



ربات‌ها و اخلاق

سینا فلاح‌زاده‌راسته‌کناری. کارشناس‌ارشد مکانیک

تهران- ایرنا- اگر کسی در حین رانندگی با خودروهای عادی دچار تصادف منجر به فوت شود و شواهد نشان‌دهنده یک تصادف غیرعمدی باشند، راننده به دلیل قتل غیرعمد مورد بازخواست قرار گرفته و موظف به پرداخت غرامت خواهد شد. حال وضعیتی را در نظر بگیرید که خودروی موردنظر مجهز به هوش مصنوعی سطح بالا بوده و رفتارش توسط الگوریتم‌های بسیار پیچیده و تکامل‌یافته تنظیم شود. در این صورت چه کسی مقصر است؟ راننده؟ شرکت سازنده خودرو؟ یا شرکت توسعه‌دهنده الگوریتم هوش مصنوعی مورد استفاده؟ اگر نهایتاً معلوم شود که سیستم هوش مصنوعی خودرو طی یک فرایند ناخوشایند، در واقع از کنترل خارج شده و با «اختیار» خودش اقدام به این کار کرده است، چطور!؟ آیا می‌توان برای ربات‌های بسیار هوشمند که احتمالاً در آینده‌ای نزدیک تمام شئون زندگی ما را تحت‌تأثیر قرار خواهند داد علاوه‌بر هوشمندی، اختیار و مسئولیت و به‌این‌ترتیب حقوق و وظایف اخلاقی در نظر گرفت؟

به گزارش روزنامه شرق؛ مورد تکان‌دهنده‌ای مثل چت‌بات Tay را که توسط شرکت مایکروسافت در ۲۳ مارس ۲۰۱۶ در توئیتر معرفی شد در نظر بگیرید. Tay بلافاصله پس از شروع به کار، پاسخ‌دادن به سایر حساب‌های توئیتری را آغاز کرد، اما پاسخ‌های برهمکنشی و تکامل‌یافته این بات پس از چندساعت حالتی زنده، تهاجمی و نژادپرستانه به خود گرفتند. البته ابتدائاً اصلاً به نظر نمی‌رسید که این چت‌بات یک ایدئولوژی فاشیستی منسجم دارد، اما کامنت‌های نژادپرستانه آن موجب شد که مایکروسافت در کمتر از ۲۴ ساعت مجبور به خاموش‌کردن آن شود. هنوز نظریه‌های فراوانی درباره علت آنچه برای Tay پیش آمد وجود دارد، اما کلیت این قضیه باز هم نشان‌دهنده چالش بزرگ اخلاقی هوش مصنوعی است: اینکه نهایتاً چه کسی مسئول است؟

واقعیت امر این است که رابطه ما با ربات‌ها دست‌کم از میانه‌های قرن بیستم وارد فاز بسیار چالش‌برانگیز خود شده است، اما سناریوهایی که بیش از همه ذهن مردم را در این زمینه درگیر می‌کنند بیشتر مسائلی هستند که به‌جای چالش‌های واقعی به مسائل موردعلاقه فیلم‌های علمی-تخیلی مربوط می‌شوند. البته یک سناریوی آخرالزمانی رباتی (Robot Armageddon) که در آن ربات‌ها سرانجام به چنان سطحی از هوشمندی برسند که به طور کامل از کنترل آفرینندگان خود خارج شده و حتی بر آنها مسلط شوند با توجه به پیشرفت‌های هوش مصنوعی و دانش رباتیک، اصلاً دور از ذهن نیست، اما تا فرارسیدن آن روز چالش‌های مهم دیگری هم با ربات‌ها داریم و خواهیم داشت. برخی از مهم‌ترین این چالش‌ها ماهیتی اخلاقی دارند و به همین دلیل است که اخیراً مبحث اخلاق رباتی (= Roboethics) توسط فیلسوفان اخلاق و دانشمندان این حوزه بسیار جدی گرفته شده است (Lin, Abney & Jenkins, ۲۰۱۷). جالب اینجاست که ما هنوز از به‌دست‌دادن یک پاسخ قطعی به پرسش دیرینه اختیار انسان یا مسئله نحوه برخورد با حیوانات ناتوانیم، اما مسائل اخلاقی و فلسفی ربات‌ها ما را با سرعت فراوانی به مبارزه می‌خوانند.

این مورد مانند سایر مسائل فلسفی مربوط به تکنولوژی نشان‌دهنده این هستند که تکنولوژی تا چه اندازه می‌تواند سریع‌تر و جلوتر از تصور ما حرکت کند و ما تا چه اندازه باید در مقابله با چالش‌های احتمالی قرارگرفته بر سر راهمان از ناحیه تکنولوژی هوشیار باشیم. ربات‌ها ماشین‌هایی با توانایی حسگری، حرکت و پردازش اطلاعاتند. توانایی پردازش اطلاعات در ربات‌ها بیشتر به شکل‌های ادراک، استدلال، طرح‌اندازی، یادگیری و نیز پردازش سیگنال و کنترل فیدبک است. از میانه قرن بیستم به‌تدریج استفاده از ربات‌ها در کارخانه‌ها شروع شد و روزبه‌روز میزان استفاده از آنها افزایش یافت. استفاده از ربات‌ها ابتدائاً برای کارهای سخت تکراری یا کارهایی که باید در محیط‌های بسیار خشن و خطرناک انجام می‌شدند گسترش یافت، اما به‌تدریج با پیشرفت‌های علم رباتیک از ربات‌ها به نحو گسترده‌ای برای منظوره‌های خدماتی، نظامی و امنیتی، تحقیقات و آموزش، بهداشت و درمان، مسائل شخصی و مسائل زیست‌محیطی استفاده می‌شود (Lin, Abney & Bekey, ۲۰۱۴). جالب اینجاست که حضور ربات‌ها در هرکدام از این بخش‌ها با چالش‌های اخلاقی و پیچیدگی‌های فراوانی همراه است که در ادامه به برخی از آنها اشاره خواهد شد.

کدام اخلاق؟

به نظر می‌رسد پیش از آنکه بتوانیم در مورد اخلاق ربات‌ها اظهارنظر کنیم باید مسائل فراوانی را بین خودمان حل کنیم. دست‌کم نمی‌توانیم منتظر بمانیم تا ربات‌ها خودشان آن‌قدر هوشمند شوند که بتوانند تصمیمات دشوار اخلاقی را که حتی ما از گرفتن آنها درمی‌مانیم به راحتی اتخاذ کنند! سابقه ما در اندیشه فلسفی در باب اخلاق بسیار مفصل است و هر اصلی که راهنمای اخلاقی ما باشد می‌تواند در موارد حساس تأثیرات شگرفی بر زندگی فردی و اجتماعی ما برجای بگذارد. پس از گذشت نزدیک به دوونیم هزاره از تولد اندیشه فلسفی که از همان سرآغاز با اخلاق و معماهای اخلاقی درگیری داشت، هنوز بسیاری از مسائل و معماهای اخلاقی در سطوح گوناگون اخلاق کاربردی و اخلاق نظری و همین‌طور فرااخلاق حل‌نشده باقی مانده‌اند و این البته از ماهیت بفرنج این مسائل و همین‌طور کمبود تجربه و اختلاف دیدگاه‌های ما ناشی می‌شود. حال فرض کنید که در چنین شرایطی حل مسائل مربوط به اخلاق ربات‌ها نیز بر دوش نحیف اندیشه بشری قرار گرفته است. مسلماً رهاکردن این مسائل بدون کوچک‌ترین تلاشی برای رفع‌وجوع آنها می‌تواند به فجایع بزرگی دامن بزند، چنان‌که برخورد ما با تکنولوژی، به‌ویژه در یک‌ونیم قرن اخیر تاکنون بارها این مهم را به ما گوشزد کرده است. بپراه نخواهد بود اگر بگوییم تکنولوژی تمام آینده ماست و از این جهت هیچ راهی در اندیشه بدون توجه به تکنولوژی و تأثیر آن به جایی نخواهد رسید. اندیشه فلسفی در باب اخلاق با توجه به سابقه طولانی و آبشخورهای متنوعش به راه‌های گوناگونی قدم

گذاشته و چشم‌اندازهای متنوعی را پیش چشم خود دیده است. به غیر از مسائل مربوط به اخلاق کاربردی که به بایدها و نبایدها می‌پردازد، مکاتب اخلاقی فراوانی به توجیه این بایدها و نبایدها همت گماشته‌اند که در حوزه اخلاق نظری قابل بحث هستند. اخلاق ربات‌ها علاوه بر به‌ارث‌بردن همه این مسائل از اخلاق انسانی دارای مسائل و معماهای خاص خود نیز هست.

مباحث فراوانی درباره چارچوب‌های اخلاقی ربات‌ها وجود دارند و بیشتر آنها ناظر بر این مسئله هستند که چارچوب‌های اخلاقی انسان‌ها در یک جهان بدون ربات نمی‌توانند در یک جهان دارای ربات (جهان ما) به‌راحتی و بی‌دردسر کاربرد داشته باشند. به‌راستی اگر بتوانیم به ربات‌ها رفتار اخلاقی دهیم، کدام چارچوب‌ها باید مدنظر ما قرار گیرند؟ آیا باید ربات‌هایی کاملاً وظیفه‌گرا طراحی کنیم که در هر حالتی کار درست را انجام دهند؟ آیا قانون طلایی کانتی درباره ربات‌ها کاربرد دارد؟ آیا باید ربات‌ها را جوری طراحی کنیم که عواقب اعمال خود را بسنجند و بر اساس عواقب احتمالی اعمال خود دست به داوری اخلاقی و عمل بزنند؟ خوب و بد را چگونه برای ربات‌ها تعریف کنیم؟ ارزش‌های اخلاقی اجتماعات گوناگون چگونه؟ آیا هر جامعه‌ای باید اخلاق ربات‌هایش را با توجه به ارزش‌های اخلاقی خودش طراحی کند؟ یا می‌توان به یک توافق جهانی رسید؟ سؤالاتی از این دست در واقع برخی از بی‌شمار سؤالاتی هستند که در این زمینه در ابتدای کار به ذهن می‌رسند.

اخلاق ربات‌ها

اخلاق ربات‌ها شامل برخی از تحلیل‌های بسیار پایه‌ای مربوط به کار با ربات‌ها در حوزه اخلاق کاربردی می‌شود. این مسائل با وجود سادگی‌شان بسیار چالش‌برانگیزند و نیاز به تحلیل‌های دقیق فلسفی برای رسیدن به پاسخ‌های ولو اجمالی درباره آنها احساس می‌شود. برخی از این پرسش‌ها عبارت‌اند از (Lin et al., ۲۰۱۷; Tamburrini, ۲۰۰۹):

چه کسی مسئول آسیب‌های احتمالی ایجادشده از طریق ربات‌هاست؟

آیا می‌توان ربات‌های بسیار هوشمند را مانند انسان‌ها دارای حقوق و مسئولیت‌های اخلاقی دانست؟

آیا حریم شخصی افراد به‌وسیله ربات‌های شخصی دارای دسترسی به اینترنت تهدید می‌شود؟

آیا چارچوب‌های اخلاقی‌ای که از طریق آنها اخلاق انسان‌ها را بررسی می‌کنیم، درباره ربات‌ها نیز کاربرد دارند؟ اگر نه، ویژگی‌های چارچوب‌های جدید مورد نیاز چیستند؟

استفاده از ربات‌ها در جنگ در چه صورتی توجیه‌پذیر است؟

البته در این مقاله کوتاه نمی‌توانیم به‌طور کامل به این پرسش‌های پیچیده بپردازیم. درباره مسائل پایه‌ای مربوط به اخلاق ربات‌ها دیدگاه‌های گوناگونی وجود دارد. گروهی از دانشمندان معتقدند ربات‌ها هر چقدر هم که هوشمند و دارای توانایی تحلیل منطقی باشند، باز هم در دسته ماشین‌ها قرار می‌گیرند و گروهی دیگر درجاتی از تشخیص اخلاقی را برای آنها قائل می‌شوند اما نه به صورت دارابودن اختیار و مسئولیت، چنان‌که برای شخص انسانی می‌توان قائل شد

(Tzafestas, ۲۰۱۸). البته همه اینها به این معنا نیست در آینده با ربات‌هایی که دارای اختیار به معنی انسانی کلمه هستند، روبرو خواهیم شد. مفهوم پاسخ‌گویی (Responsibility/Verantwortung) یک مفهوم کانونی در مسائل مربوط به اخلاق ماشین‌هاست (Loh, ۲۰۱۹). ما غالباً عادت داریم فقط انسان‌ها را پاسخ‌گو و مسئول بدانیم و برای مثال، در صورتی که یک حیوان، یک ماشین و یک انسان در صحنه‌ای که در آن حادثه‌ای ناگوار رخ داده است حاضر باشند، ابتدا انسان را مسئول خواهیم دانست و اگر او را مقصر آن حادثه تشخیص دهیم، بازخواست کرده و اگر لازم باشد، مجازات می‌کنیم. اما درباره حیوانات و ماشین‌ها چنین تلقی‌ای از مسئولیت وجود ندارد. یک حیوان یا یک ماشین را (هر قدر هم که پیچیده باشد) با این طرز تلقی از مسئولیت و پاسخ‌گویی نمی‌توان مقصر یک حادثه قلمداد و بازخواست کرد. البته اگر متوجه شویم ماشین یا ربات مورد نظر ایراد دارد، برای برطرف‌کردن ایراد تلاش می‌کنیم اما هرگز انتظار پاسخ‌گویی از یک ماشین را نخواهیم داشت. اگر ایرادی در رفتار یک ربات باشد، آن را به‌راحتی به قوانین فیزیک یا طراحی ربات و الگوریتم آن نسبت می‌دهیم و تقصیری متوجه خود ماشین نمی‌کنیم. درباره حیوانات هم این نداشتن انتظار برای پاسخ‌گویی وجود دارد.

به نظر می‌رسد در شرایطی که هم‌اکنون در آن قرار داریم و با توجه به پیشرفت‌های هوش مصنوعی تا به این جای کار هنوز نمی‌توان ربات‌ها را به مفهوم کلاسیک کلمه پاسخ‌گوی کارهای خود دانست، اما این احتمال وجود دارد که در آینده‌ای نه چندان دور این انحصار مفهوم پاسخ‌گویی به نیاکان انسانی ربات‌ها از میان رخت بریند و ربات‌ها به صورت مستقل مسئول اعمال خود قلمداد شوند. طبیعتاً در آن شرایط به معنای کامل کلمه باید از اخلاق ربات‌ها سخن بگوییم و صد البته امیدوار باشیم خود ربات‌ها هم در این عرصه به کمک ما بیایند و اخلاقی فکر و عمل کنند! با توجه به اینکه ربات‌ها به‌زودی به‌طور گسترده در خانه‌های مردم حضور خواهند یافت، نگرانی‌های فراوانی درباره حریم شخصی افراد در منازل یا جاهای دیگر و احتمال تجاوز گسترده به آن از طریق ربات‌هایی که به اینترنت وصل هستند و می‌توانند از سوی هکرها و سرویس‌های جاسوسی استفاده شوند، وجود دارد. از سوی دیگر، زندگی با ربات‌ها می‌تواند موجب وابستگی‌های عاطفی میان انسان و ربات‌های هوشمند شود. همه این مسائل باید در طراحی اخلاقی ربات‌ها مدنظر قرار گیرند تا وجود ربات‌ها بتواند به جای سخت‌ترکردن زندگی به آسان‌ترشدن آن برای انسان‌ها کمک کند.

ربات‌های جنگی

جنگ یک پدیده بسیار شوم است و امروزه کمتر کسی پیدا می‌شود که از عقل سلیم بهره‌ای داشته باشد و در لزوم اجتناب هرچه بیشتر انسان‌ها از جنگ شک کند. در صورتی که جنگی به صورت ناخواسته درگیرد باز هم تلاش می‌شود تا به مدد تکنولوژی از میزان حضور مستقیم انسان‌ها (نظامی و غیرنظامی) در محل‌های بسیار خطرناک کاسته شود. این مسئله به نوبه خود موجب شده است که جنگ‌ها به شدت تکنولوژیک شوند و بسیاری از شاخه‌های تکنولوژی نیز جنگی شده‌اند،

یعنی به هر صورت نوعی استفاده جنگی برای آنها ارائه شده است. فناوری رباتیک نیز به دلیل توانایی‌های منحصر به فرد ربات‌ها در کارکردن در محیط‌های خشن و ناامن جنگی از این قاعده مستثنا نیست.

امروزه ده‌ها کشور در جهان دارای فناوری ساخت ربات‌های جنگی هستند و همین امر ملاحظات اخلاقی مربوط به این عرصه را پررنگ‌تر از گذشته می‌کند. مسئله ربات‌های قاتل (Killer Robots) که امروزه به صورت گسترده‌ای مورد توجه اندیشمندان و دانشمندان قرار گرفته است، در واقع یکی از مهم‌ترین چالش‌های اخلاق تکنولوژی به‌شمار می‌رود. این نوع ربات‌ها در صورتی که با روند حاکم امروز به پیشرفت خود ادامه دهند، در آینده‌ای نه چندان دور به سلاح‌های کشتار جمعی بدل خواهند شد که به دلیل دسترسی آسان‌تر و ارزان‌تر از سلاح‌های کشتار جمعی امروزه می‌توانند بسیار خطرناک باشند. سربازان به دلیل داشتن عواطف انسانی در رویارویی با صحنه‌های پیچیده و خطرناک جنگ به لحاظ اخلاقی دارای نقاط قوت و ضعف متعددی هستند. از جمله نقاط قوت ربات-سربازها نسبت به انسان‌ها این است که آنها اساساً دچار خشم یا فروپاشی عاطفی نمی‌شوند و زیر فشارهای بسیار زیاد که می‌تواند برای سربازان عادی خردکننده باشد وظایف محوله را با دقت و سرعت انجام دهند.

اما از طرفی چنین ربات‌هایی باز هم به دلیل فقدان عواطف بشری یا ضعف در برنامه‌ریزی ممکن است در میدان نبرد صحنه‌های دلخراشی را ایجاد کنند که اگر یک سرباز به جای آنها بود، هرگز دست به چنان جنایاتی نمی‌زد. البته روش‌های گوناگونی وجود دارد که بتوان به کمک آنها ربات‌ها را به نحوی برنامه‌ریزی کرد که برای مثال به غیرنظامیان و نیروهای خودی آسیب نرسانند، اما معماهای اخلاقی استفاده از ربات‌ها در جنگ مانند استفاده از هر نوع دیگر تکنولوژی در شرایط جنگی به این موارد محدود نیستند و تحلیل آنها نیازمند مباحث گسترده‌تری است (Arkin & Moshkina, ۲۰۰۷). برای مثال هنوز مشخص نیست چگونه می‌توان ربات‌هایی طراحی کرد که در میدان نبرد به کنوانسیون‌های بین‌المللی احترام بگذارند یا مثلاً با اسیران رفتار درستی داشته باشند.

چه می‌توان کرد؟

هوش مصنوعی و رباتیک با بسیاری از شئون زندگی ما درآمیخته‌اند و از شواهد چنان برمی‌آید که وسعت و عمق این درآمیختگی در آینده‌ای نه چندان دور به مراتب بیشتر از امروز خواهد شد. امروزه هیچ شاخه‌ای از اندیشه بشری نمی‌تواند بدون توجه به تکنولوژی کوچک‌ترین قدمی در راه فهم مشکلات انسان و حل آنها بردارد و از سویی دیگر رشد تکنولوژی بدون توجه به ملاحظات اخلاقی می‌تواند موجب بروز صدمات جبران‌ناپذیری به جوامع انسانی شود. تکنولوژی همواره ثابت کرده است که توانایی خروج از کنترل ما را دارد و درست به همین دلیل است که باید در مطالعات اخلاقی مربوط به هوش مصنوعی و رباتیک که به نحوی نشئت‌گرفته از مهم‌ترین انقلابات علمی-تکنولوژیکی عصر حاضر هستند، بسیار دقیق و پرتلاش باشیم چراکه از قلم‌انداختن هر جنبه‌ای از روابط پرجنب‌وجوش و پیچیده میان انسان‌ها و ربات‌ها می‌تواند برای ما بسیار گران تمام شود.

منابع:

۱- Arkin, R., & Moshkina, L. (۲۰۰۷). Lethality and Autonomous Robots: An Ethical Stance

۲- Lin, P., Abney, K., & Bekey, G. A. (۲۰۱۴). Robot ethics: the ethical and social implications of robotics: The MIT Press

۳- Lin, P., Abney, K., & Jenkins, R. (۲۰۱۷). Robot Ethics ۲.۰: From Autonomous Cars to Artificial Intelligence: Oxford University Press

۴- Loh, J. (۲۰۱۹). Verantwortung und

(Roboterethik. In M. Rath, F. Krotz, & M. Karmasin (Eds.), Maschinenethik: Normative Grenzen autonomer Systeme (pp. ۹۱-۱۰۵

.Wiesbaden: Springer Fachmedien Wiesbaden

۵- Tamburrini, G. (۲۰۰۹). Robot ethics: A view from the philosophy of science. Ethics and

robotics, ۱۱-۲۲

۶- Tzafestas, S. (۲۰۱۸). Ethics in robotics and automation: A general view. International

Robotics & Automation Journal, ۴, ۲۲۹-۲۳۴