



حیات ۳:
انسان در عصر هوش مصنوعی

حیات ۳: انسان در عصر هوش مصنوعی

مکس تگمارک

ترجمه‌ی جمیل آریایی

انتشارات مازیار

سرشناسه	: تگمارک، مکس، ۱۹۶۷ - م. Tegmark, Max
عنوان و نام پدیدآور	: حیات ۳: انسان در عصر هوش مصنوعی / مکس تگمارک؛ ترجمه‌ی جمیل آریایی.
مشخصات نشر	: تهران: مازیار، ۱۳۹۶.
مشخصات ظاهری	: ۳۸۴ ص.: مصور؛ ۱۴/۵×۲۱/۵ س.م.
شابک	: ۹۷۸-۶۰۰-۶۰۴۳-۸۷-۶
فروست	: قلمرو علم.
وضعیت فهرست‌نویسی: فیبا	
یادداشت	: عنوان اصلی: Life 3.0 : being human in the age of artificial intelligence, 2017.
یادداشت	: نمایه
موضوع	: هوش مصنوعی -- فلسفه
موضوع	: Artificial intelligence -- Philosophy
موضوع	: هوش مصنوعی -- جنبه‌های اجتماعی
موضوع	: Artificial intelligence -- Social aspects
موضوع	: هوش مصنوعی -- جنبه‌های اخلاقی
موضوع	: Artificial intelligence -- Moral and ethical aspects
موضوع	: خودکاری -- جنبه‌های اجتماعی
موضوع	: Automation -- Social aspects
موضوع	: خودکاری -- جنبه‌های اخلاقی
موضوع	: Automation -- Moral and ethical aspects
موضوع	: تکنولوژی -- آینده‌نگری
موضوع	: Technological forecasting
شناسه افزوده	: آریایی، جمیل، ۱۳۳۰ - مترجم
رده بندی کنگره	: ۹۱۳۹۶ ز ۸ت/۷/۴۳۳۴ Q
رده بندی دیویی	: ۰۰۶/۳۰۱
شماره کتابشناسی ملی	: ۵۰۰۴۲۳۰

www.mazyarpub.ir
mazyarpub@yahoo.com

انتشارات مازیار

مقابل دانشگاه تهران، ساختمان ۱۲۹۶ (ظروفچی) طبقه اول، واحد ۴، تلفن ۶۶۴۶۲۴۲۱

حیات ۳: انسان در عصر هوش مصنوعی

مکس تگمارک

ترجمه‌ی جمیل آریایی با همکاری فرید آریایی

صفحه‌آرایی مرواک.

چاپ اول ۱۳۹۷

شمارگان ۱۱۰۰

شابک ۹۷۸-۶۰۰-۶۰۴۳-۸۷-۶

فهرست مطالب

	آغاز سخن: داستان گروه اُمگا ۹
۲۹	۱. به مهم‌ترین گفتگوی روزگار ما خوش آمدید
۳۰	تاریخچه‌ی مختصر پیچیدگی
۳۱	سه مرحله‌ی حیات
۳۷	مناقشات
۳۷	آرمان‌شهروندان دیجیتالی
۴۰	تکنولوژی‌های هراس‌ها
۴۰	جنبش هوش مصنوعی سودمند
۴۵	برداشت‌های نادرست
۴۸	افسانه‌های زمان‌بندی
۵۱	افسانه‌های مناقشه‌برانگیز
۵۲	افسانه‌هایی درباره‌ی خطرات هوش مصنوعی
۵۴	راه پیش‌رو
۶۱	۲. ماده باهوش می‌شود
۶۱	هوش چیست؟
۶۷	حافظه چیست؟
۷۳	محاسبه چیست؟
۸۴	یادگیری چیست؟
۹۷	۳. آینده‌ی نزدیک: دست‌آوردها، خطاها، قوانین، سلاح‌ها و اشتغال
۹۸	دست‌آوردها
۱۰۹	خطاها
۱۲۱	قوانین

۶ انسان در عصر هوش مصنوعی

- ۱۲۷ سلاح‌ها
- ۱۳۷ اشتغال
- ۱۵۰ هوش هم‌تراز انسان؟
- ۱۵۵ ۴. انفجار هوش مصنوعی؟
- ۱۵۷ خودکامگی
- ۱۵۹ پرومته جهان را تسخیر می‌کند
- ۱۷۲ سناریوهای پرواز آهسته و چندقطبی
- ۱۷۶ سایبورگ‌ها (انسان‌های سایبری) و آپلودها
- ۱۷۶ (انسان‌های شبیه‌سازی شده)
- ۱۷۹ چه رخ خواهد داد؟
- ۱۸۵ ۵. پیامد: ۱۰,۰۰۰ سال بعد
- ۱۸۹ آرمان‌شهروندی آزادی‌خواه
- ۱۹۴ دیکتاتور خیرخواه
- ۱۹۸ آرمان‌شهر برابری‌خواه
- ۲۰۲ دربان
- ۲۰۳ خدای نگهبان
- ۲۰۵ خدای برده
- ۲۱۱ سلطه‌گران
- ۲۱۵ فرزندان
- ۲۱۷ نگهبان باغ وحش
- ۲۱۸ ۱۹۸۴
- ۲۲۱ بازگشت
- ۲۲۳ خودنابودی
- ۲۲۸ چه می‌خواهید؟
- ۲۳۱ ۶. موقوفه‌ی کیهانی ما، یک میلیارد سال بعد

۲۳۳	استفاده‌ی بهینه از منابع
۲۴۸	دسترسی به منابع از طریق استقرار در کیهان
۲۶۶	سلسله مراتب کیهانی
۲۷۸	چشم‌انداز

۷. اهداف ۲۸۳

۲۸۳	فیزیک: منشأ اهداف
۲۸۷	زیست‌شناسی: تکامل اهداف
۲۸۹	روان‌شناسی: تعقیب اهداف و سرپیچی از آنها
۲۹۱	مهندسی: برون‌سپاری اهداف
۲۹۴	هوش مصنوعی دوستانه: هم‌سوسازی اهداف
۳۰۴	اخلاقیات: انتخاب اهداف
۳۱۰	اهداف نهایی؟

۸. آگاهی ۳۱۹

۳۱۹	چه کسی اهمیت می‌دهد؟
۳۲۰	خودآگاهی چیست؟
۳۲۱	مسئله چیست؟
۳۲۵	آیا خودآگاهی ورای علم است؟
۳۲۸	سرنخ‌های آزمایشگاهی درباره‌ی خودآگاهی
۳۳۷	نظریه‌های خودآگاهی
۳۴۴	مناقشات خودآگاهی
۳۴۸	خودآگاهی هوش مصنوعی چه احساسی دارد؟
۳۵۲	معنی

۳۵۷ پایان سخن: داستان گروه انستیتو آینده‌ی حیات

۳۵۸	انستیتو حیات آینده متولد می‌شود
۳۶۱	ماجراجویی پورتو ریکو

۸ انسان در عصر هوش مصنوعی

۳۶۶

راه‌اندازی پژوهش‌های امنیّت هوش مصنوعی

۳۷۰

اصول هوش مصنوعی آسیلومار

۳۷۴

خوش‌بینی فکورانه

به مهم ترین گفتگوی روزگار ما خوش آمدید

تکنولوژی این توانایی را به حیات داده است تا چنان شکوفا شود یا چنان خود را نابود سازد که در گذشته امکان آن را نداشت.

— انستیتو آینده‌ی حیات

جهان هستی پس از نزدیک به چهارده میلیارد سال که از تولدش می‌گذرد تازه از خواب برخاسته و از هستی خود آگاه شده است. از روی این سیاره‌ی آبی‌رنگ کوچک، بخش‌های ناچیز هوشمندی در جهان ما با تلسکوپ‌های خود به سوی کیهان خیره شده و پی در پی کشف می‌کنند که آنچه فکر می‌کردند وجود داشته تنها پاره‌ی اندکی از جهانی فراخ‌تر است: منظومه‌ی شمسی، کهکشان، و جهانی با صدها میلیارد کهکشان که در الگوی زیبایی از گروه‌ها، خوشه‌ها، و ابرخوشه‌ها آرایش یافته‌اند. هرچند این ستاره‌نگران خودآگاه در بسیاری موارد با هم اختلاف دارند، اما با یکدیگر هم عقیده‌اند که این کهکشان‌ها زیبا و تحسین‌برانگیزند.

زیبایی، نه در قوانین فیزیک بلکه در چشمان این نظاره‌گران است و از این رو پیش از این که جهان ما از خواب برخیزد، زیبایی نبود. بیداری جهان ما شکوهمند است و ارزش دارد آن را جشن بگیریم چون جهان ما را از زامبی بی‌مغز ناخودآگاهی به سامانه‌ی زیستی زنده‌ای بدل ساخته که خوداندیشی، زیبایی، امید، و دنبال کردن اهداف، معنی و مقصود را در خود پناه داده است. به گمان من اگر جهان ما از خواب بیدار نمی‌شد همه چیز معنی خود را از دست می‌داد و جهان چیزی به جز فضای فراخ بیهوده‌ای بیش نبود. اگر جهان ما بر اثر فاجعه‌ای کیهانی یا حادثه‌ی خودخواسته‌ی ناگواری به خواب فرو رود، وای که همه چیز بی‌معنی می‌شود.

از سوی دیگر، اگر این اتفاق‌ها نیافتد، اوضاع می‌تواند بهتر از این‌ها باشد. هنوز نمی‌دانیم که آیا ما انسان‌ها تنها ستاره‌نگران، چه بسا نخستین آن‌ها، در کیهان‌مان هستیم. با وجود این، آن قدر درباره‌ی جهان آموخته‌ایم که بدانیم این جهان توانایی آن را دارد که بیش از پیش از خواب بیدار شود. شاید ما کسی را می‌مانیم که تازه از خواب بیدار شده و بارقه‌ای از خودآگاهی را تجربه می‌کند و آن پیش‌آگاهی از

خودآگاهی بیشتری است که فرد پس از باز کردن کامل چشمان‌اش به او دست می‌دهد. شاید حیات تا میلیاردها یا صدها میلیارد سال در کیهان ما نشو و نما کند و شاید این‌ها همه از تصمیماتی ناشی خواهد شد که ما در طول حیات خود در اینجا روی سیاره‌ی کوچک‌مان می‌گیریم.

تاریخچه‌ی مختصر پیچیدگی

حال این بیداری حیرت‌انگیز چگونه اتفاق افتاد؟ این رویداد، رویدادی تک‌افتاده نبود بلکه مرحله‌ای در فرایند ۱۳/۸ میلیارد ساله‌ی بی‌قراری بود که جهان ما را پیچیده‌تر و جالب‌تر کرد و هم‌چنان با شتابی روزافزون ادامه دارد.

من فیزیکدان، خوشحال هستم از این که بیشتر وقت خود را در طول ربع قرن گذشته برای درک تاریخ کیهانی‌مان صرف کرده‌ام و در این گذر عمر به کشف‌های شگفت‌انگیزی دست یافته‌ام. از روزهایی که دانشجوی دوره‌ی تحصیلات تکمیلی بودم، فراهم آمدن ترکیبی از تلسکوپ‌های دقیق‌تر، کامپیوترهای قوی‌تر، و درک بهتر، ما را بر آن داشته تا بحث ۱۰ میلیارد ساله یا ۲۰ میلیارد ساله بودن جهان را به بحث ۱۳/۷ میلیارد ساله یا ۱۳/۸ میلیارد ساله بودن آن بکشانیم. ما فیزیکدان‌ها به یقین نمی‌دانیم که چه عاملی سبب مهبانگ شده و آیا مهبانگ آغاز همه چیز بوده یا این که مهبانگ دنباله‌ی سلسله رویدادهای آغازین بوده است. با وجود این، با فراهم شدن آشنایی از اندازه‌گیری‌های دقیق، درک دقیق‌تری از آنچه از زمان مهبانگ به این سو رخ داده به دست آمده است و اینک می‌خواهم تاریخ ۱۳/۸ میلیارد ساله‌ی کیهانی را در چند دقیقه خلاصه کنم.

در آغاز، نور بود. در کسر ناچیزی از ثانیه‌ی نخست پس از مهبانگ، کل فضای که با تلسکوپ‌های امروزی می‌بینیم («جهان قابل مشاهده‌ی ما» یا به اختصار «جهان ما») به مراتب داغ‌تر از مرکز خورشید بود و به سرعت انبساط یافت. هرچند این خود شکوهی ویژه داشت، اما از این که جهان شامل چیزی بیش از سوپ بی‌روح، غلیظ، داغ، و یکنواخت ذرات بنیادی نبود، ملال‌آور می‌نمود. همه جای آن یکسان به نظر می‌رسید و تنها ساختار جالب موجود در آن امواج صوتی پراکنده‌ای بودند که در جاهایی این سوپ را ۰/۰۰۱٪ غلیظ‌تر کرده بودند. باور بر این است که این امواج کوچک از افت و خیزهای کوانتومی ناشی شده بودند، چون اصل عدم قطعیت مکانیک کوانتومی نمی‌گذارد چیزی کاملاً بی‌روح و یکنواخت بماند.

جهان که انبساط می‌یافت و سرد می‌شد، ذرات آن با هم ترکیب می‌شدند و اشیای

پیچیده‌تری می‌ساختند و جهان مکانی جالب‌تر می‌شد. در کسر ناچیزی از ثانیه‌ی نخست، نیروهای هسته‌ای قوی کوارک‌ها را به هم پیوند دادند و پروتون‌ها (هسته‌های هیدروژن) و نوترون‌ها را ساختند و در عرض چند دقیقه تعدادی از آن‌ها در هسته‌های هلیوم گرد هم آمدند. در حدود ۴۰۰,۰۰۰ سال بعد، نیروی الکترومغناطیسی این هسته‌ها را به الکترون‌ها آراست و اتم‌های نخستین را ساخت. جهان ما هم‌چنان در حال انبساط بود که این اتم‌ها به تدریج سرد شدند و به گاز تیره‌ی سردی بدل گشتند و تیرگی این شب نخست تا ۱۰۰ میلیون سال بعد بر جهان سایه افکند. آنگاه که نیروی گرانشی توانست افت و خیزهای این گاز را تقویت کند و اتم‌ها را به سوی هم کشاند تا ستارگان و کهکشان‌های آغازین شکل گرفتند، این شب تیره پگاه کیهانی ما شد. این ستارگان، اتم‌های هیدروژن را به هم چسبانند و اتم‌های سنگین‌تری چون اتم‌های کربن، اکسیژن، و سیلیکون را آفریدند و در این فرایندها گرما و نور آزاد شد. عمر این ستارگان که به سر آمد، اتم‌هایی که خلق کرده بودند در کیهان پراکنده شدند و سیاراتی را گرد ستارگان نسل دوم به وجود آوردند.

در لحظه‌ای از زمان، گروهی از اتم‌ها در الگوی پیچیده‌ای گرد هم آمدند که هم پایدار بود و هم می‌توانست تولیدمثل کند. دیری نپائید که از این الگوها دو نسخه پدید آمد و این دو برابر شدن‌ها تداوم یافت. چهل بار دو برابر شدن کافی بود تا هزار میلیارد نسخه از این الگوی اولیه ساخته شود و در چشم به هم زدنی جهان را به تسخیر خود درآورد. آنچه پدیدار شد، حیات بود.

سه مرحله‌ی حیات

این که حیات را چگونه باید تعریف کرد، بد جوری جنجال‌برانگیز است. تعدادی از این تعاریف، شرایط ویژه‌ای را بر حیات تحمیل می‌کنند، از جمله تشکیل حیات از یاخته‌ها، که هم ماشین‌های هوشمند آینده و هم تمدن‌های برون‌زمینی را ناممکن می‌شمارند. از آنجا که ما نمی‌خواهیم افکار خود درباره‌ی آینده‌ی حیات را به گونه‌هایی محدود سازیم که تاکنون دیده‌ایم، حیات را فرایندی تعریف می‌کنیم که هم پیچیدگی خود را حفظ می‌کند و هم توان تولیدمثل دارد. حاصل این تولیدمثل مهم نیست (از اتم ساخته شده است) بلکه اطلاعات مهم است (از بیت ساخته شده است) که معلوم می‌کند اتم‌ها چگونه آرایش یافته‌اند. آنگاه که باکتری نسخه‌ای از DNA خود را می‌سازد، اتم جدیدی خلق نمی‌شود، بلکه مجموعه‌ای از اتم‌ها در همان الگوی اولیه مرتب می‌شوند و این‌گونه است که اطلاعات کپی می‌شوند. به بیان دیگر،







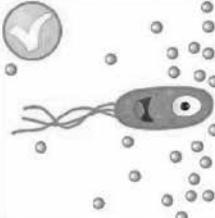


می‌توانیم حیات را سامانه‌ی پردازنده‌ی اطلاعات خودتکثیری تصور کنیم که اطلاعات آن (نرم‌افزار) هم رفتار و هم طرح کلی سخت‌افزارش را تعیین می‌کند. حیات نیز مثل جهان هستی ما رفته‌رفته پیچیده‌تر و جالب‌تر شده است^۱ و آن‌گونه که توضیح خواهیم داد به من کمک کرده است تا اشکال حیات را در سه سطح از پیچیدگی طبقه‌بندی کنم. من این سه سطح را در شکل ۱.۱ خلاصه کرده‌ام.

هنوز پرسش بی‌پاسخی هست که حیات برای نخستین‌بار چگونه، کی، و کجا در جهان ما پدیدار گشته است، اما قراین قانع‌کننده‌ای در دست داریم که حیات روی زمین ۴ میلیارد سال پیش شکل گرفته است. سیاره‌ی ما مدت‌ها پیش از آن ملامت از گونه‌های متنوع حیات بوده است. کامیاب‌ترین گونه‌ها که بر گونه‌های دیگر حاکم شده‌اند، قادر بودند به طریقی در برابر محیطشان عکس‌العمل نشان دهند. به ویژه، آن‌ها گونه‌هایی بودند که دانشمندان کامپیوتر «عوامل هوشمند» می‌خوانند، یعنی موجوداتی که از طریق حسگرهایی که دارند اطلاعاتی درباره‌ی محیط به دست می‌آورند و آنگاه این اطلاعات را پردازش می‌کنند تا تصمیم بگیرند که در برابر محیط خود چه عکس‌العملی نشان دهند، برای مثال ما با استفاده از چشم و گوش خود دریافت می‌کنیم و تصمیم می‌گیریم که به پرسش مخاطب‌مان باید چه پاسخی بدهیم. با وجود این، این گونه‌های زیستی می‌توانند سخت‌افزار و نرم‌افزار بسیار ساده‌ای داشته باشند.

برای مثال، بسیاری از باکتری‌ها حسگری دارند که با آن غلظت شکر مایعات اطراف خود را اندازه می‌گیرند و با استفاده از ساختارهای پارو شکلی که تاژک نام دارند می‌توانند شنا کنند. سخت‌افزاری که رابط بین این حسگر و تاژک‌هاست چه بسا الگوریتم ساده اما مفید زیر را اجرا می‌کند: «اگر حسگر غلظت شکر را مقداری گزارش کند که از مقدار چند ثانیه پیش آن کمتر است، آنگاه چرخش تاژک‌های مرا وارونه کن تا من تغییر جهت بدهم.»

ما صحبت کردن و مهارت‌های بی‌شماری آموخته‌ایم. از سوی دیگر، باکتری‌ها دانش‌آموزان کُندذهنی هستند. DNA باکتری‌ها نه تنها طرح سخت‌افزارها، از جمله حسگرهای شکر و تاژک‌ها، بلکه طرح نرم‌افزار آن‌ها را نیز مشخص می‌کند. آن‌ها هیچ‌گاه یاد نمی‌گیرند که به سوی شکر شنا کنند و از این‌رو این الگوریتم از آغاز به

۱. چرا حیات به سوی پیچیدگی رشد می‌کند؟ تکامل، حیاتی را می‌طلبد که آن قدر پیچیده باشد که نظم حاکم بر محیط اطراف‌اش را پیشگویی و درک کند و از این‌رو در محیطی که پیچیده باشد حیاتی که پیچیده‌تر و هوشمندتر باشد تکامل می‌یابد. حال این حیات هوشمند محیط پیچیده‌تری را برای رقابت گونه‌های حیات خلق می‌کند و این محیط نیز به نوبه‌ی خود به سوی پیچیدگی تکامل یافته و دست آخر سامانه‌ی زیستی از حیات فوق‌العاده پیچیده‌تر را می‌آفریند.

<p>آیا می تواند سخت افزارش را طراحی کند؟</p>			
<p>آیا می تواند نرم افزارش را طراحی کند؟</p>			
<p>آیا می تواند به حیات ادامه دهد و تولیدمثل کند؟</p>			
	<p>حیات ۱/۰ (زیست شناختی ساده)</p>	<p>حیات ۲/۰ (فرهنگی)</p>	<p>حیات ۳/۰ (تکنولوژیکی)</p>

شکل ۱.۱ سه مرحله‌ی حیات: تکامل زیست‌شناختی، تکامل فرهنگی، و تکامل تکنولوژیکی. حیات ۱/۰ نمی‌تواند در طول عمر خود سخت‌افزار و نرم‌افزارش را بازطراحی کند، چون این دو را DNA حیات تعیین می‌کند و تنها در طول تکامل به چند نسل آینده تغییر می‌کند. برعکس، حیات ۲/۰ می‌تواند بخش بزرگی از نرم‌افزار خود را بازطراحی کند، چون انسان می‌تواند مهارت‌های تازه‌ی پیچیده، از جمله زبان، ورزش، و حرفه، را بیاموزد و جهان‌بینی و اهداف خود را به روز کند. حیات ۳/۰ که هنوز روی زمین پدید نیامده است، به جای این که صبر کند تا در اثنای تکامل به چند نسل آینده، نرم‌افزار و سخت‌افزارش را اندک اندک بازطراحی کند می‌تواند در آنی هم نرم‌افزار و هم سخت‌افزارش را بنیادی بازطراحی کند.

شکل سخت‌افزاری در DNA آن‌ها نصب می‌شود. البته نوعی فرایند یادگیری در کار بوده، اما در طول عمر این نوع خاص از باکتری‌ها رخ نداده است. این فرایند در نسل‌های بعدی این گونه‌ی باکتریایی، از طریق نوعی فرایند آزمون و خطا در اثنای طول عمر چند نسل بعدی آن‌ها رخ داده است که در آن انتخاب طبیعی آن دسته از جهش‌های کاتوره‌ای DNA را ترجیح داده است که مصرف شکر این باکتری را بهینه

کرده‌اند. تعدادی از این جهش‌ها طراحی تازک‌ها و سخت‌افزارهای دیگر را اصلاح کرده‌اند، حال آن‌که جهش‌های دیگر به بهینه‌سازی سامانه‌ی پردازش اطلاعات دست زده‌اند که الگوریتم یافتن شکر و دیگر نرم‌افزارها را پیاده می‌کند.

این چنین باکتری‌هایی مثال حیاتی هستند که من آن را «حیات ۱/۰» نامیده‌ام، یعنی، حیاتی که سخت‌افزارها و نرم‌افزارهای آن به جای این که طراحی شوند، تکامل می‌یابند. از سوی دیگر، من و شما مثال «حیات ۲/۰» هستیم، یعنی، حیاتی که سخت‌افزارهای آن تکامل می‌یابند اما نرم‌افزارهای آن طراحی می‌شوند. منظورم از نرم‌افزارها همه‌ی الگوریتم‌ها و دانشی است که برای پردازش اطلاعات دریافتی از حس‌هامان به کار می‌بریم و تصمیم می‌گیریم که چه باید بکنیم، از توانایی ما برای تشخیص دوستان‌مان هنگام دیدن آن‌ها گرفته تا توانایی راه رفتن، خواندن، نوشتن، محاسبه کردن، آواز خواندن و شوخی کردن.

آنگاه که به دنیا می‌آییم هیچ یک از این توانایی‌ها را نداریم و همه‌ی این نرم‌افزارها بعدها طی فرایندی به نام یادگیری، در مغز ما برنامه‌ریزی می‌شوند. در حالی که برنامه‌ی دوران کودکی ما را بیشتر خانواده و آموزگاران طراحی می‌کنند و آن‌ها هستند که تصمیم می‌گیرند ما باید چه چیزهایی یاد بگیریم، اما اندک اندک که رشد می‌کنیم توانا تر می‌شویم تا خود نرم‌افزارهامان را طراحی کنیم. چه بسا مدرسه از ما بخواهد که زبان بیگانه‌ی دوم یاد بگیریم: آیا می‌خواهیم مدول نرم‌افزار یادگیری زبان فرانسه را در مغزمان پیاده کنیم یا مدول نرم‌افزار یادگیری زبان اسپانیایی را؟ می‌خواهیم بازی تنیس را یاد بگیریم یا بازی شطرنج را؟ می‌خواهیم درس بخوانیم تا آشبز، وکیل، یا داروساز شویم؟ آیا با خواندن این کتاب می‌خواهیم درباره‌ی هوش مصنوعی و آینده‌ی حیات بیشتر بدانیم؟

حیات ۲/۰ که می‌تواند نرم‌افزارهای خود را طراحی کند، از حیات ۱/۰ باهوش‌تر است. هوش بالا، هم نیاز به سخت‌افزارهای زیاد (که از اتم ساخته می‌شوند) دارد و هم به نرم‌افزارهای زیاد (که از بیت‌ها ساخته می‌شوند). این واقعیت که اغلب سخت‌افزارهای ما انسان‌ها پس از تولد به ما افزوده می‌شوند (از طریق فرایند رشد) مفید است چون جثه‌ی ما به بزرگی مجرای زایمان مادرمان بستگی ندارد. به همین منوال، این واقعیت که قریب به اتفاق نرم‌افزارهای ما انسان‌ها پیش از تولد به ما افزوده می‌شوند (از طریق فرایند یادگیری) مفید است چون هوش نهایی ما به این بستگی ندارد که هنگام حاملگی چه مقدار اطلاعات از طریق DNA نوع اول ما به ما منتقل می‌شود. وزن من بیست و پنج برابر وزن زمانی است که به دنیا آمدم و اتصالات سیناپسی که

نورون‌های مغز ما به هم وصل می‌کنند می‌توانند، در مقایسه با DNA که با آن متولد شده‌ام، در حدود صد هزار مرتبه اطلاعات بیشتری را ذخیره کنند. سیناپس‌های ما همه‌ی دانسته‌ها و مهارت‌های ما، به ارزش ۱۰۰ ترابایت اطلاعات، را در خود ذخیره می‌کنند، حال آن که DNA تنها در حدود یک گیگابایت اطلاعات را می‌تواند در خود ذخیره کند که ظرفیت لازم برای ذخیره کردن یک فیلم سینمایی کامل هم نمی‌شود. بنابراین از نظر فیزیکی امکان ندارد که نوزاد تازه متولد بتواند به طور کامل انگلیسی صحبت کند و در امتحانات ورودی کالج قبول شود؛ هیچ راهی نیست که اطلاعات از پیش در ذهن او بارگذاری شده باشد چون مدول اطلاعات اولیه‌ای که وی از پدر و مادر خود به ارث می‌برد (DNA او)، ظرفیت ذخیره‌سازی اطلاعات کافی را ندارد.

حیات ۲/۰ که قادر است نرم‌افزارهای خود را طراحی کند از حیات ۱/۰ باهوش‌تر و نیز انعطاف‌پذیرتر است. هرگاه محیط تغییر کند، حیات ۱/۰ فقط می‌تواند در طول چند نسل بعدی به کندی تکامل یافته و با محیط سازگار شود. از سوی دیگر، حیات ۲/۰، با به روز کردن نرم‌افزارهای خود، در آنی با محیط جدید سازگار می‌شود. برای مثال، باکتری‌هایی که با آنتی‌بیوتیک‌ها از بین برده می‌شوند، پس از گذشت چند نسل در برابر این داروها مقاوم می‌شوند، با وجود این، تک باکتری‌ها به هیچ‌روی رفتارشان را تغییر نمی‌دهند. برای مثال، کسی که پی می‌برد به بادام‌زمینی حساسیت دارد، بلافاصله رفتارشان را تغییر می‌دهد و بادام‌زمینی نمی‌خورد. انعطاف‌پذیری حیات ۲/۰ در سطح جمعیتی تأثیر فراوانی دارد، چون هرچند DNA ما انسان‌ها در طول پنجاه هزار سال گذشته دستخوش تحول بنیادینی نشده است، اما اطلاعاتی که به طور جمعی در ذهن ما، کتاب‌ها، و کامپیوترها ذخیره شده به حد انفجار رسیده است. پیاده‌سازی مدول نرم‌افزاری که ما را قادر ساخته تا با زبان گفتاری پیچیده‌ای ارتباط برقرار کنیم، این امکان را فراهم آورده است تا اطلاعات مفید مغز هر فرد را به مغز افراد دیگر کپی کنیم تا پس از مرگ آن فرد، اطلاعات از دست نروند. پیاده‌سازی مدول نرم‌افزاری که ما را قادر کند بخوانیم و بنویسیم، به ما این توانایی را می‌دهد که اطلاعاتی بیش از آنچه را مردم می‌توانند حفظ کنند، ذخیره کنیم و به اشتراک بگذاریم. با طراحی نرم‌افزاری مغزی که بتواند تکنولوژی تولید کند (یعنی، با مطالعه‌ی علوم و مهندسی) ما توانسته‌ایم بیشتر اطلاعات موجود در جهان را، با چند کلیک، در اختیار بسیاری از انسان‌های دنیا قرار دهیم.

این انعطاف‌پذیری، حیات ۲/۰ را بر زمین حاکم کرده است. دانش جمعی آدمی که از قیود ژنتیکی رهایی یافته، بر اثر دست‌آوردهایی که یکی دیگری را به وجود

آورده، از جمله زبان گفتار، نوشتن، صنعت چاپ، علوم مدرن، کامپیوتر، اینترنت، و غیره، با آهنگ روزافزونی هم‌چنان در حال بالندگی است. تکامل فرهنگی رو به رشد نرم‌افزار اشتراکی ما به نیروی غالبی بدل گشته که آینده‌ی آدمی را شکل می‌دهد و تکامل زیست‌شناختی کند یخچالی ما را به انزوا می‌راند.

با این همه تکنولوژی‌های قدرتمند که امروزه داریم، همه‌ی گونه‌های حیاتی که می‌شناسیم از لحاظ سخت‌افزارهای زیست‌شناختی محدودیت بنیادینی دارند. این سخت‌افزارها هیچ‌کدام میلیون‌ها سال عمر نمی‌کنند تا همه‌ی ویکیپدیا را از بر کنند، همه‌ی علوم شناخته‌شده را یاد بگیرند یا از پروازهای فضایی بدون فضاپیما لذت ببرند. هیچ‌یک از این سخت‌افزارها قادر نیستند کیهان ملال‌آور ما را به کره‌ی زیستی رنگارنگی بدل سازند که میلیاردها یا میلیون‌ها سال نشو و نما کند و جهان ما را قادر سازد که دست آخر همه‌ی توانایی‌های خود را به فعل برساند و کامل از خواب بیدار شود. این‌ها همه محتاج این هستند که حیات به نسخه‌ی حیات ۳/۰ به‌روزرسانی شود که این توانایی را دارد که نه تنها نرم‌افزارها، بلکه سخت‌افزارهای خود را نیز طراحی کند. به بیان دیگر، حیات ۳/۰ سر سلسله‌ی دودمان خود است و قیود دست و پاگیر تکامل را ندارد.

مرز دقیق میان مراحل سه‌گانه‌ی حیات، مشخص نیست. اگر باکتری‌ها حیات ۱/۰ و انسان‌ها حیات ۲/۰ باشند، آنگاه شاید بتوانیم موش‌ها را جزو حیات مرحله‌ی ۱/۱ طبقه‌بندی کنیم که می‌توانند بسیاری از چیزها را یاد بگیرند اما نه آن قدر که بتوانند به زبان مشترکی صحبت کنند یا اینترنت را اختراع کنند. فزون بر آن، از آنجا که موش‌ها زبان مشترکی ندارند، آنچه را آموخته‌اند به هنگام مرگ از دست می‌دهند و به نسل‌های بعدی خود منتقل نمی‌کنند. به همین منوال، می‌توانیم استدلال کنیم که انسان‌های امروزی را باید جزو حیات مرحله‌ی ۲/۱ به حساب آوریم که می‌توانند به‌روزرسانی سخت‌افزاری جزئی، از جمله کاشت دندان‌ها، زانوان، تنظیم‌کننده‌های ضربان قلب مصنوعی را انجام دهند، اما توانایی‌های آن‌ها به اندازه‌ای نیست که بتوانند قدشان را ده برابر بلندتر یا مغزشان را هزار برابر بزرگتر کنند.

خلاصه این که می‌توانیم مراحل رشد حیات را به سه مرحله تقسیم کنیم که ویژگی توانایی حیات برای طراحی خود، آن‌ها را از هم متمایز می‌کند:

- حیات ۱/۰ (مرحله‌ی زیست‌شناختی): سخت‌افزارها و نرم‌افزارهای خود را رشد می‌دهد.
- حیات ۲/۰ (مرحله‌ی فرهنگی): سخت‌افزارهای خود را رشد می‌دهد و بیشتر

نرم افزارهای خود را طراحی می کند.

- حیات ۳/۰ (مرحله ی تکنولوژیکی): سخت افزارها و نرم افزارهای خود را طراحی می کند.

پس از ۱۳/۸ میلیارد سال تکامل کیهانی، رشد بنیادینی بر روی زمین شتاب گرفته است: حیات ۱/۰ چهار میلیارد سال پیش شکل گرفت، حیات ۲/۰ (ما انسان ها) در حدود صد میلیون سال پیش پدیدار شد، و بسیاری از پژوهشگران هوش مصنوعی بر این باورند که چه بسا حیات ۳/۰، با پیشرفت پژوهش های هوش مصنوعی، در اثنای قرن آتی، شاید هم در اثنای طول عمر ما، رخ دهد. اگر حیات ۳/۰ پدیدار شود، آنگاه چه اتفاق خواهد افتاد و چه تأثیری بر ما خواهد گذاشت؟ این است عنوان کتابی که هم اینک پیش رو دارید.

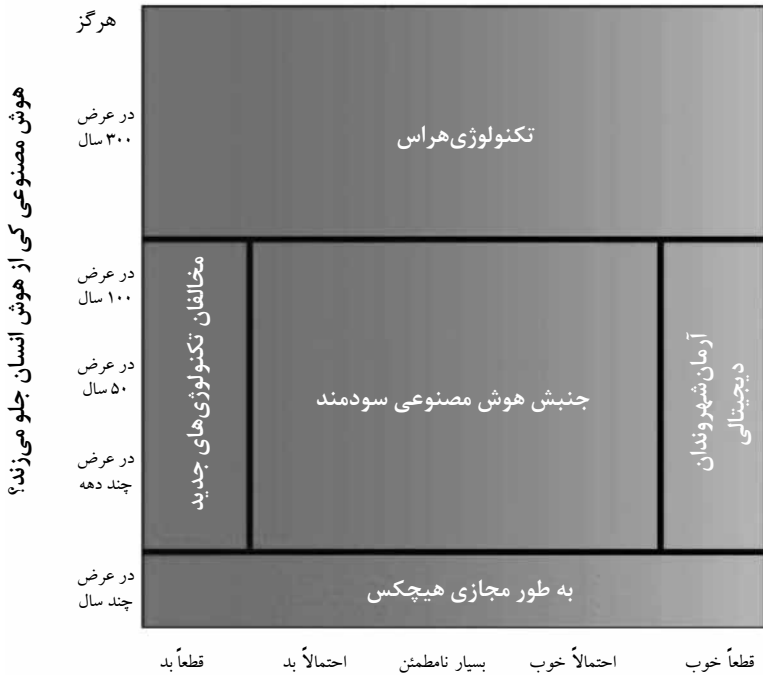
مناقشات

این پرسش بیش از اندازه مناقشه برانگیز است و پژوهشگران برجسته ی هوش مصنوعی دنیا نه تنها در پیش بینی ها بلکه در عکس العمل های احساسی خود، از خوش بینی مطمئن گرفته تا بدبینی جدی، نیز به شدت با هم اختلاف نظر دارند. آنان به پرسش هایی درباره ی تأثیر کوتاه مدت هوش مصنوعی بر اقتصاد، قانون و ارتش اتفاق آرا ندارند و آنگاه که افق زمان را گسترش می دهیم و از هوش مصنوعی کلی، به ویژه آن که هوش او به انسان و فراتر می رسد که همان حیات ۳/۰ است می برسیم، این اختلاف نظرها عمیق تر می شوند. هوش کلی، برخلاف هوش محدود برنامه ی بازی شطرنج، قادر است هر کاری، از جمله یادگیری، را انجام دهد.

جالب است که بدانیم مناقشه بر سر حیات ۳/۰ به جای یک پرسش، حول دو پرسش دور می زند. یکی این که این مرحله ی حیات، اگر عملی شود، کی رخ خواهد داد؟ و دوم این که معنی آن برای انسانیت چه خواهد بود؟ از دیدگاه من سه مکتب فکری هست که باید آن ها را جدی گرفت، چون هر کدام از آن ها شامل گروهی از افراد صاحب نام است. در شکل ۲.۱ من این گروه ها را به ترتیب شهروندان آرمانی دیجیتال، تکنولوژی هراس، و اعضای جنبش هوش مصنوعی سودمند نامیده ام.

آرمان شهروندان دیجیتالی

آنگاه که کودکی بیش نبودم تصور می کردم که میلیاردرها فخر فروش و خودپسند هستند.



اگر هوش مصنوعی آبرانسان پدیدار شود، خوب خواهد بود؟

شکل ۲.۱ بیشتر مناقشات بر سر هوش مصنوعی قوی (آن که می‌تواند با انسان برابری کند) حول دو پرسش دور می‌زند. اگر هوش مصنوعی قوی بخواهد رخ دهد، چه زمانی خواهد بود و آیا برای انسانیت چیز خوبی به ارمغان خواهد آورد؟ تکنولوژی‌های هراسناک و شهروندان آرمانی دیجیتال هر دو می‌گویند که نباید نگران بود اما دلایل آن‌ها با هم کاملاً فرق دارد: دسته‌ی اول معتقدند که هوش مصنوعی کلی که با هوش انسان برابری کند به این زودی‌ها رخ نمی‌دهد، حال آن که دسته‌ی دوم معتقدند که رخ خواهد داد و به نفع بشریت خواهد بود. دسته‌ی جنبش هوش مصنوعی درست فکر می‌کنند که نگرانی درباره‌ی آن به‌جا و مفید است، زیرا پژوهش نکات ایمنی هوش مصنوعی و بحث درباره‌ی آن احتمال پیامدهای خوب آن را افزایش می‌دهد. مخالفان تکنولوژی‌های جدید پیامدهای هوش مصنوعی را ناگوار می‌دانند و با آن مخالف هستند. بخشی از این شکل را از تیم اوربان الهام گرفته‌ام. [۱]

در سال ۲۰۰۸ که لری پیچ را در گوگل ملاقات کردم، او کل این کلیشه‌ها را خرد و خمیر کرده بود، شلوار جین و پیراهن معمولی به تن داشت که گویی می‌خواهد به یکی از سفرهای تفریحی MIT برود. شیوه‌ی گفتار نرم فکورانه و لبخند دوستانه‌اش به من آرامش داد و فکر نکردم که از حرف زدن با او وحشت دارم. روز ۱۸ جولای

سال ۲۰۱۵ در میهمانی الون ماسک و همسرش تالولا در ناپاولی به او برخوردیم و بحث ما به علاقه‌ی فرزندانمان به حرف‌های رکبیک کشید که من کتاب کلاسیک ادبی روزی که چوبدستی من روانی شد، اثر اندی گریفیس را توصیه کردم و لری درجا آن را سفارش داد. داشتم فکر می‌کردم که چه بسا نام لری به عنوان یکی از تأثیرگذارترین انسان‌هایی که تاکنون زیسته‌اند در تاریخ ثبت خواهد شد و حدس می‌زنم که اگر در طول عمر من، حیات دیجیتالی ابرهوشمندی جهان ما را فرا بگیرد، به خاطر تصمیماتی خواهد بود که لری گرفته است.

به اتفاق همسرانمان، لوسی و میا، شام خوردیم و بحث‌مان به آنجا کشید که آیا ماشین‌ها الزاماً خودآگاه خواهند شد یا نه که لری اعتقاد داشت موضوعی گمراه‌کننده است. اواخر شب پس از این که کوکتل نوشیدیم مناظره‌ای طولانی و الهام‌بخش میان لری و الون درگرفت که درباره‌ی آینده‌ی هوش مصنوعی بود و این که چه باید کرد. ساعت از نیمه شب می‌گذشت که مردم زیادی اطراف‌مان جمع شده بودند و مدام تعدادشان رو به افزایش بود. لری جانانه از دیدگاهی که من دوست دارم آن را آرمان‌شهروندی دیجیتالی بنامم، دفاع کرد و آن نوعی حیات دیجیتالی است که گام طبیعی و مطلوب بعدی در تکامل کیهانی است و آنگاه که جلو افکار دیجیتالی خود را آزاد بگذاریم و محدود نکنیم، پیامدهای آن قطعاً خوب خواهند بود. من لری را بانفوذترین نمود آرمان‌شهروندی دیجیتالی می‌دانم. وی استدلال کرد که اگر قرار باشد حیات در کهکشان ما و فراتر گسترش یابد که به گمان او چنین خواهد شد، آنگاه ناچار است به شکل دیجیتالی باشد. نگرانی عمده‌ی لری این بود که هوش مصنوعی بدگمانی پیدایش آرمان‌شهر دیجیتالی را به تعویق خواهد انداخت یا زمام امور هوش مصنوعی را به ارتش خواهد سپرد که با شعار گوگل «شیطان نباش» در تناقض است. الون مدام بر لری می‌تاخت و از او می‌خواست جزئیات بحث خود را روشن کند و برای مثال بگوید که از کجا تا این اندازه مطمئن است که حیات دیجیتالی همه‌ی دلبستگی‌های ما را نابود نخواهد ساخت. لری هر از گاه الون را به «گونه‌پرستی» متهم می‌کرد که گونه‌های خاصی از حیات را دون می‌شمرد چرا که آن‌ها به جای این کربن‌پایه باشند، سیلیکون‌پایه هستند. از فصل ۴ به بعد، به کندوکاو این موضوعات و استدلال‌های جالب خواهیم پرداخت.

گرچه لری آن شب گرم تابستانی در کنار استخر در اقلیت بود، اما آن آرمان‌شهروندی دیجیتالی که او با بلاغت قهرمان آن بود، حامیان زیاد گردن‌کلفتی دارد. هانس موراوک که روایت‌گرا و آینده‌گرا بود با کتاب ماندگار سال ۱۹۸۸ به نام مواظب

کودکان باشید الهام‌بخش نسلی از آرمان‌شهروندان دیجیتالی شد، سنتی که مخترعی به نام ری کروتسویل آن را دنبال و اصلاح کرد. ریچارد ساتون، یکی از پیشگامان زیرشاخه‌ای از هوش مصنوعی که یادگیری تقویتی نام دارد، در کنفرانس پورتو ریکو که به زودی آن را شرح خواهیم داد، از آرمان‌شهروندی دیجیتالی جانانه دفاع کرد.

تکنولوژی‌هراس‌ها

گروه دیگری از متفکران نیز هستند که به دلیل کاملاً متفاوتی دغدغه‌ی هوش مصنوعی ندارند. این گروه می‌گوید ساختن هوش مصنوعی کلی آبرسان به قدری سخت است که تا صد سال آینده رخ نخواهد داد و بنابراین ساده‌لوحانه است از حالا به آن فکر کنیم. من این دیدگاه را تکنولوژی‌هراس می‌نامم که آندرو نگ آن را به شیوایی چنین بیان می‌کند: «ترس از پیدایش روبات‌های قاتل مثل این است که نگران ازدیاد جمعیت در سیاره‌ی مریخ باشیم.» آندرو دانشمند ارشد بایدو، گوگل چین، بود و اخیراً که در کنفرانسی در بوستون از او نظرش را جویا شدم، همین حرف خود را تکرار کرد. به گمان او نگران خطر هوش مصنوعی بودن چه بسا پیشرفت در زمینه‌ی هوش مصنوعی را کند کند. دیگر تکنولوژی‌هراس‌ها چون رادنی بروکس، استاد پیشین MIT که سرپرست گروه سازنده‌ی جاروبرقی روباتیک رومبا و روبات صنعتی باکستر بود نیز اظهارنظر مشابهی کرده‌اند. برای من جالب است که هرچند آرمان‌شهروندان دیجیتالی و تکنولوژی‌هراس‌ها هر دو می‌پذیرند که ما نباید نگران هوش مصنوعی باشیم اما روی مسائل دیگر توافق ندارند. بیشتر آرمان‌شهروندان دیجیتالی برآنند که هوش مصنوعی هم‌تراز با هوش انسان طی بیست تا صد سال آینده محقق خواهد شد اما تکنولوژی‌هراس‌ها آن را خیالی واهی می‌دانند و به تمسخر غیب‌گویی شگفت‌انگیزش می‌خوانند که از «شور و شوق علاقه‌مندان کامپیوتر» نشأت گرفته است. در ماه دسامبر ۲۰۱۴ که رادنی بروکس را در یکی از میهمانی‌های جشن تولد دیدم به من گفت که او صددرصد یقین دارد هوش مصنوعی در طول عمر من به وقوع نخواهد پیوست. در ایمیلی که بعد از این میهمانی به او زدم پرسیدم «آیا مطمئنی که منظورت نود و نه درصد نبود؟» و او در پاسخ نوشت «نه بُردل نود و نه درصدی. منظورم همان صددرصد بود. بدون که اتفاق نمی‌افتد.»

جنبش هوش مصنوعی سودمند

نخستین بار که در ماه ژوئن سال ۲۰۱۴ استوارت راسل را در یکی از قهوه‌خانه‌های