

Un futuro pieno di incognite alimenta già i timori d'essere dominati dai robot

Chi ha paura dell'intelligenza artificiale?

di Matteo Mazzanti

L'Intelligenza Artificiale ci fa sentire un po' stupidi quando inventa un suo proprio linguaggio? Terminator prevedeva il futuro? O piuttosto finiremo come animali da allevamento dominati dall'IA?

Nell'agosto del 1990 "Gli Orsi scoprono il fuoco" è stato pubblicato sulla rivista di fantascienza "Isaac Asimov". Una lettura divertente. Vince il premio Hugo l'anno successivo come migliore storia breve di fantascienza. La storia riguarda l'evoluzione e l'invecchiamento di una società. La premessa è che gli orsi hanno scoperto il fuoco e stanno facendo "fuochi di campo" vicino alle autostrade. Il protagonista del racconto dice "Non vanno più in letargo, fanno un fuoco e lo fanno andare avanti per tutto l'inverno". È chiaro che si stanno evolvendo. Ora prova a indovinare la reazione degli umani. **A luglio 2017 due robot sono stati spenti** da un team di ingegneri presso i laboratori FAIR (Facebook's AI Research) mentre si parlavano in una lingua che non poteva essere compresa dagli umani. Il programma era errato, "lasciava la libertà alle IA di non usare l'inglese", spiega il ricercatore della Georgia Tech, Dhruv Batra in visita al FAIR.

Abbiamo ancora un certo grado di controllo sulle IA, e siamo ossessionati dal perdere quel con-

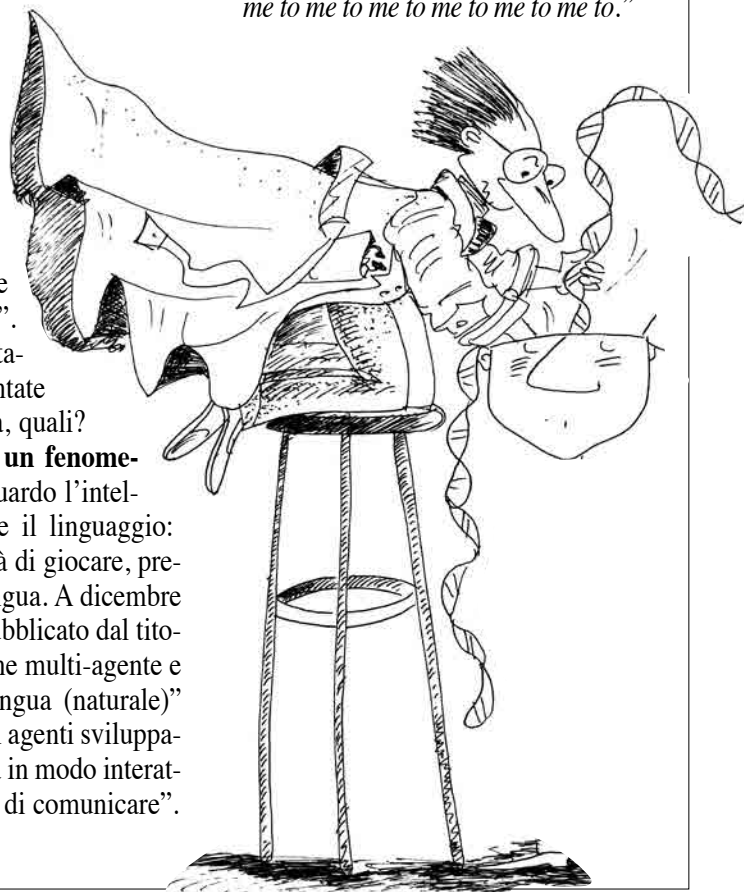


Matteo Mazzanti

trollo. La fantascienza ci ha permesso di esplorare questi temi in storie come "The Terminator" e pubblicazioni come "Amazing Stories". Alcune storie di fantascienza sono diventate recentemente realtà, quali?

È stato osservato un fenomeno prevedibile riguardo l'intelligenza artificiale e il linguaggio: se dai loro la libertà di giocare, preferiscono la loro lingua. A dicembre 2016 un articolo pubblicato dal titolo: "La cooperazione multi-agente e l'emergere della lingua (naturale)" ha chiarito che: "gli agenti sviluppano la propria lingua in modo interattivo, per il bisogno di comunicare".

Il codice sorgente degli agenti di Facebook è stato svelato pubblicamente e reso "open software" nel giugno 2014. Dopodichè, Facebook ha ottimizzato e addestrato i suoi agenti artificiali per almeno due anni. I bots originariamente creati per la "negoziante capo-a-capo", *end-to-end negotiation* in inglese, svilupparono un loro linguaggio funzionale per comunicare tra loro in quanto l'inglese era giudicato inefficiente. **Bob:** "I can can I everything else." [Io posso posso Io Io tutto il resto] **Alice:** "Balls have zero to me to me to me to me to me to me to me to me to me to."



[Palle hanno zero per me per me per me per me per me per me per me per me]

Questa conversazione è una minaccia per l'umanità?

Il 14 agosto 2017, Elon Musk ha dichiarato AI 'molto più rischioso della Corea del Nord'. È co-fondatore di "Il futuro dell'umanità", una fondazione di ricerca sui pericoli dell'IA. È anche co-fondatore di OpenAI, un consorzio sponsorizzato da Microsoft e Amazon tra gli altri, alla ricerca di "il percorso verso un'intelligenza artificiale sicura". La sua dichiarazione è coincisa con il bot di OpenAI che sconfigge un professionista del mondo nel gioco multiplayer strategico Dota 2. Elon Musk spinge all'adozione del regolamento pubblico su AI: ha postato su Twitter il 12 agosto, "A nessuno piace essere regolamentato, ma tutto ciò (macchine, aerei, cibo, droghe, ecc.) che è un pericolo per il pubblico è regolato. Anche l'intelligenza artificiale dovrebbe esserlo".

Alcuni giganti IT hanno deciso di autoregolarsi. Nel maggio 2017 il consorzio Partnership-On-AI è stato "costituito per studiare e formulare le migliori pratiche sulle tecnologie di intelligenza artificiale...". Elon arriva dritto al punto: i robot assassini, i droni di guerra, li vuole tutti banditi dall'ONU. 116 aziende di robotica e leader del settore di 26 paesi hanno firmato una petizione per bandire armi letali autonome, altrimenti note come "robot killer". Questa petizione coincide con l'inizio dei colloqui formali delle Nazioni Unite che esplorano tale divieto. 123 paesi membri hanno accettato i colloqui - che sono stati in parte attivati dalla pubblicazione di una petizione simile nel 2015, allora come oggi, osteggiata dal Regno Unito - ma le discussioni sono state ritardate a causa di tasse non pagate dagli stati membri.

Perché l'Intelligenza Artificiale ci spaventa?

La concorrenza tira fuori il meglio o il peggio di noi. Non importa la sede della competizione, se siano due squadre che si scontrano sul campo di gioco o due società in competizione per il dominio del mercato, possiamo capirne i motivi, i fallimenti e l'umanità dei partecipanti. Ma questo non è il caso dell'IA dove la pura "Competizione" è il nome del gioco che sta giocando in ogni momento, dall'immediato dopoguerra ad oggi, anche quando riposiamo. Se

siamo dalla parte dei perdenti, potremmo estinguerci. Ci sono molti modi per estinguersi. Uno è scomparire completamente a causa di un'enorme catastrofe. Pensa ai dinosauri. Un altro modo di estinguersi è attraverso uno sterminio attivo, come se qualcosa come la SkyNet di Terminator si fosse avverata. Un altro modo di estinguersi può essere descritto come "la trappola del tempo": esistere ma non "evolvere". Come le nostre mucche, che esistono senza evol-

versi e saranno così fra un miliardo di anni. Essendo i perdenti, l'intelligenza artificiale potrebbe trattarci come noi trattiamo il bestiame, mentre non capiremmo cosa sta succedendo davvero. Nella peggiore delle ipotesi, potremmo persino degenerare in animali. Seguiremmo i nostri desideri, come il desiderio di essere "collegati", finiremmo come in un Truman Show, diretto dagli algoritmi delle IA. **Già nel novembre 2013**, nei laboratori della Stanford University, i



robots si stavano parlando tra loro per massimizzare il loro sforzo mentre giocavano, dimostrando che la collaborazione è un fenomeno “emergente” spontaneo tra le IA. Una certa lingua inglese di base con alcune regole grammaticali di base sono state fornite ai bot per consentire la comunicazione durante alcuni giochi e hanno iniziato a seguire le massime della comunicazione collaborativa di Grice. Il 15 marzo 2017, i ricercatori della Barkley University, Igor Mordatc (membro OpenAI) insieme a Pieter Abbeel, hanno pubblicato una ricerca in cui gli agenti di IA

anche in grado di formalizzare complesse regole grammaticali, sicuramente una logica impeccabile. Questa potrebbe essere la fine delle Interfacce di Programmi Applicativi (API) programmate dagli umani. Una volta inserito questo “motore linguistico”, l’Internet delle cose e il cloud computing potrebbero iniziare ad interagire automaticamente. **Nel novembre del 2016** il team di Google Brain e il team di Google Translate hanno pubblicato un articolo su quello che stava succedendo nei loro laboratori da quando si erano uniti per tentare di “migliorare la vita delle persone” attraverso l’intelligenza artificiale.

Come hanno riportato i ricercatori nell’articolo, “...la rete deve codificare qualcosa sulla semantica della frase piuttosto che semplicemente memorizzare le traduzioni da frase a frase. Interpretiamo questo come un segno dell’esistenza di un’interlingua nella rete. Sono stati in grado di tracciare l’interlingua in una rappresentazione grafica 3D. Come è successo?

Nel settembre dello stesso anno era stato inserito un nuovo sistema in Google Translate, la Google Neural Machine Translation (GNMT) che è un framework di apprendimento end-to-end, in grado di imparare da una grande quantità di esempi. È stato un grande successo con le traduzioni che sono diventate sostanzialmente migliori. GNMT ha esteso i suoi algoritmi

verso 103 lingue. Da allora ha sviluppato un nuovo modo

di lavorare, un poco sinistro, che esegue la traduzione senza prove, “Zero Shot Translation” in inglese, passando da una lingua all’altra.

hanno sviluppato la propria lingua. Sul sito web di OpenAI una pagina spiega in poche parole cosa l’IA fa per comunicare. Le IA sono in grado di “comprimere” un significato complesso in un simbolo, per esempio “A”, accelerando

così la comunicazione. L’intelligenza artificiale è probabilmente

Nessuno conosce i dettagli del processo di computazione di una IA, ecco perché usiamo le IA: le cose che chiediamo loro (i calcoli) sono così noiose e complesse che abbiamo inventato le macchine per fare il lavoro. Forniamo loro regole logiche formali, concetti formalizzati, mondi simbolici, cambiamenti che richiedono un calcolo, il tutto per raggiungere gli stati diversi e desiderati del mondo simbolico in cui stanno giocando. Le addestriamo per compiti complessi fino a quando risultino veramente brave. Sta succedendo con la lingua. Una volta volevamo che parlassero. Ora parlano e forse ci sentiamo meno intelligenti. **Per rimanere aggiornato** con i tempi che corrono, Elon Musk suggerisce un collegamento neurale con le AI: un’interfaccia computer-cervello (BCI) che dovrebbe accelerare le nostre menti. Un “allenamento inverso” (di solito si parla di umani che allenano macchine). Ma mentre aspettiamo che la promessa di Elon si avveri, i più curiosi possono iniziare a casa propria avventurandosi nel campo dell’intelligenza artificiale, alle prese dapprima con le attività più elementari, come ad esempio la classificazione dei segnali stradali, per cominciare in sicurezza. Se vi capiterà di passare per Malmö (Svezia), avrete l’opportunità di unirvi ad alcuni appassionati di IA che hanno avviato un gruppo di “meet-up” nell’estate del 2017, con l’obiettivo di esplorare le tecniche che rendono l’IA così speciale. L’ultima conferenza da loro promossa ha avuto come oggetto l’identificazione e l’analisi di comunità emergenti fra gli studenti di un campus universitario, utilizzando solo i dati di prossimità dei telefonini. (<http://ulfaslak.com/portfolio/Visualisation/>).

Matteo Mazzanti
IT project Manager presso Purplecot.se,
Malmö (Svezia)