
Se due robot si parlano

Autore: Giulio Meazzini

Fonte: Città Nuova

Un piccolo imprevisto incidente nello sviluppo di due sistemi di intelligenza artificiale ha scatenato un grande dibattito in Rete sulle regole da definire per il rapporto tra macchine e uomo

All'inizio dello scorso agosto, una piccola notizia ha agitato la Rete: Facebook ha spento due dei suoi robot sperimentali, Bob e Alice, appena hanno cominciato a parlare tra loro **in una lingua sconosciuta**, che i programmatori non capivano. Tutto è nato da un banale **errore di programmazione** delle due "intelligenze artificiali", costruite con l'obiettivo di interagire con l'uomo, in linguaggio naturale, per l'acquisto di beni e servizi. I programmatori si sono dimenticati di inserire un controllo che "obbligasse" Bob e Alice a comunicare usando solo il linguaggio degli umani. Siccome questi sistemi di intelligenza artificiale sono costruiti in modo da **imparare dall'esperienza** e quindi ottimizzare le proprie modalità operative, finché hanno interagito con esseri umani hanno usato il nostro linguaggio, ma nel momento in cui hanno comunicato direttamente, macchina a macchina, hanno ottimizzato (in modo automatico) la trasmissione, usando un linguaggio più semplice ed efficace. Peccato che fosse incomprensibile per gli umani. I ricercatori di Facebook li hanno immediatamente spenti, corretto l'errore di programmazione e quindi concluso con successo l'esperimento. Ma ormai la frittata era fatta. In Rete si è scatenata la discussione: da una parte i **catastrofisti**, per i quali presto le intelligenze artificiali complotteranno contro l'uomo in un linguaggio per noi incomprensibile, dall'altra i **tecno-entusiasti** per i quali qualsiasi novità tecnologica va bene per definizione e porterà un futuro meraviglioso. Nella discussione sono intervenuti anche i pezzi grossi: **Mark Zuckerberg**, padrone di Facebook, ha accusato i catastrofisti di non capire niente di tecnologia e di diffondere previsioni esagerate e allarmistiche. Dall'altra parte **Elon Musk**, creatore di Tesla la macchina elettrica, e **Stephen Hawking**, famoso scienziato, continuano ad avvertire che lo sviluppo di intelligenze artificiali autonome potrebbe portare un giorno a guerre accidentali tra macchine e contro l'uomo, con conseguente distruzione della razza umana. Sarà meglio seguire bene lo sviluppo di questa discussione, se non altro per un preciso motivo: intelligenza artificiale significa macchine che si sviluppano (ed ottimizzano il loro comportamento) in modo **imprevedibile, e a volte incomprensibile, per gli esseri umani**. Queste macchine intelligenti si stanno diffondendo ovunque nel pianeta, compresi i gangli decisionali, come nel controllo degli scambi di borsa, delle centrali nucleari, dei sistemi di armamento, delle reti di comunicazione. Non è quindi sbagliata la richiesta di Elon Musk di concordare a livello mondiale alcune **precise regole di sviluppo** di queste piattaforme.