

# زندگی ۳/۰

انسان بودن در عصر هوش مصنوعی

مکس تگمارک

(استاد MIT)

ترجمه

میشم محمد امینی

(عضو هیئت علمی دانشگاه شهید بهشتی)

فرهنگ‌نشر نو

با همکاری نشر آسیم

تهران - ۱۳۹۷

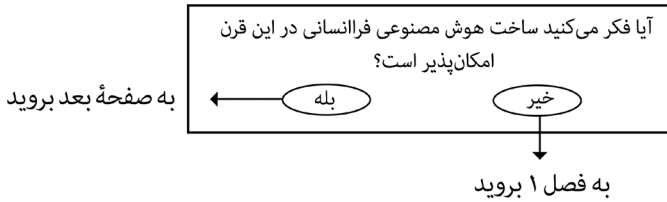
## فهرست

۱۱	سپاسگزاری	
۱۵	پیش درآمد: داستان گروه اُمگا	
۴۷	به مهم‌ترین بحث زمانه ما خوش آمدید	۱
۴۸	تاریخچه مختصری از پیچیدگی	
۵۰	سه مرحله زندگی	
۵۸	بحث‌ها و اختلاف نظرها	
۷۰	برداشت‌های نادرست	
۸۱	مسیر پیش رو	
۸۹	ماده هوشمند می‌شود	۲
۹۰	هوش چیست؟	
۹۸	حافظه چیست؟	
۱۰۶	محاسبه چیست؟	
۱۲۱	یادگیری چیست؟	

- ۳
- آینده نزدیک: پیشرفت‌های شگرف، .... ۱۳۹
- پیشرفت‌های شگرف ۱۴۱
- ایرادهای نرم‌افزاری و هوش مصنوعی مستحکم ۱۵۷
- قوانین ۱۷۵
- تسلیحات ۱۸۳
- شغل و دستمزد ۱۹۷
- هوش سطح انسانی؟ ۲۱۵
- ۴
- انفجار هوش؟ ۲۲۳
- تمامیت‌خواهی ۲۲۶
- پرومتئوس جهان را در اختیار می‌گیرد ۲۲۹
- پیشرفت آهسته و سناریوهای چندقطبی ۲۴۷
- سایبورگ‌ها و آپلودها ۲۵۳
- واقعاً چه رخ خواهد داد؟ ۲۵۷
- ۵
- پیامدها در ۱۰,۰۰۰ سال آینده ۲۶۳
- آرمانشهر آزادی‌باورانه ۲۶۸
- دیکتاتور خیرخواه ۲۷۵
- آرمانشهر برابری‌خواهانه ۲۸۲
- نگهبان ۲۸۷
- خدای حامی ۲۸۹
- خدای اسیر ۲۹۲
- فاتحان ۳۰۰
- فرزندان جدید ۳۰۶
- مسئول باغ وحش ۳۱۰
- ۱۹۸۴ ۳۱۱
- بازگشت به عقب ۳۱۴
- خودویرانگری ۳۱۷
- شما چه می‌خواهید؟ ۳۲۵

۳۲۹	استعداد کیهانی ما: چند میلیارد سال آینده و فراتر از آن	۶
۳۳۲	استفادهٔ بیشینه از منابعمان	
۳۵۲	کسب منابع بیشتر از راه استقرار در کیهان	
۳۷۴	سلسله‌مراتب کیهانی	
۳۹۱	چشم‌انداز	
۳۹۷	اهداف	۷
۳۹۸	فیزیک: منشأ اهداف	
۴۰۲	زیست‌شناسی: تکامل اهداف	
۴۰۶	روان‌شناسی: تعقیب اهداف و شورش در برابرشان	
۴۰۸	مهندسی: برون‌سپاری اهداف	
۴۱۲	هوش مصنوعی انسان‌دوست: همسو کردن اهداف	
۴۲۶	اخلاق: گزینش اهداف	
۴۳۵	اهداف غایی	
۴۴۵	آگاهی	۸
۴۴۶	چه کسی اهمیت می‌دهد؟	
۴۴۷	آگاهی چیست؟	
۴۴۹	مسئله چیست؟	
۴۵۳	آیا آگاهی فراتر از قلمرو علم است؟	
۴۵۸	سرنخ‌های تجربی دربارهٔ آگاهی	
۴۶۹	نظریه‌های آگاهی	
۴۷۹	مناقشات در باب آگاهی	
۴۸۴	آگاهی هوش مصنوعی چگونه احساسی می‌تواند داشته باشد؟	
۴۹۱	معنا	
۴۹۷	سخن پایانی: داستان گروه مؤسسهٔ آیندهٔ زندگی	
۵۲۷	یادداشت‌ها	

## زندگی ۳/۰



## داستان گروه امگا

گروه امگا چشم و چراغ شرکت بود. بخش‌های دیگر از طریق کاربردهای مختلف تجاری از هوش مصنوعی محدود<sup>۱</sup> درآمد لازم برای ادامه کار را نصیب شرکت می‌کردند، اما گروه امگا برای دستیابی به چیزی می‌کوشید که رؤیای همیشگی مدیر عامل بود: ساخت هوش مصنوعی عمومی<sup>۲</sup>. بیشتر کارکنان شرکت اعضای این گروه، یا به تعبیر خودمانی‌تر خودشان «امگایی‌ها» را یک مشت آدم بسیار خیالپرداز می‌دانستند که همیشه چند دهه از هدف خود فاصله دارند. با این حال همه خواسته‌های آنها را با روی گشاده می‌پذیرفتند، چون از حیثیت و اعتباری که کارهای پیش‌تازانه امگایی‌ها به شرکتشان می‌داد خوششان می‌آمد، و همین‌طور بابت الگوریتم‌های پیشرفته‌ای که امگایی‌ها گاه‌گاه در اختیارشان می‌گذاشتند قدردان آنها بودند.

اما چیزی که سایر کارکنان شرکت نمی‌دانستند این بود که امگایی‌ها این تصویر ساختگی از خودشان را هوشیارانه شکل داده بودند تا رازی را پنهان سازند: آنها به تحقق جسورانه‌ترین برنامه در تاریخ بشریت بی‌اندازه

1. narrow AI

2. general artificial intelligence

نزدیک بودند. مدیرعامل پرجذب‌ه‌شان کسانی را برای عضویت در این گروه دست‌چین کرده بود که نه فقط بهره‌ هوشی بسیار بالایی داشتند، بلکه بسیار بلندپرواز و آرمان‌خواه و متعهد به یاری بشریت بودند. مدیرعامل همیشه به آنها تأکید می‌کرد که برنامه‌شان فوق‌العاده خطرناک است و اگر دولت‌های قدرتمند از آن آگاهی یابند، تقریباً به هر کاری - از جمله آدم‌ربایی - دست می‌زنند تا آنها را متوقف کنند، و حتی ترجیح‌شان این است که کد برنامه‌هایشان را بدزدند. اما چون هیچ‌کدام جا نزدند و همه برای فعالیت در این زمینه اعلام آمادگی کردند، دقیقاً به همان دلیلی که بسیاری از فیزیکدانان برجسته جهان برای همکاری در ساخت سلاح‌های هسته‌ای به پروژه مَنهتن پیوستند: چون متقاعد شده بودند که اگر خودشان این کار را اول انجام ندهند، کسان دیگری انجام خواهند داد که شاید چنین اهداف و آرمان‌های والایی نداشته باشند.

برنامه هوش مصنوعی آنها، که نام پرومیتئوس بر آن گذاشته بودند، پیوسته پیشرفت می‌کرد و توانایی‌های بیشتری به دست می‌آورد. اگرچه قابلیت‌های شناختی‌اش در بسیاری از حوزه‌ها، مانند مهارت‌های اجتماعی، هنوز خیلی عقب‌تر از انسان‌ها بود، اما امگایی‌ها خیلی زحمت کشیده بودند تا آن را در انجام یک وظیفه خاص فوق‌العاده توانمند سازند؛ یعنی در برنامه‌ریزی سامانه‌های هوش مصنوعی. آنها آگاهانه چنین راهکاری را انتخاب کرده بودند، زیرا استدلال انفجار هوش را که ریاضیدان بریتانیایی اروینگ گود<sup>۱</sup> در ۱۹۶۵ مطرح کرده بود پذیرفته بودند. استدلال گود از این قرار است: «ماشین فوق‌هوشمند را ماشینی تعریف می‌کنیم که بتواند در همه فعالیت‌های فکری، از هر انسانی، هر اندازه هم که هوشمند باشد، بسیار فراتر عمل کند. طراحی ماشین‌ها هم یکی از همین فعالیت‌های فکری است، پس یک ماشین فوق‌هوشمند می‌تواند ماشین‌هایی طراحی کند که

حتی از خودش هم بهتر باشند؛ و بنابراین بی‌تردید «انفجار هوش» رخ می‌دهد، و هوش ماشین‌ها از سطح هوش انسان بسیار فراتر خواهد رفت. بنابراین نخستین ماشین فوق‌هوشمند آخرین اختراعی است که انسان به آن نیاز خواهد داشت، [پس از این اختراع، هیچ اختراع دیگری لازم نیست]؛ البته با فرض آنکه این ماشین به اندازه‌ای مطیع انسان باشد که به ما بگوید چگونه باید بر آن تسلط داشته باشیم».

اعضای تیم امگا دریافتند که اگر بتوانند این فرایند بازگشتی پیشرفت را به راه اندازند، آنگاه ماشین خیلی زود به اندازه‌ای هوشمند خواهد شد که می‌تواند همه مهارت‌های بشری مفید دیگر را به خودش آموزش دهد.

## چند میلیون اول

ساعت نُه صبح جمعه بود که آنها تصمیم گرفتند شروع کنند. پرومیتوس در یک محفظه مخصوص در ردیف طولانی قفسه‌ها در اتاقی بزرگ با دسترسی محدود و تهویه مطبوع مشغول کار شد. به دلایل امنیتی، اتصال آن به اینترنت را کاملاً قطع کرده بودند، اما درون حافظه برنامه یک کپی از بخش عمده محتوای شبکه جهانی (از جمله ویکی‌پدیا، کتابخانه کنگره، توئیتر، گلچینی از یوتیوب، بخش عمده‌ای از فیس‌بوک، و غیره) وجود داشت تا از این داده‌های آموزشی برای یادگیری برنامه استفاده شود.<sup>۱</sup> اعضای گروه این زمان را برای شروع برنامه انتخاب کرده بودند تا بتوانند بدون مزاحم به کارشان ادامه دهند: خانواده‌ها و دوستانشان فکر می‌کردند

۱. برای سادگی، در این داستان شرایط فعلی اقتصاد و فناوری را پیش‌فرض گرفته‌ام، هرچند که بیشتر پژوهشگران حدس می‌زنند که دست کم چند دهه با هوش مصنوعی عمومی در سطح بشر فاصله داریم. اجرای برنامه امگا در آینده راحت‌تر هم خواهد بود اگر اقتصاد دیجیتال به رشد خود ادامه دهد و دامنه خدماتی که به صورت آنلاین و بدون هیچ پرسشی عرضه می‌شود گسترش یابد.



## به مهم‌ترین بحث زمانه ما خوش آمدید

فناوری به حیات قابلیت آن را داده است که به چنان سطحی از شکوفایی برسد که پیشتر هرگز سابقه نداشته است - یا خود را به کلی نابود کند.

مؤسسه آینده‌زندگی

جهان ما سیزده میلیارد و هشتصد میلیون سال پس از پیدایش بیدار شده و نسبت به خودش آگاهی یافته است. بخش‌های خودآگاه بسیار کوچکی از جهان ما، از یک سیاره کوچک آبی‌رنگ، با تلسکوپ به کیهان نگریده‌اند و بارها کشف کرده‌اند که تمام چیزهایی که پیشتر تصور می‌کردند وجود دارد فقط بخش کوچکی از چیزی بزرگ‌تر است: یک منظومه خورشیدی، یک کهکشان، و جهانی با بیش از صد میلیارد کهکشان دیگر که در قالب الگویی پیچیده از گروه‌ها و خوشه‌های ستاره‌ای و آبرخوشه‌ها نظام یافته است. اگرچه این اخترشناسان خودآگاه درباره بسیاری مسائل اختلاف نظر دارند، اما تقریباً همه می‌پذیرند که این کهکشان‌ها زیبا و با بهت‌اند.

البته زیبایی در چشم ناظر است، نه در قوانین فیزیک؛ پس پیش از آنکه جهان ما بیدار شود، هیچ زیبایی‌ای وجود نداشت. همین سبب می‌شود بیداری کیهانی ما امری حیرت‌انگیز و مایه شادمانی باشد؛ به این ترتیب جهان ما از یک مرده متحرک بی‌شعور و فاقد آگاهی به زیست‌بومی زنده بدل شد که تأمل درباره‌ی خویشتن، زیبایی، و امید - و دنبال کردن هدف و معنا و مقصود - را در دل می‌پروراند. اگر جهان ما هرگز بیدار نشده بود، از نگاه من کاملاً بی‌معنا و بیهوده می‌بود - چیزی نمی‌بود جز هدر دادن فضا در ابعاد عظیم. اگر جهان ما بر اثر فاجعه‌ای کیهانی یا دردسری خودساخته، دوباره و برای همیشه به خواب برود، متأسفانه بی‌معنا خواهد شد.

از سوی دیگر، وضع از اینکه هست بهتر هم می‌تواند بشود. هنوز نمی‌دانیم که آیا ما انسان‌ها تنها موجودات، یا حتی نخستین موجودات اخترشناس در کیهانمان هستیم یا نه، اما به اندازه‌ای درباره‌ی جهانمان آموخته‌ایم که بدانیم قابلیت آن را دارد که بسیار بیش از آنچه اینک می‌بینیم بیدار شود. شاید بتوان گفت حال کنونی ما شبیه نخستین جرقه کمرنگ خودآگاهی است که امروز صبح وقتی که از خواب بیدار شدید حس کردید: نوعی پیش‌آگاهی از یک آگاهی بسیار بزرگ‌تر که وقتی چشمانتان را باز کردید و کاملاً خودآگاه شدید فرا می‌رسد. شاید حیات میلیاردها یا تریلیون‌ها سال در سراسر کیهان ما گسترش یابد و ببالد - و شاید علت این امر تصمیم‌هایی باشد که ما در سیاره کوچکمان و در دوران زندگی مان اتخاذ می‌کنیم.

### تاریخچه مختصری از پیچیدگی

خب، این بیداری حیرت‌انگیز چگونه رخ داد؟ یک رویداد واحد و منفرد نبود، بلکه فقط یک گام از فرایند بی‌وقفه ۱۳/۸ میلیاردساله‌ای بود که جهان ما را پیوسته پیچیده‌تر و جالب‌تر می‌کند - و هنوز هم با سرعتی شتابنده ادامه دارد. به عنوان فیزیکدان، احساس خوشبختی می‌کنم که در طی ربع قرن

گذشته فرصت آن را داشته‌ام که به مشخص کردن تاریخ کیهانی مان کمک کنم، و پیمودن این مسیر سفر اکتشافی بسیار شگفتی‌آوری بوده است. زمانی که دانشجوی مقطع کارشناسی بودم، در این باره بحث می‌کردیم که آیا جهان ما ۱۰ میلیارد ساله است یا ۲۰ میلیارد ساله، اما امروز با بهره‌مندی از آمیزه‌ای از تلسکوپ‌های بهتر و رایانه‌های بهتر و درک بهتر، در این باره بحث می‌کنیم که سن جهان ۱۳/۷ میلیارد سال است یا ۱۳/۸ میلیارد سال. ما فیزیکدان‌ها هنوز دقیقاً نمی‌دانیم چه چیزی علت مهبانگ شد، یا اینکه این رویداد به‌راستی آغاز همه‌چیز بود یا اینکه صرفاً دنباله مرحله‌ای قبلی بود. اما درک نسبتاً کاملی داریم از اینکه پس از مهبانگ چه اتفاقاتی رخ داده است، و این در نتیجه حجم عظیمی از اندازه‌گیری‌های بسیار دقیق میسر شده است، پس خواهش می‌کنم به من اجازه دهید که در چند دقیقه، ۱۳/۸ میلیارد سال تاریخ کیهانی را خلاصه کنم.

در آغاز، نور بود. در نخستین کسر از ثانیه پس از مهبانگ ما، کل آن بخش از فضا که تلسکوپ‌های ما علی‌الاصول می‌توانند مشاهده کنند («جهان قابل مشاهده ما» یا به اختصار «جهان ما») بسیار داغ‌تر و روشن‌تر از هسته خورشید ما و به‌سرعت در حال انبساط بود. شاید مسئله بسیار چشمگیر و خارق‌العاده به نظر آید، اما از جهتی خیلی هم ملال‌آور بود، چون جهان ما چیزی نبود مگر سوپ خالی از حیات و چگال و داغ ذرات بنیادین، و کاملاً یکنواخت و تهی از هرگونه جذابیت. همه‌جا وضع تقریباً یکسانی حاکم بود، و یگانه ساختار جذاب امواج صوتی ضعیف و ظاهراً تصادفی‌ای بود که این سوپ را در بعضی نقاط در حدود ۰/۰۰۱ درصد چگال‌تر می‌کرد. عموماً تصور بر این است که همین امواج ضعیف منشأ به‌اصطلاح نوسانات کوانتومی بوده‌اند، چون اصل عدم قطعیت هایزنبرگ مانع از آن می‌شود که همه‌چیز کاملاً ملال‌آور و یکنواخت باشد.

هرچه جهان منبسط‌تر و سردتر می‌شد، جالب‌تر هم می‌شد، چون ذرات آن پیوسته با یکدیگر ترکیب می‌شدند و چیزهایی پیچیده‌تر پدید می‌آوردند.

در نخستین لحظه، نیروی هسته‌ای قوی از پیوند کوارک‌ها با یکدیگر پروتون (هسته هیدروژن) و نوترون پدید آورد، و بعضی از این پروتون‌ها و نوترون‌ها هم در عرض چند دقیقه به هم پیوستند و هسته‌های هلیوم به وجود آوردند. تقریباً ۴۰۰,۰۰۰ سال بعد، نیروی الکترومغناطیسی این هسته‌ها را در کنار الکترون‌ها قرار داد و به این ترتیب نخستین اتم‌ها ساخته شدند. با ادامه روند انبساط جهان، این اتم‌ها به تدریج سرد شدند و یک گاز تاریک سرد را به وجود آوردند، و تاریکی این شب نخست حدود ۱۰۰ میلیون سال طول کشید. هنگامی که نیروهای گرانشی توانستند با تشدید نوسانات درون آن گاز، اتم‌ها را کنار هم جمع کنند و نخستین ستاره‌ها و کهکشان‌ها را به وجود آورند این شب بلند سرآغاز پیدایش کیهان ما شد. این ستاره‌های نخستین به واسطه همجوشی هیدروژن و پیدایش اتم‌های سنگین‌تر مانند کربن و اکسیژن و سیلیسیم، گرما و نور تولید کردند. با مرگ این ستاره‌ها، بسیاری از اتم‌هایی که ساخته بودند درون کیهان بازیابی شدند و سیاره‌هایی را که دور ستاره‌های نسل دوم می‌گشتند پدید آوردند.

در یک مرحله، گروهی از اتم‌ها در قالب الگویی پیچیده سامان یافتند که می‌توانست خود را حفظ و تکثیر کند. خیلی زود دو نمونه از این الگو به وجود آمد، و شمار آنها پیوسته دوبرابر می‌شد. فقط چهل بار دوبرابر شدن کافی بود تا شمار این نمونه‌ها به یک تریلیون برسد، پس این نخستین موجود خودتکثیرگر به سرعت به نیرویی قابل توجه تبدیل شد. حیات از راه رسیده بود.

### سه مرحله زندگی

مسئله تعریف حیات محل بحث و مناقشه فراوان است و از این حیث شهرت سوئی یافته است. تعاریف مختلف و فراوانی ارائه شده که برخی از آنها شامل شرایطی بسیار مشخص و محدودکننده مانند متشکل بودن

## ماده هوشمند می شود

هیدروژن...، اگر زمان کافی داشته باشد، سرانجام به انسان تبدیل می‌شود.

ادوارد رابرت هریسن<sup>۱</sup>

در طول ۱۳/۸ میلیارد سالی که از مهبانگ می‌گذرد، یکی از چشمگیرترین تحولاتی که رخ داده این است که ماده گنگ و بی‌جان هوشمند شده است. چطور چنین چیزی امکان‌پذیر شد و موجودات در آینده تا چه اندازه می‌توانند باهوش‌تر شوند؟ علم درباره سرگذشت و سرنوشت هوش در کیهان ما چه چیزی برای گفتن دارد؟ این فصل را به بررسی مبانی و بنیادهای سازنده هوش اختصاص می‌دهیم، تا در ادامه بهتر بتوانیم به این پرسش‌ها پردازیم. اینکه می‌گوییم توده‌ای از ماده هوشمند است چه معنایی دارد؟ یعنی چه که یک شیء می‌تواند به یاد آورد، محاسبه کند، و یاد بگیرد؟

1. Edward Robert Harrison

## هوش چیست؟

من و همسر من به تازگی بخت آن را داشتیم که در همایشی دربارهٔ هوش مصنوعی که بنیاد نوبل سوئد ترتیب داده بود شرکت کنیم، و آنجا وقتی که از گروهی از محققان برجستهٔ هوش مصنوعی خواسته شد هوش را تعریف کنند، بحثی دامنه‌دار درگرفت و در نهایت هم اجماعی حاصل نشد. این قضیه برایمان خیلی عجیب بود: حتی میان پژوهشگران باهوش متخصص هوش، اتفاق نظری دربارهٔ اینکه هوش چیست وجود ندارد! پس روشن است که هیچ تعریف «صحیح» بی‌مناقشه‌ای از هوش وجود ندارد. در عوض، بسیاری تعاریف مختلف هست مانند قابلیت انجام عملیات منطقی، فهم، برنامه‌ریزی، شناخت عاطفی، خودآگاهی، خلاقیت، حل مسئله، و یادگیری.

در بررسی خود دربارهٔ آیندهٔ هوش می‌خواهیم دیدگاهی تا حد امکان گسترده و جامع برگزینیم که هوش را فقط به شکل‌هایی از هوش که تا به حال وجود داشته است محدود نکند. به همین دلیل تعریفی که در فصل پیش به دست دادم و همین‌طور نحوهٔ کاربرد من از این واژه در طول کتاب، بسیار گسترده است:

هوش = توانایی دستیابی به اهداف پیچیده

این تعریف به اندازه‌ای کلی هست که همهٔ تعاریف پیشتر یادشده را دربربگیرد، چون فهم و خودآگاهی و حل مسئله و یادگیری و غیره همه نمونه‌هایی از اهداف پیچیده‌ای هستند که موجود هوشمند می‌تواند داشته باشد. همچنین تعریف من به اندازه‌ای کلی هست که این تعریف فرهنگ آکسفورد را نیز شامل شود: «توانایی کسب و به‌کارگیری دانش و مهارت»

– چون ممکن است هدف موجود هوشمند به کارگیریِ دانش و مهارت باشد.

اهدافِ ممکنِ متعددی وجود دارد، بنابراین انواعِ ممکنِ متعددی از هوش نیز وجود دارد. از این رو بر پایهٔ تعریف ما، کمی‌سازیِ هوش انسان‌ها یا حیوانات غیرانسان یا ماشین‌ها با یک عدد واحد مانند IQ معقول نیست.<sup>۱</sup> کدامیک از این دو باهوش‌ترند: برنامهٔ رایانه‌ای که فقط می‌تواند شطرنج بازی کند، یا برنامه‌ای که فقط می‌تواند گو بازی کند؟ پاسخ به این پرسش ممکن نیست، چون این دو برنامه در انجام دو کار مختلف توانایی دارند که به طور مستقیم با هم قابل‌قیاس نیستند. اما می‌توانیم بگوییم برنامهٔ سومی که در دستیابی به همهٔ اهداف دست کم به اندازهٔ آن دو برنامه توانا است و در دست کم یک کار (مثلاً در شطرنج) آشکارا بهتر است، از هر دو برنامه باهوش‌تر است.

همچنین بحث و جدل دربارهٔ هوشمند بودن یا نبودن چیزی در موارد بینابینی چندان عاقلانه نیست، چون توانایی یک ویژگیِ مدرج یا طیفی است و حالت همه یا هیچ ندارد. چه کسانی توانایی دستیابی به هدفِ سخن گفتن را دارند؟ نوزادان؟ خیر. مجریان رادیو؟ بله. در مورد خردسالانی که می‌توانند ده کلمه بگویند چه می‌توان گفت؟ یا پانصد کلمه؟ مرز را کجا ترسیم می‌کنید؟ من عمداً از واژهٔ مبهم «پیچیده» در تعریف بالا استفاده کرده‌ام، چون کار زیاد جالبی نیست که بخواهیم به شکلی مصنوعی مرزی میان هوش و غیرهوش بکشیم، بلکه بهتر است درجهٔ تواناییِ دستیابی به اهداف مختلف را با معیاری کمی بیان کنیم.

برای دسته‌بندی انواع مختلف هوش تمایز اساسی دیگر تمایز میان هوش محدود و گسترده است. رایانهٔ شطرنج‌باز دیپ‌بلو ساخت IBM

۱. برای درک بهتر این موضوع تصور کنید کسی ادعا کند که توانایی انجام حرکات ورزشی در سطح قهرمانان المپیک را می‌توان فقط با یک عدد به نام «ضریب ورزشی» (AQ) کمی کرد، بنابراین ورزشکاری که بالاترین AQ را داشته باشد در تمام رشته‌های ورزشی مدال طلا خواهد گرفت.

## آینده نزدیک: پیشرفت‌های شگرف، خطاهای نرم‌افزاری، قوانین، سلاح‌ها، مشاغل

اگر جهت حرکت را زود عوض نکنیم، سرانجام به همان جایی  
که می‌رویم خواهیم رسید.

اروین کوری<sup>۱</sup>

انسان بودن در عصر و دوره کنونی چه معنایی دارد؟ مثلاً آن ارزشی که واقعاً  
برای خود قائلیم و ما را از دیگر موجودات زنده و همین‌طور از ماشین‌ها  
متفاوت می‌کند چیست؟ افراد دیگر برای چه چیزی در ما ارزش قائل هستند  
که سبب می‌شود برخی از آنها به ما پیشنهاد کار بدهند؟ هر پاسخی که برای  
این پرسش‌ها در یک زمان مشخص داشته باشیم، روشن است که ظهور  
فناوری ناگزیر این پاسخ‌ها را به تدریج تغییر می‌دهد.

مثلاً خود من را در نظر بگیرید. من در مقام دانشمند به این می‌بالم  
که می‌توانم هدف‌هایم را خودم مشخص کنم، به این می‌بالم که می‌توانم  
با خلاقیت و شهود به گستره وسیعی از مسائل حل‌نشده پردازم، و به



این می‌بالم که می‌توانم زبان را برای بیان و انتقال آنچه کشف می‌کنم به کار بگیرم. برای من مایه خوشوقتی است که جامعه حاضر است در ازای این کار به من پول بدهد. صدها سال پیش، احتمالاً من هم مانند بسیاری دیگر، به جای این کار به کشاورزی یا صنعتگری مشغول بودم و همین را اساس هویت خود می‌دانستم، اما اکنون رشد فناوری سبب شده است که این‌گونه مشاغل فقط کسر بسیار کوچکی از نیروی کار را به خود اختصاص دهد. این یعنی دیگر ممکن نیست که همه هویتشان را بر پایه کشاورزی یا صنعتگری برسانند.

برای شخص من ناراحت‌کننده نیست که امروز ماشین‌ها در کارهاییدی مانند حفاری و بافندگی ماهرترند، چون اینها نه جزء علائق من است، نه منبع درآمد یا عزت نفسم. راستش، هرگونه توهمی که ممکن بود درباره توانایی‌هایم در این زمینه داشته باشم وقتی که هشت سالم بود کاملاً از میان رفت. در مدرسه مجبورم کردند یک درس بافندگی بگذرانم، اما من به‌زحمت نمره قبولی گرفتم و فقط با کمک یک کلاس‌پنجمیِ مهربان که دلش به حالم سوخته بود توانستم پروژه پایانی را انجام بدهم.

اما آیا با رشد پیوسته فناوری، پیشرفت‌های هوش مصنوعی سرانجام آن توانایی‌هایی را که اکنون به من احساس عزت نفس و به کارم در بازار ارزش می‌دهد نیز تحت تأثیر قرار خواهد داد؟ استیوارت راسل به من گفت که برای او و بسیاری از همکاران پژوهشگرش در زمینه هوش اخیراً لحظاتی پیش آمده است که از شدت شگفتی فریاد بکشند، چون دیده‌اند که هوش مصنوعی می‌تواند کارهایی را انجام دهد که پیشتر اصلاً انتظار نداشتند تا چندین و چند سال قادر به انجام آنها باشد. من هم از شما اجازه می‌خواهم که به چند مورد از این لحظه‌ها که برای خودم پیش آمده است اشاره کنم، که به نظرم نشانه آن است که در چند حوزه دیگر نیز به‌زودی هوش مصنوعی به سطح توانایی‌های بشر خواهد رسید.

## پیشرفت‌های شگرف

### کنشگرهای برخوردار از قابلیت یادگیری با تقویت عمیق

یکی از دفعاتی که به‌راستی دهانم از تعجب باز مانده بود سال ۲۰۱۴ بود هنگام تماشای فیلمی از سامانه هوش مصنوعی دیپ‌مایند در حال یادگیری نحوه انجام بازی‌های رایانه‌ای. دقیق‌تر بگویم، این سامانه هوش مصنوعی در حال انجام بازی فرار یا بریک‌آت<sup>۱</sup> بود (شکل ۳.۱ را ببینید). این بازی یکی از بازی‌های قدیمی آتاری است که از دوره نوجوانی خوب آن را به خاطر دارم. هدف در این بازی حرکت دادن یک راکت است به طوری که توپ بتواند پس از برخورد با سطح راکت بالا بجهد و به دیوار آجری بالای صفحه بازی برخورد؛ هر دفعه که توپ به یک آجر برخورد، آن آجر ناپدید می‌شود و امتیاز شما افزایش می‌یابد.

من هم در گذشته‌ها چند بازی رایانه‌ای طراحی و برنامه‌نویسی کرده بودم و از همین رو خوب می‌دانستم که نوشتن برنامه‌ای که بتواند بریک‌آت بازی کند چندان دشوار نیست - اما کاری که گروه دیپ‌مایند انجام داده بودند این نبود. آنها سامانه هوش مصنوعی‌ای ساخته بودند که همانند یک لوح سفید بود و از این بازی - یا از هیچ بازی دیگر، یا حتی از مفاهیمی مانند بازی و راکت و آجر و توپ - هیچ چیز نمی‌دانست. تنها چیزی که این سامانه هوش مصنوعی می‌دانست این بود که فهرست بلندی از اعداد در بازه‌های زمانی منظم به عنوان ورودی به آن داده می‌شود: امتیاز در آن زمان مشخص و دنباله درازی از اعداد که ما (و نه هوش مصنوعی) به عنوان مشخصات صفحه بازی و رنگ قسمت‌های مختلف آن می‌شناسیم.

## انفجار هوش؟

اگر ماشینی بتواند بیندیشد، ممکن است هوشمندانه‌تر از ما بیندیشد، در این صورت ما چه جایگاهی خواهیم داشت؟ حتی اگر بتوانیم ماشین‌ها را مطیع خود نگاه داریم... حتماً در مقام یک گونه زیستی، به شدت احساس حقارت می‌کنیم.

الن تورینگ، ۱۹۵۱

نخستین ماشین فوق‌هوشمند آخرین اختراعی است که انسان‌ها به آن نیاز خواهند داشت، البته با فرض آنکه این ماشین به اندازه‌ای مطیع انسان باشد که به ما بگوید چگونه باید بر آن تسلط داشته باشیم.

اروینگ ج. گود، ۱۹۶۵

نمی‌توانیم این امکان را کاملاً نادیده بگیریم که سرانجام روزی بتوانیم هوش عمومی مصنوعی در سطح انسان بسازیم، بنابراین فصل حاضر را به بررسی این امکان و پیامدهای احتمالی آن اختصاص می‌دهیم. بحث را با اساسی‌ترین مشکل آغاز می‌کنیم:

آیا هوش مصنوعی واقعاً می‌تواند بر جهان تسلط یابد، یا انسان‌ها را قادر به این کار سازد؟

اگر وقتی که سخن از تسلط بر جهان با استفاده از ربات‌های ترمیناتور شکل تفنگ به دست به میان می‌آید ناباورانه به گوینده نگاهی می‌اندازید گویی حرفی بسیار احمقانه زده است، کاملاً حق دارید: چنین سناریویی خیلی غیرواقعی و ابلهانه است. این ربات‌های هالیوودی آنقدرها از ما باهوش‌تر نیستند و اصلاً به موفقیت هم نمی‌رسند. به گمان من، خطر داستان ترمیناتور این نیست که در آینده رخ خواهد داد، بلکه خطرش این است که توجه را از مخاطرات و فرصت‌های واقعی که هوش مصنوعی پیش روی ما قرار داده است منحرف می‌کند. برای اینکه جهان واقعاً از وضع کنونی به حالتی برسد که تحت تسلط قدرتی مجهز به هوش عمومی مصنوعی است، باید سه گام منطقی طی شود:

- گام ۱: ساخت هوش عمومی مصنوعی در سطح بشر
- گام ۲: استفاده از این هوش عمومی مصنوعی برای ساخت ابرهوش
- گام ۳: استفاده از این ابرهوش یا رهاسازی آن برای تسلط بر جهان.

در فصل پیش دیدیم که نمی‌توان وقوع گام ۱ را به عنوان چیزی که تا ابد غیرممکن است نادیده گرفت. همچنین دیدیم که اگر گام ۱ محقق شود، برداشته شدن گام ۲ هم چندان بعید نیست، چون هوش عمومی مصنوعی این قابلیت را دارد که به طور مداوم سامانه‌های هوش عمومی مصنوعی بهتری طراحی کند و تنها محدودیت در این عرصه قوانین فیزیک هستند – و به نظر می‌رسد بر پایه این قوانین دستیابی به هوش بسیار فراتر از هوش انسان امکان‌پذیر است. نهایتاً هم اینکه چون ما انسان‌ها توانسته‌ایم با هوش برتر خود بر سایر موجودات زنده زمین تسلط یابیم، پس این امکان چندان دور از ذهن نیست که موجودی ابرهوشمند نیز بتواند بر ما تسلط پیدا کند.

البته این استدلال‌ها که ادعا می‌کنند چنین چیزی امکان‌پذیر است به طرز ناامیدکننده‌ای مبهم و غیردقیق هستند، و دشواری اصلی هم در تبیین جزئیات است. آیا هوش مصنوعی در عمل می‌تواند جهان را به تصرف خود آورد؟ برای بررسی این پرسش، ترمیناتورهای ابله را فراموش کنید، و در عوض به بعضی سناریوهای دقیقی که ممکن است محقق شوند توجه کنید. سپس هریک از این داستان‌ها را به دقت و ارسی می‌کنیم و می‌کشیم نقایص و کاستی‌هایشان را بیابیم، پس لطفاً با نگاهی شکاک و منتقد آنها را بخوانید - نکته اصلی‌ای که این سناریوها نشان می‌دهند این است که ما چه اندازه درباره اینکه چه چیزی رخ خواهد داد و چه چیزی رخ نخواهد داد، ناآگاه و بی‌اطلاعم، و اینکه گستره امکانات در این زمینه چقدر وسیع است. سناریوهای نخست ما متعلق به آن سرطیف هستند که سریع‌ترین و چشمگیرترین تغییرات در آنها اتفاق می‌افتد. به عقیده من، بررسی دقیق این سناریوها کار بسیار ارزشمندی است - نه به این دلیل که لزوماً محتمل‌ترین سناریوها هستند، بلکه چون اگر به این نتیجه نرسیم که محقق شدن این وضعیت‌ها خیلی بعید است، آنگاه باید درک مناسبی از آنها داشته باشیم تا پیش از آنکه خیلی دیر شود بتوانیم اقدامات پیشگیرانه‌ای انجام دهیم و از بروز پیامدهای ناخوشایند جلوگیری کنیم.

پیش‌درآمد این کتاب سناریویی بود که در آن انسان‌ها با استفاده از ابرهوش جهان را به تصرف خود در آوردند. اگر هنوز این بخش را نخوانده‌اید، به عقب بازگردید و اکنون بخوانید. حتی اگر قبلاً خوانده‌اید، برگردید و نگاهی به آن بیندازید، تا پیش از آنکه وارد نقد و بررسی آن شویم داستان خوب در خاطرتان باشد.

\*\*\*

در ادامه خیلی زود به نقاط ضعف جدی برنامه‌امگایی‌ها خواهیم پرداخت، اما اگر یک لحظه فکر کنید که این برنامه موفقیت‌آمیز از کار درآید، چه

احساسی درباره آن دارید؟ موضوع بسیار خوبی است برای بحث‌های بعد از شام! پس از اینکه امگایی‌ها استیلای خود را بر جهان استحکام بخشیدند، چه خواهد شد؟ بستگی دارد به اینکه هدفشان چه باشد و صادقانه بگویم، نمی‌دانم چه هدفی دارند. اگر خودتان در قدرت بودید، چه نوع آینده‌ای دلخواه شما بود که به وجود آورید؟ در فصل ۵، مجموعه‌ای از گزینه‌های پیش رو را بررسی خواهیم کرد.

### تمامیت‌خواهی

فرض کنید مدیرعاملی که گروه امگا زیر نظر او کار می‌کند هدف‌های بلندمدتی نظیر اهداف آدولف هیتلر و یوسیف استالین دارد. کسی چه می‌داند، شاید واقعاً همین‌طور باشد و او این هدف‌ها را از همه پنهان نگه داشته است تا زمانی که به اندازه‌ای قدرت یابد که بتواند اهدافش را محقق کند. حتی اگر هدف‌های اولیه مدیرعامل اهداف والایی باشد، باید هشدار لرد اکتن را به یاد داشت که در سال ۱۸۸۷ گفته بود: «قدرت فساد می‌آورد، و قدرت مطلق فساد مطلق». مثلاً مدیرعامل به‌سادگی می‌تواند از پرومتئوس استفاده کند و دولتی به وجود آورد که بر شهروندانش نظارت کامل داشته باشد. عملیات فضولی دولت که ادوارد اسنودن<sup>۱</sup> افشا کرد با این هدف دنبال می‌شد که در نهایت به وضع «ثبت کامل» دست یابد - یعنی همه ارتباطات الکترونیکی برای تحلیل‌های احتمالی بعدی ضبط شود - اما پرومتئوس می‌تواند هدفی بالاتر را دنبال کند و به فهم همه ارتباط‌های الکترونیکی برسد. پرومتئوس از راه خواندن همه ایمیل‌ها و پیام‌های متنی فرستاده شده، گوش کردن به همه تماس‌های تلفنی، نگاه کردن فیلم همه دوربین‌های امنیتی و ترافیکی، تحلیل همه تراکنش‌های کارت‌های اعتباری

1. Edward Snowden



## پیامدها در ۱۰,۰۰۰ سال آینده

به سادگی می‌توان تصور کرد که اندیشه بشر از بند تن فانی رها شود - عقیده به زندگی پس از مرگ باوری رایج است. اما لازم نیست برای پذیرش این امر دیدگاهی عرفانی یا مذهبی اتخاذ کنیم. رایانه می‌تواند نمونه‌ای عالی حتی برای سفت و سخت‌ترین هواداران دیدگاه مکانیک باورانه باشد. هانس موراوک، فرزندِان ذهن

من به سهم خودم به فرمانروایان جدید رایانه‌ای مان خوشامد می‌گویم.  
کن جنینگز، هنگام باخت در مسابقه چپردی!

به رایانه واتسن ساخت IBM

انسان همانند سوسک به موجودی بی‌اهمیت تبدیل می‌شود.

مارشال برین

مسابقه برای دستیابی به هوش عمومی مصنوعی آغاز شده است، و هیچ نمی‌دانیم چه پیش خواهد آمد. اما این نباید مانع از اندیشیدن ما شود درباره اینکه می‌خواهیم آینده چگونه باشد، زیرا خواست ما بر نتیجه حاصل اثرگذار است. ترجیح شخص شما چیست، و چرا؟

۱. آیا می‌خواهید ابرهوش وجود داشته باشد؟
۲. آیا می‌خواهید که انسان‌ها باز هم وجود داشته باشند، یا موجودات دیگری جایگزینشان شوند، یا سایبورگی شوند، و/یا آلود/شبییه‌سازی شوند؟
۳. آیا می‌خواهید کنترل اوضاع دست انسان باشد یا ماشین؟
۴. آیا می‌خواهید سامانه‌های هوش مصنوعی خودآگاه باشند یا نه؟
۵. آیا می‌خواهید تجربه‌های خوشایند را بیشینه و رنج را کمینه کنید، یا می‌خواهید این مسئله را به حال خود رها کنید تا خودش چاره‌ای پیدا کند؟
۶. آیا می‌خواهید حیات در سراسر کیهان گسترش یابد؟
۷. آیا می‌خواهید تمدنی شکل بگیرد که می‌کوشد به هدفی بزرگتر دست یابد که شما با آن همدلی دارید، یا شکل‌های آینده حیات برایتان مطلوب‌اند، حتی اگر اهدافشان از نظر شما پوچ و بی‌معنی باشد؟

برای آغاز اندیشه و بحث در این باره، مجموعه گسترده‌ای از حالات ممکن را در جدول ۵.۱ به طور خلاصه آورده‌ام. بدیهی است که این فهرست جامع نیست، اما موارد را طوری انتخاب کرده‌ام که کل طیف حالات ممکن را بپوشاند. روشن است که نمی‌خواهیم در اثر ضعف برنامه‌ریزی حالت نهایی نامطلوبی بر جهان حاکم شود. توصیه می‌کنم پاسخ‌هایتان به پرسش‌های ۱-۷ را جایی یادداشت کنید و پس از خواندن این فصل، نگاهی به آن بیندازید و ببینید آیا نظرتان عوض شده است! این کار را می‌توانید در وبسایت <http://AgeOfAi.org> نیز انجام دهید. به علاوه، در این وبسایت می‌توانید از دیدگاه خوانندگان دیگر نیز آگاه شوید و به بحث و تبادل نظر بپردازید.





## استعداد کیهانی ما: چند میلیارد سال آینده و فراتر از آن

سیر پایان نظریه پردازی‌های ما سرانجام به یک ابرتمدن می‌رسد، یعنی ترکیبِ کلِ حیاتِ منظومهٔ خورشیدی که پیوسته در حال پیشرفت و گسترش است، و از خورشید به سوی فضای خارج انتشار می‌یابد و موجود غیرزنده را به ذهن تبدیل می‌کند.  
هانس موراوک، فرزندان ذهن

از نظر من، هیجان‌انگیزترین کشف علمیِ کل تاریخ این است که تخمین ما از امکانات بالقوهٔ حیات در آینده بیش از انداز ناچیز بوده است. لازم نیست سقف رؤیایها و آرزوهایمان زندگی صدساله همراه با مشکلاتی چون بیماری و فقر و سرگردانی باشد. بلکه زندگی به کمک فناوری این قوه را دارد که میلیاردها سال شکوفا باشد، و نه فقط اینجا در منظومهٔ خورشیدی، بلکه در تمام کیهانی که بسیار عظیم‌تر و شورانگیزتر از چیزی است که نیاکانمان تصور می‌کردند. برای آرزوهایمان هیچ سقفی نباید قائل باشیم.

این خبر خوبی است برای گونه‌ای که در طول دوران‌ها به شوق جابه‌جا کردنِ مرزها کوشیده است. بازی‌های المپیک محل ستایش از جابه‌جاییِ مرزهای قدرت و سرعت و چابکی و استقامت است. در علم جابه‌جا کردنِ مرزهای معرفت و فهم ستوده می‌شود. و هنر و ادبیات ستایشگرِ جابه‌جاییِ مرزهای آفرینش تجربیات زیبا و غنابخش است. بسیاری انسان‌ها و سازمان‌ها و کشورها گسترش منابع و قلمرو و طول عمر را ارزشمند می‌شمارند. با در نظر داشتنِ دغدغه و سواس‌گون ما انسان‌ها دربارهٔ مرزها، عجیب نیست که از میان کتاب‌های مشمول حقوق مالکیت فکری، پرفروش‌ترین کتاب کتاب رکوردهای گینس است.

پس اگر مرزهایی که ما پیشتر برای حیات قائل بودیم قابل شکستن است، حدودِ نهایی کدام‌اند؟ چه اندازه از کیهان ما می‌تواند دارای زندگی باشد؟ حیات تا کجا می‌تواند برسد و چه اندازه می‌تواند دوام آورد؟ زندگی از چقدر ماده می‌تواند بهره‌بردار و چه اندازه انرژی و اطلاعات و محاسبه می‌تواند استخراج کند؟ این حدود نهایی را فهم ما وضع نمی‌کند، بلکه کار قوانین فیزیک است. طرفه اینجاست که در این شرایط تحلیل آینده بلندمدت حیات ساده‌تر از آینده کوتاه‌مدت آن است.

اگر کل تاریخ ۱۳/۸ میلیاردسالهٔ کیهان را در یک هفته فشرده کنیم، آنگاه داستان ۱۰,۰۰۰ سالهٔ دو فصل پیش در کمتر از نیم ثانیه به پایان می‌رسد. یعنی با اینکه نمی‌توانیم پیش‌بینی کنیم که آیا انفجار هوش رخ خواهد داد و اگر رخ دهد چه زمانی خواهد بود و پیامدهای بلافاصله‌اش چه خواهد بود، اما همهٔ این افت و خیرها فقط لحظه‌ای گذرا در تاریخ کهکشان است که جزئیات آن بر حدود نهایی حیات تأثیری ندارد. اگر در دورهٔ پس انفجار هوش نیز موجودات زنده مانند انسان‌های امروزی دغدغهٔ گذشتن از مرزها را داشته باشند، آنگاه فناوری‌ای به‌وجود خواهند آورد که در عمل، به این حدود نهایی می‌رسد - چون می‌تواند. در این فصل، بررسی خواهیم کرد که این حدود کدام است، و به این ترتیب شمه‌ای خواهیم دید از اینکه

آینده بلندمدتِ حیات چگونه می‌تواند باشد. این حدود مبتنی بر درک کنونی ما از فیزیک است، بنابراین باید آن را به چشم کران پایینِ حالات ممکن دید: شاید کشفیات علمی آینده دریچه‌ای به سوی امکانات تازه برای رشد حیات بگشاید.

اما آیا به‌راستی می‌دانیم که حیات آینده این اندازه بلندپرواز خواهد بود؟ نه، نمی‌دانیم؛ شاید موجودات زندهٔ آینده به اندازهٔ معتادان هروئین قانع باشند یا مثل معتادان تلویزیون باشند و همیشه یک جا بنشینند و دائم پایه‌پای خانوادهٔ کارداشیان<sup>۱</sup> تماشا کنند. اما دلیلی هست که احتمال دهیم بلندپروازی و جاه‌طلبی یکی از ویژگی‌های صورت‌های پیشرفتهٔ حیات است. چنین موجودی، صرف نظر از آنکه به دنبال پیشینه کردن چه چیزی است، هوش یا طول عمر یا دانش یا تجربیات هیجان‌انگیز، به هر حال به منابع نیاز دارد. بنابراین انگیزه دارد که فناوری خود را تا حد امکان پیشرفت دهد تا بتواند از بیشتر منابعی که دارد استفاده کند. پس از آن، تنها راه برای پیشرفت بیشتر در اختیار گرفتن منابع بیشتر از طریق گسترش یافتن به مناطق هرچه بزرگ‌تری از کیهان خواهد بود.

همچنین این امکان هست که حیات در چندین نقطهٔ مختلف از کیهان ما به طور مستقل پدید آید. در این صورت، تمدن‌های غیربلندپرواز اساساً نقش مهمی در کیهان نخواهند داشت، و با گذشت زمان بخش‌های بزرگ و بزرگ‌تری از منابع کیهانی را بلندپروازترین صورت‌های حیات در اختیار می‌گیرند. بنابراین انتخاب طبیعی در سطح کیهانی عمل می‌کند و پس از مدتی، تقریباً تمام شکل‌های موجود حیات بلندپرواز و جاه‌طلب خواهند بود. به بیان مختصر، اگر علاقه داشته باشیم بدانیم کیهان ما در نهایت تا چه اندازه می‌تواند زنده شود، باید حد و مرز این جاه‌طلبی را که قوانین فیزیک تعیین می‌کنند مطالعه کنیم. بیایید همین کار را بکنیم! ابتدا بپردازیم به حد و

1. *Keeping Up with the Kardashians*