندارم. در این‌چنین فضایی هرگز هوش مصنوعی عمومی اتفاق نمی‌افتد. اتاق چینی سرل (Chinese Room) اگر درست تقریر شود، مطلب درستی است. او تعبیری به هوش قوی می‌کرد. اتاق چینی را که سرل (Searle)، برای هوش قوی مثال می‌زند، می‌بینید برای آن جوابی ندارند و نخواهند داشت. البته به آن معانی که منظور او بود؛ یعنی «آگاهی». عده‌ای با حرف او مقابله کردند و خواستند به «آگاهی» هم برسند. عرض من این است که روی نسل فعلی از کامپیوترهای دیجیتالی ممکن نیست. قبلاً بحث‌های علمی آن را از چند جهت انجام داده‌ایم. قبل از کرونا هم مقاله فرگه را که بحث کرده‌ایم، دقیقاً این مطالب را ثابت می‌کند. همچنین مسائل افلاطون گرایی در آنالیز.

جلوتر عرض کرده‌ام در وضوح مسائل متافیزیکی نزد کل بشر، امروزه خلأ تحقیقی نداریم، بلکه خلأ تبیینی داریم. اگر هم عقب بیافتد به گردن ما است. یعنی خلأ تبیینی این است که زمینه‌اش فراهم است؛ آنالیز ریاضی که در زمان ما پیشرفت بالایی دارد، تنها نیاز به این است که شما از آن استفاده کنید و تبیینی کنید. و لذا با مسائل تبیینی که قبلاً عرض کرده بودم؛ با مسائلی که فرگه در [مقاله] اندیشه می‌گفت و چند مطلب دیگر...؛ اگر یک ماشین در عصر دیجیتال تست تورینگ را هم پشت سر بگذارد و هیچ مشکلی هم نداشته باشد، (هیچ مشکلی نیست و دلتان از این‌جهت جمع باشد که لا یقف است؛ نمی‌دانم چرا دغدغه می‌کنند؟! در این‌که یک ماشینی بخواهد تست تورینگ را برای رفتارگرائی پشت سر بگذارد، ذهن من اصلاً مشکلی ندارد و حد یقف ندارد. هوش عمومی به‌معنای رفتار گرائی به‌راحتی می‌تواند. گمان من این است که

می‌شود و اصلاً مبهم نیست.) ولی صحبت سر این است که هوش عمومی به‌معنای خودآگاهی و ذهن باشد؛ این‌ها را نمی‌توان. تا زمانی‌که عصر دیجیتال است، نمی‌شود. آن‌هایی هم که در عصرهای بعدی می‌آیند -زیست رایانه و کوانتوم رایانه- زمانی‌که بیت‌ها به کیوبیت (Qubit) تبدیل شود...؛ امروز زمان بیت است و در آینده کامپیوترها کیوبیت می‌شوند؛ کوانتوم بیت (Quantum bit). آن زمان دو-سه گزینه دارد؛ اینکه آیا ممکن است ای نه، عرض می‌کنم.

فعلاً در محل نزاع خودمان فرض می‌گیریم که یک ربات است با هوش عمومی رفتارگرایانه، یعنی به نعبیری کاملاً تست تورینگ را پشت سر می‌گذارد. تست تورینگ یعنی اگر دو نفر باشند که یکی از آن‌ها ربات باشد و دیگری انسان، و اصلاً معلوم نباشند، انسان‌های متعددی هم باشند نمی فهمند کدام یک از این‌ها انسان است و کدام یک از آن‌ها ماشین است. از نظر رفتارگرائی هم یک امر ممکنی است. قطع نظر از مباحث فلسفی.

شاگرد: هنوز نرسیده‌اند.

استاد: نرسیده‌اند. اما چرا می‌گوئی نمی‌رسند؟!

لزوم آموختن زبان‌های برنامه‌نویسی و نقص ندانستن آن در مسیر طلبگی

من به‌صورت طلبگی ذهنم را عرض می‌کنم. ممکن است بگویند تو اشتباه می‌کنی. اما ذهنیت من این است. جلوتر هم مکرر عرض کرده‌ام. وقتی ما المنطق مباحثه می‌کردیم، حیثیاتی که سر از هزاران حیثیت در می‌آورد ما را مأیوس می‌کرد. بعد از این‌که با زبان‌های برنامه‌نویسی آشنا شدم، نفس راحتی کشیدم. شما هم که در سن جوانی هستید! یک زمانی بود که اگر کسی این زبان‌ها را نمی‌دانست کمالی را نداشت، اما وقتی جلو می‌روید اگر این زبان‌ها را ندانید کم‌بود دارید. حدود ده سال پیش به یکی از اعزه عرض کردم؛ پاراسال به من گفت تو قبلاً به من گفتی اما حالا دارم می‌بینم که چرا به حرف تو گوش ندادم. زیاد هم لازم نیست وقت بگذارید. روزی ده دقیقه ای پنج دقیقه برای آن وقت بگذارید، بعد از دو- سه سال می‌بینید انواع این زبان‌ها را می‌دانید و ذهن شما با آن مانوس است و دستتان به کد می‌رود. این‌ها چیزهایی است که نیاز است. این‌طور نیست که بخواهید همین‌طور فقهی بحث کنید و از نزدیک چیزی را نبینید.

چنین رباتی را در نظر بگیرید که تست تورینگ را پشت سر گذاشته است؛ از نظر رفتارگرائی واجد هوش عمومی است. یعنی فرض می‌گیریم یک برنامه‌نویس کارهای آن ربات را انجام داده و آن را رها کرده و الآن یادگیری ماشین به‌خاطر این هوش عمومی در حد بالا است. اَبَر هوش هم اگر بعداً بیاید جای خودش. فعلاً فرض مثال ما این است.

خُب مهندسی که این ربات را تهیه کرده، از نظر فنی چون هوش عمومی دارد، هیچ مشکلی نیست که بعداً خود این ربات سرقت مسلحانه را یاد بگیرد. سرقت مسلحانه‌ای را یاد می‌گیرد که خود مهندسی که برنامه را نوشته اولاً بلد نیست؛ ثانیاً اگر بفهمد قرار است ربات او بعداً سرقت مسلحانه بکند هرگز این کار را نمی‌کند. ولی خُب خبر ندارد. او می‌خواهد رباتی باشد که کاری را انجام بدهد. بالاترین فرض را داریم می‌گوییم تا محل نزاع به‌خوبی روشن شود.

حال این ربات آمد و خودش سرقت مسلحانه را یاد گرفت. بعداً هم اتومبیلی را گرفت و راننده اش را کشت و بعد از قتل هم اتومبیل را برد. الآن محل نزاع ما چنین چیزی است. مثالی را گفتم که مقصود من را برساند. الآن که این ربات این کار را کرد، کشتن راننده با الگوریتم سرقتی که یاد گرفته و خودش به دست آورده، مستند به چه کسی است؟

شاگرد: احتمالش را هم نمی‌داد که ممکن است یاد بگیرد؟

استاد: احتمال نمی‌داد. مثل چاقویی در آشپزخانه. کارخانه آن چاقو را نساخته بود تا با آن سر ببرد. مغازه‌دار هم آن را نفروخته بود تا او سر کسی را ببرد. اصلاً این را احتمال نمی‌دادند.

شاگرد۲: مستند به صاحب آن است. می‌گوید آن را فروخته ای واگذار شده.

استاد: من حرفی در این جواب‌ها ندارم. اما می‌خواهم دقیقاً محل نزاع روشن شود. منظور من نقص و خرابی مثال نیست. منظور من این است که به جایی می‌رسیم که هوش مصنوعی آن راول، اقدام و ترفند کار را خودش به دست آورده است. به خودش استناد دارد. این‌جا محل نزاع ما است. یعنی الآن در فضای فقه به جایی رسیده‌ایم که خود هوش مصنوعی این کار را کرده است. شما نمی‌توانید بگویید چون فلانی برنامه آن را نوشته باید هم او را گرفت. این اصلاً عقلائی نیست. مثل اینکه عقلائی نیست که به مغازه‌دار بگویید چرا چاقو فروختی؟! ای سارق کارخانه را بگیرید که چرا چاقو ساختی که شخصی با آن آدم کشت؟!

شاگرد: شاید در یک پله دیگری هم بتوان برای استنادش بحث کرد که علم هنوز به جایی نرسیده که بتواند این احتمال را تأیید بکند. و الا اگر علم به جایی برسد که بگوید این چیزی که شما می‌سازید ممکن است این کار را بکند –ولو شخصش احتمال را نمی‌داده- باز بحث استناد مطرح می‌شود.

استاد: فعلاً من به مقصود خودم رسیده‌ام. اما اینکه شما بگویید مستند می‌شود یا نمی‌شود من تنها محل نزاع را گفته‌ام. شما دارید آن را حل می‌کنید. می‌گویید چون احتمالش می‌رود به او هم مستند می‌شود. ولی علی ای حال خیلی صاف نیست. مخصوصاً با بحث‌هایی که بعداً می‌شود. این‌طور نیست به صرف احتمالی که می‌رود مستند کنیم. مثلاً کارخانه‌ای که چاقوی تیز به دست مردم می‌دهد، با آن نمی‌توانند سر ببرند؟! ای صاحب کارخانه تو که این همه چاقوی تیز می‌سازی و به داخل جمعیت می‌دهی، احتمال عقلائی نمی‌دهی که با این چاقو سر کسی را ببرند؟!

شاگرد۲: عقلاء بین این دو مثال تفاوتی نمی‌گذارند؟!

استاد: نه، اگر شما رونده هوش مصنوعی را بدانید می‌بینید که تفاوتی نیست. وقتی راول پیشرفت هوش مصنوعی را بدانید، می‌بینید همین کارخانه است. ولی من به مقصود اولی خودم رسیدم.

شاگرد: اگر کسی سگس را همراه خودش دارد که این‌قدر به او آموزش داده که برای او معلوم است که این سگ اهلی است. ولی یک دفعه پاچه می‌گیرد، استناد این سگ به چه شکل است؟ عرض من این است که باز این استناد دارد. یعنی هنوز احتمال استنادش هست. در آن جایی که به‌صورت عقلائی این احتمال هست که سگی که دست پرورده ی شما است ممکن هست که باز در برود. یعنی درصد خطای آن به قدری کم نشده که احتمالش غیر عقلائی شود.

استاد: من حرفی ندارم و الآن نخواستم به مسأله جواب بدهم. وجهی که شما می‌گویید تقریری برای پیشرفت بحث است. ما باید یک مقدماتی را طی کنیم تا به این فرمایش شما برسیم که ببینیم الآن این استناد صحیح است یا نه. این مثال سگ و ... در فقه هست. شما از حیوان مواظبت بکنید یا نکنید. مواردی هست که اگر از حیوان کاری سر بزند مالک آن ضامن نیست. بحمدالله فقه بسیار غنی است. یعنی حتی این محل نزاعی که برای هوش مصنوعی عرض می‌کنم مطالب مبنایی آن تمام در فقه هست. قواعد ضمان با آن ده بندی که در ایتا گذاشته‌ام را بررسی می‌کنیم.

علی ای حال آن چه که مقصود من بود را خدمت شما عرض کردم. یعنی پس ما که بحث فقهی میکنیم که مسئولیتپذیری هوش مصنوعی از نظر فقهی چیست، نمیخواهیم جایی را بگوییم که استاندارد به برنامهنویس و ... روشن است. دقیقاً میخواهیم جایی را بگوییم که خود ماشین ای گرفته و خود هوش مصنوعی تر فند کار را ای گرفته و الآن کسی را کشته است. عمد و خطای آن را باید بعداً بررسی کنیم.

شاگرد۲: بحث یادگیری را فرمودید. به تبع آن بحث کلان داده هم مطرح میشود که برای یادگیری در اختیار آن قرار داده میشود. آیا عناصر دیگری مانند افرادی که کلان داده را آماده میکنند و در اختیار او قرار میدهند هم مؤثر است؟

استاد: یکی از بندهایی که در این ده بندی که عرض شد، آمده است، یکی این است که استاندارد نقطهای نیست. استاندارد توزیع پذیر است. یعنی شما یک جدولی را از تمام چیزهایی که میتوانند داخل باشند، ارائه میدهید؛ البته با منطق دو ارزشی و صفر و یکی نمیگوییدا این مستند به این است، بلکه شما بهصورت منطق فازی و درصدی میگوییدا بیست درصد مستند به این است؛ سی درصد مستند به آن است. پس این هم خودش یک اصل است که مسأله استاندارد، توزیع پذیر است.

بنابراین محل نزاع در ذهن شریفتان معلوم شده. آن چه را هم که شما میفرمایید، کسی که کلان داده (Big Data) را تهیه میکند، هر چه راول بیشتر میشود بخش عظیمی از آن را بعداً خود هوش مصنوعی تولید میکند. اینکه میگویند به تکینگی هم میرسد ای نه...؛ بهخاطر همین است که بعد میگویند کنترل آن از دست بشر در میرود و دیگر هوش مصنوعی نیست و ابَر هوش است؛ (super Artificial Intelligence).

شاگرد: یک وقتی هست که از هوش مصنوعی خطایی سر نمیزند ولی چون در جامعهای تطبیق شده که هنوز کاملاً جا نیافتاده، در تطبیق آن در فضای غیر هوش مصنوعی مشکل پیش میآید. مثلاً در فضای راهنمایی و رانندگی با وجود چراغ گاهی افسر پلیس هم هست، اما طراحی هوش مصنوعی طوری تنظیم نشده که وقتی افسر هست به حرف او گوش بدهی و تنها به آن چراغ توجه میکند، لذا به فرمان افسر گوش نمیدهد و با دیگری تصادف میکند.

استاد: اگر منظورتان را درست فهمیدهام، یعنی گاهی است که خطا مستند به این نیست که خود هوش کاری را کشف کرده باشد و طبق آن اقدامی کرده باشد، بلکه مستند به کمبود و نقص و عدم پیشرفت در الگوریتم هایی که او باید اعمال کند، این هم مانعی ندارد. بعداً ما به همه اینها میرسیم. یعنی وقتی جدول استاندارد را دقیقاً بسط میدهیم، ذره به ذره چیزهایی که میتواند حتی بهعنوان درصد و زیر یک درصد داخل باشد در اینکه این خطا رخ داده، همه را بررسی میکنیم.

ولی آن چه که مهم است برای این است که بحث راه بیافتد. آقایان هم که فرمودند خدمتشان گفتم: ماشینهای جیپ که اول بود، استارت نداشت. هنوز استارت نیامده بود. آنها را با هندل روشن میکردند. کسی که هندل میزد لازم نبود راننده باشد. اصلاً هم رانندگی بلد نبود. ولی خُب میتوانست هندل بزند. هندل هم برای این بود که موتور روشن شود و راه بیافتد. حالا هم به اندازه هندل زدن که راه بیافتد، وقتی من طلبه میبینم میتوانم هندل بزنم تا راه بیافتد...؛ وقتی هم راه افتاد برعهده جوانان و استادی است و اگر اشکالی در مباحث هست، از این ناحیه ببینید.

عبور از کامپیوترهای
سیلیکونی به زیستی و
کوانتومی، و طرح سؤالی در
مقدمه کتاب «جادوی هومص»

پیشنهاد اسم «خودرایان» برای هوش مصنوعی

چند کتاب به من داده بودند؛ می‌خواستم چند سطری از ابتدای آن‌ها را بخوانم. یکی از آن‌ها اولین مورد از انتشارات مجموعه کتب تخصصی هوش مصنوعی از حوزه علمیه است؛ «**فهم اصول اخلاقی و امنیتی حاکم بر هوش مصنوعی**» که دوسطر از صفحه اول و آخر آن را خدمت شما می‌خوانم. کتاب دیگری هم به من دادند به نام «**جادوی هومَص**». همه میل دارند کلمات را مختصر کنند. وقتی هم کلمه‌ای زیاد تکرار می‌شود سخت است آن را مدام تکرار کنند؛ لذا به جای هوش مصنوعی می‌گویند هوص. من که این را دیدم به ذهنم آمد بگویند هوشوارە.مانند ماهواره. بعد به ذهنم اینطور آمد: الآن رایانه جا گرفته است. رایانیدن هم به‌معنای مرتب کردن و منظم کردن است. در فارسی به هوش مصنوعی می‌گوییم خودرایان. خودرایان واژه خوبی است. خودروی‌ِخودران‌ِخودرایان. خودرو است یعنی اتوموبیل است در مقابل دوچرخه. خودران است یعنی راننده ندارد و توسط هوش مصنوعی اداره می‌شود. خودرایان است، یعنی الآن آن چیزی که مسئولیت اداره آن را دارد، صرفاً راننده نداشتن نیست، بلکه حال دیگر خودش رفتار خودش را تنظیم می‌کند. رایان به‌معنای منظم کردم؛ به‌دست‌آوردن و فهمیدن الگوریتم اقدام. نمی‌دانم این را کسی گفته یا نگفته و اصلاً به درد می‌خورد یا نه. مباحثه طلبگی است.

این کتاب برای استادی است که شاید چند سالی از من بزرگتر باشند. علی ای حال برای من جالب بود. ایشان می‌گویند:

> در سن دبستان که بودم تهران با دوران سلجوقیان هفتصد سال پیش تفاوت چندانی نداشت. هنوز برق به خانه ما نرسیده بود. هنوز در تهران برای رفتوآمد از کالسکه و درشکه استفاده می‌کردیم. هنوز شترها در جمعه بار می‌آوردند. هر روز باید گوشت می‌خریدیم. حمام در منزل معنا نداشت. برای رفتوآمد شبانه باید چراغ نفتی به همراه می‌بردیم.

«در سن دبستان که بودم تهران با دوران سلجوقیان هفتصد سال پیش تفاوت چندانی نداشت»؛ ایشان این‌ها را می‌گویند، من هم یادم هست. من هم در یزد بودم.

«هنوز برق به خانه ما نرسیده بود»؛ بله برق نبود؛ حتی در یزد کارخانه‌های بود که سر شب، ساعت شش-هفت روشن می‌شد و دوباره ایزده و دوازده آن را خاموش می‌کردند. در زمانی‌که تازه برق آمده بود. شبکه برق که نبود.

«هر روز باید گوشت میخریدیم»! همینطور بود. یخچال نبود. یخچال هایی هم که تازه آمده بود، نفتی بود. بچه بودیم میرفتیم نگاه میکردیم و مثل یک شمعی در آن روشن بود. با آتش و شمع روشن یخ میکرد. درش را که باز میکردی داخل آن یخ زده بود. میرفتیدی که ببینید چرا یخ زده، میدیدید که شمع روشن است! اینها یخچال های نفتی بود. باید به آن نفت میکردیم. نفتش که تمام میشد دوباره خاموش میماند.

«برای رفتوآمد شبانه باید چراغ نفتی به همراه میبردیم»! خدا رحمت کند آن آقا را! دو سه سالشان بود. میگفت سی سال اخیر عمر من یک طرف، شصت سال قبلی یک طرف. ایشان میگفت من که شانزده سالم بودم اولین بار دوچرخه به دست آمد. آن وقت توضیح میداد و خیلی جالب بود. میگفت چطور آمده بودند. من وقتتان را نمیگیرم. بله، این پیشرفتهای سریعی است که صورت گرفته است.

از خاطراتی که من دارم و برای خود من جالب است، این است: سن من خیلی کم بود. بعد که جلوتر آمدیم، ده سالم که بودیم تازه تلویزیون به دست آمد. کسی تلویزیون ندیده بود. فرستنده آن را گذاشتند و آمد. رواجی هم بود؛ قبلش بین متدینین معروف بود. در بحارالانوار هم هست. جلد شصت اسلامیه. من بچه بودم که میگفتند؛ میگفتند حضرت فرمودند اهل قم برای زمین حجت میشوند و مستضعف نمیمانند. آن چه که منظور من است و خاطره است، این است: حضرت فرمودند «حتی لایبقی مستضعف فی الارض حتی المخدرات فی الحجال»[1]؛ یعنی زنهای کنج خانه. خاطره من این است: بعد از اینکه تلویزیون به دست آمد، این روایت مرتب تکرار میشد. یعنی اینکه گفتند مخدرات هم آگاه میشوند بهخاطر این تلویزیون است. مثلاً فرستنده را در حوزه میگذارند و همه از آن استفاده میکنند حتی المخدرات فی الحجال. این برای حدود سال پنجاه بود. حدود سال پنجاه یادم است. این خاطره مرتب در ذهنم تکرار میشد. در جلد ۵۷ بحار که در نرمافزارها هست، باب «المدوح من البلدان و المذموم منها و غرائبها» این روایت را ذکر میکنند. روایاتی که مرحوم مجلسی از تاریخ قم آوردهاند خیلی جالب است.

در روایت هست که امام صادق علیهالسلام سؤال کردند اهل کجایید؟ عرض کردند: «نحن من اهل الری»[2]. فرمودند: **«مرحبا باخوانان من اهل قم»**. دو-سه بار تکرار شد. تا اینجا که فرمودند: **«ستدفن فیها امرأة من أولاد موسی تسمّی فاطمة فمن زارها وجبت له الجنة»**. وقتی منزل رفتید دنبالهٔ روایت را نگاه کنید: میگوید وقتی امام صادق علیهالسلام این را فرمودند، هنوز امام کاظم علیهالسلام متولد نشده بودند. هنوز پدر حضرت معصومه متولد نشده بودند، حضرت فرمودند: **«مرحبا باخوانان من اهل قم»**.

منظور اینکه در این روایت دارد؛ فرمودند علم از نجف میرود و در قم ظهور میکند، **«تصیر معدن للعلم و الفضل حتی لایبقی فی الارض مستضعف فی الدین حتی المخدرات فی الحجال»**؛ حتی زنهای کنج خانه هم دیگر مستضعف نیستند.

منظور اینکه ایشان که میگویند به این صورت بوده، درست است. و با این فاصله کوتاه این پیشرفتها حاصل شده است. بعد از ناصر خسرو میگویند و ... . البته در ادامه میگویند (که البته نمیدانم منظور چه بوده):

> هوموص کلمه جاودی روز جهان است. مترادف با آیندبی بدیل برای انسانها، ما انسانها موجودات کربن پایه، موجوداتی هوشمندتر از خود آفریدههایی که

> سیلیکون پایه اند.

درحالیکه سیلیکون پایه‌ها معلوم نیست هوشمندتر از انسان باشند. اولاً که آینده سیلیکون پایه نیست؛ عصر دیجیتال تمام شده است. الآن از عصر برق گذشته است. در سه-چهار انقلاب صنعتی ما مرتب داریم از برق استفاده می‌کنیم، اما عصر برق گذشته است. عصر دیجیتال هم تمام شد. یعنی کسانی که می‌دانند، می‌گویند الآن سیلیکون پایه‌ها دارند می‌روند. الآن دیگر **DNA** هست؛ به جای ترانزیستور، ترانسکریپتوم (Transcriptome) **RNA** آمده. ترانسکریپتوم چیزی است که بشر می‌خواهد بعداً از آن برای کامپیوترهای آینده استفاده کند.

چیز عجیبی دیدم. عجائب است. نوشته بود یک گرم DNA قابلیت ذخیره اطلاعات یک تریلیون سی دی را دارد. یک تریلیون سی دی چیز کمی نیست! وقتی رایانه‌هایی زیستی که بعداً می‌آید و می‌خواهند با DNA ذخیره اطلاعات شود، سیلیکون پایه نیستند. آنها طور دیگری هستند. کامپیوتر کوانتومی هم که باز یک فضای دیگری دارد. البته شاید من سر در نمی‌آورم؛ یعنی ایشان مقصود دیگری دارد؛ ما طلبگی می‌گوییم که سر نخی برای ادامه باشد.

شاگرد: شاید به اعتبار آینده نوشته است.

استاد: سیلیکون پایه برای کامپیوترهای دیجیتالی است. وقتی پایه و مبناء بر ترانسکریپتوم شد؛ وقتی مبناء بر کیوبیت شد، دیگر سیلیکون چه کاره است؟! ممکن است در آن زمان هم به وفور از سیلیکون استفاده کنند –مثل برقی که الآن ما استفاده می‌کنیم- اما عصری نیست که آنها پایه باشند.

---

[1] بحار الأنوار ج ۵۷ ص ۲۱۳؛ «عَنِ الصَّادِقِ عَلَيْهِ السَّلَامُ: أَنَّهُ ذُكِرَ كُوفَةُ وَ قَالَ سَتَخْلُو كُوفَةُ مِنَ الْمُؤْمِنِينَ وَ يأرز [أَيْ أَرْزٌ] عَنْهَا الْعِلْمُ كَمَا تأرز [أَرْزٌ] الْحَيَّةُ فِي جُحْرِهَا ثُمَّ يَظْهَرُ الْعِلْمُ بِبَلْدَةٍ يُقَالُ لَهَا الدِّينِ حَتَّى لَا يَبْقَى فِي الْأَرْضِ مُسْتَضْعَفٌ فِي الدِّينِ حَتَّى الْمَخَدَّرَاتِ فِي الْحِجَالِ وَ ذَلِكَ عِنْدَ قُرْبِ ظُهُورِ قَائِمِنَا فَيَجْعَلُ اللَّهُ مُقَامَ أَهْلِ قَائِمِينَ مَقَامَ الْحُجَّةِ وَ لَوْ لَا ذَلِكَ لَسَاخَتِ الْأَرْضُ بِأَهْلِهَا وَ لَمْ يَبْقَ فِي الْأَرْضِ حُجَّةٌ فَيُفِيضُ الْعِلْمُ مِنْهُ إِلَى سَائِرِ الْبِلَادِ فِي الْمَشْرِقِ وَ الْمَغْرِبِ فَيَتِمُّ حُجَّةُ اللَّهِ عَلَى الْخَلْقِ حَتَّى لَا يَبْقَى أَحَدٌ عَلَى الْأَرْضِ لَمْ يَبْلُغْ إِلَيْهِ الدِّينُ وَ الْعِلْمُ ثُمَّ يَظْهَرُ الْقَائِمُ عَلَيْهِ السَّلَامُ وَ يَسِيرُ سَبَاً لنَقِمَةِ اللَّهِ وَ سَخَطِهِ عَلَى الْعِبَادِ لِأَنَّ اللَّهَ لَا يَنْتَقِمُ مِنَ الْعِبَادِ إِلَّا بَعْدَ إِنْكَارِهِمْ حُجَّةً».

[2] بحار الأنوار الجامعة لدرر أخبار الأئمة الأطهار علیهم السلام , جلد ۵۷ , صفحه ۲۱۶؛ «وَ رُوِيَ عَنْ أَبَانٍ الْأَحْمَرِ قَالَ: سَمِعْتُ أَبَا عَبْدِ اللَّهِ عَلَيْهِ السَّلَامُ وَ قَالُوا نَحْنُ مِنْ أَهْلِ الرِّيِّ فَقَالَ مَرْحَباً بِكُمْ أَهْلَ الرِّيِّ أَنْتُمْ مِنْ أَهْلِ الدَّيْلَمِ قَالُوا نَعَمْ فَقَالَ أَنْتُمْ مِنْ أَهْلِ الْكَرَامَةِ قَالُوا نَحْنُ مِنْ أَهْلِ الرِّيِّ فَأَعَادَ الْكَلَامَ قَالُوا ذَلِكَ مِرَاراً وَ أَجَابَهُمْ بِمِثْلِ مَا أَجَابَ بِهِ أَوَّلًا فَقَالُوا نَحْنُ مِنْ أَهْلِ قُمٍّ فَقَالَ مِنْ أَهْلِ الْمَكْرُمَاتِ وَ أَهْلِ الْجَنَّةِ إِنَّ لِأَهْلِ قُمٍّ عَلَى اللَّهِ حَرَماً وَ هُوَ مَكَّةُ وَ إِنَّ لِلرَّسُولِ حَرَماً وَ هُوَ الْمَدِينَةُ وَ إِنَّ لِأَمِيرِ الْمُؤْمِنِينَ حَرَماً وَ هُوَ الْكُوفَةُ وَ إِنَّ لَنَا حَرَماً وَ هُوَ بَلْدَةٌ مُقَدَّسَةٌ تُسَمَّى طَامَةَ فَمَنْ زَارَهَا وَجَبَتْ لَهُ الْجَنَّةُ قَالَ الرَّاوِي وَ كَانَ هَذَا الْكَلَامُ مِنْهُ قَبْلَ أَنْ يُولَدَ الْكَاظِمُ عَلَيْهِ السَّلَامُ».

فقره‌ای از کتاب «فهم اصول اخلاقی و امنیتی حاکم بر هوش مصنوعی»

عبارتی را از این کتاب می‌خوانم؛ یک سطر از صفحه اول و یک سطر از صفحه آخر می‌خوانم.

> در بره زمانی شگفت‌انگیزی به سر می‌بریم که براساس آن بشریت وعده‌های متعددی پیرامون همگرائی روز افزون در دسترسی به کلان داده، سرعت بالا و توسعه پلتفرم محاسبات ابری، و پیشرفت فزاینده در الگوریتم های یادگیری ماشین پیشرفته را طلیعه بخشیده است.

ببینید این‌ها مطالب مهمی است. الآن شما کنار رفیقتان نشسته‌اید؛ می‌گویید شماره تلفن شما چند است؟ فوری آن را در موبایلتان ثبت می‌کنید. الآن شما آن را به بیگ دیتا اضافه کرده‌اید. یعنی دقیقاً معلوم شد که رفیق فلانی هستید، شماره تلفن او معلوم، شماره تلفن شما هم معلوم، و ارتباط شما هم معلوم است. این خودش یکی از چیزهایی است که در بیگ دیتا موجود است. این بیگ دیتا دم و دستگاهی است. «و توسعه پلتفورم محاسبات ابری»؛ محاسبات ابری‌ای که صورت گرفته و می‌گیرد و الآن هم به دنبالش محاسبات اِج (Edge computing) می‌آید.

«و پیشرفت فزاینده در الگوریتم‌های یادگیری ماشین پیشرفته را طلیعه بخشیده است»؛ شما باید روی یک چیزهایی همت کنید و در اطراف آن فکر کنید. یکی از چیزهایی که در این فضا بسیار مهم است، همین الگوریتم است. باید درک درستی از الگوریتم داشته باشید که چیست. درک درستی داشته باشید که بعداً علی بصیرة جلو بروید. در حاشیه رئوس ثمانیة بود. این‌ها خیلی مهم هستند. در کتب قدیمی ما هستند. تفتازانی می‌گوید و مرحوم ملاعبدالله آن را توضیح داده‌اند. علماء رئوس ثمانیة را در ابتدا می‌آوردند. رئوس ثمانیة از مطالبی است که ما در اینجا هم به آن نیاز داریم.

الآن در کل دنیا می‌گویند الگوریتم از چه گرفته شده؟ اسم خوارزمی است؛ الخوارزمی. جالب هم این است که ما فارسی‌ها «ال» عربی را می‌اندازیم ولی آن‌ها «ال» را حفظ می‌کنند و در لغتشان هست. الآن به علم جبر چه می‌گویند؟ الجبرا (algebra). «ال» آن هست. به الکل چه می‌گویند؟ الکحول (alcohol).

همین کلمه الگوریتم. کلمه «الجزری» که نام پدر علم ربات‌یک است؛ ابن رزاز جزری. در زمان ناصر بالله عباسی بوده. کتابی به نام «الجامع بین العلم والعمل فی معرفة الحیل الهندسیة» نوشته که الآن در کل رفرنس‌های جهانی می‌گویند پدر علم ربات‌یک بدیع الزمان جزری است که آن‌ها می‌گویند الجزری. شرح حال او هم معروف است. همه قبول دارند. یک کتاب نوشته که در آن پنجاه اختراع مهم در آن قرار دارد. آن هم در زمان ابن ادریس!

شاگرد: شاید یک علت آن این است که آن‌ها نمی‌دانند که «ال» این اسامی زائد است.

استاد: نه، من نکته دیگری عرض می‌کنم. این‌ها یک نحو اتمام حجتی است که خدای متعال مدیریت می‌کند و باقی می‌گذارد؛ در زبان آن‌ها که زبان بین‌المللی می‌شود باید این «ال» آن باشد تا به دنبالش بلند شوند. لذا این «ال» نیازی است که به‌دنبال این «ال» بلند شوند.

صفحه آخر این کتاب را هم بخوانم؛ نتیجه‌گیری در صفحه آخر:

> در سال 1936 یک ریاضی دان بیست و سه‌ساله از شهر میدوایل به نام آلن تورینگ با مداد و کاغذ تصویری از یک نوار خطی را ترسیم نمود که به‌طور مساوی به مربع ها، فهرستی از نمادها و چند قاعده اصلی تقسیم شده بود. وی درواقع یک طرح کلیای را رسم نمود تا فرایند گام به گام نحوه انجام هر گونه محاسبات از ساده‌ترین تا پیچیده را نشان دهد.

بعد می‌گویند:

> اختراع منحصر به فرد تورینگ که امروزه به سادگی ماشین تورینگ می‌گوییم درواقع راه‌حلی برای مسأله پیچیده...

این هم صفحه آخر کتاب بود که ان شاء الله به آن نگاه کنید. در اطراف آن سه چیز را مطالعه کنید. اول: ماشین تورینگ. دوم: تست تورینگ و سوم: اتاق چینی. این‌ها مطالبی است که بعداً هم از نظر علمی و فلسفی و هم از نظر فنی با آن کار داریم.

والحمد لله رب العالمین