Pechino, Cina. Uso di un software di riconoscimento facciale nella sede dell'azienda tecnologica Megvii

GILLES SABRIÉ (THE NEW YORK TIMES/CONTRASTO)

Il codice Frankenstein

Andrew Smith, The Guardian, Regno Unito

I software che hanno la capacità di apprendere e svilupparsi da soli hanno dato vita a un sistema che rischia di sfuggire al controllo degli esseri umani

Il 18 marzo 2018 è stato il giorno che gli esperti di tecnologia temevano sarebbe arrivato. Quella sera la luna nuova non aggiungeva luce a una strada a quattro corsie già scarsamente illuminata di Tempe, in Arizona. Un modello speciale di Volvo Xc90 sviluppato per Uber ha localizzato un oggetto davanti a sé. L'auto si guidava da sola e procedeva da 19 minuti senza alcun intervento da parte del suo conducente di sicurezza, una persona in carne e ossa. Una serie di radar e sensori ha permesso al software di bordo di calcolare che, procedendo alla velocità di settanta chilometri all'ora, se l'oggetto fosse rimasto fermo l'auto l'avrebbe raggiunto in sei secondi. Ma gli oggetti che s'incontrano sulle strade raramente restano immobili. Perciò un altro software ha cominciato a scorrere una banca dati di entità meccaniche e biologiche riconoscibili per identificare quell'oggetto e prevedere il suo comportamento. All'inizio non ha trovato niente, ma qualche secondo dopo ha deciso che si trattava di un'altra auto, che si sarebbe allontanata senza richiedere un intervento particolare. Solo all'ultimo secondo ha identificato l'oggetto: era una donna in bicicletta con le buste della spesa appese ai due lati del manubrio. Senza dubbio la donna si aspettava che la Volvo la evitasse, come avrebbe fatto qualsiasi altro veicolo. Non potendo prendere una decisione da solo, il computer ha passato i comandi al conducente, che però in quel momento era distratto. Una donna di 49 anni, Elaine Herzberg, è stata investita e uccisa, costringendo il settore dell'alta tecnologia a farsi due domande scomode: quella tragedia era inevitabile? Ci dovremmo abituare all'idea che questi incidenti sono possibili?

"In un certo senso abbiamo perso il controllo. Quando scriviamo il codice di un programma e il codice diventa un algoritmo e l'algoritmo comincia a creare nuovi algoritmi, la cosa sfugge sempre di più al controllo umano. Il software finale è un universo di codici che nessuno capisce fino in fondo". Se queste parole vi sembrano agghiaccianti, è perché lo sono, visto che le ha pronunciate Ellen Ullman che, oltre a essere una programmatrice nota fin dagli anni settanta, è anche una delle poche persone che scrivono cose illuminanti sul processo di codifica. "Qualcuno dice: 'E allora Facebook? Crea algoritmi, li usa e può cambiarli'. Ma non funziona così. Una volta partiti, gli algoritmi imparano, si modificano e si gestiscono da soli. Facebook interviene ogni tanto, ma non li controlla davvero. E certi programmi non solo si gestiscono da soli: attingono a librerie software, a sistemi operativi profondi".

Cos'è un algoritmo?

Di poche cose in questo momento si discute così spesso e con tanta passione come degli algoritmi. Ma cos'è un algoritmo? Dalla nascita di internet negli anni novanta il concetto è cambiato molto. Di base, un algoritmo è una cosa semplice: una regola che rende automatico il trattamento di un certo tipo di dati. Se succede A, allora fai B, altrimenti fai C. È la logica dell'informatica classica. Se un utente dichiara di avere 18 anni, consentigli di accedere al sito, altrimenti scrivi: "Mi dispiace, devi avere 18 anni per entrare". I programmi informatici sono essenzialmente composti da una serie di questi algoritmi. Se ci sembra che i computer facciano miracoli è perché sono veloci, non perché sono intelligenti.

Parallelamente negli ultimi anni è emerso un significato più ambiguo e misterioso della parola "algoritmo", che ormai indica qualsiasi grande sistema in grado di prendere decisioni complesse, qualsiasi mezzo per raccogliere una serie di dati in ingresso e valutarli velocemente in base a una serie di criteri (o regole). Una cosa che ha rivoluzionato alcuni settori della medicina, della scienza, dei trasporti e delle comunicazioni. Gli algoritmi hanno migliorato la nostra vita in molti modi.

Così nel 2016 ha cominciato a prendere forma una visione più sfumata degli algoritmi. Se tendiamo a parlarne in termini quasi biblici, come entità indipendenti dotate di una vita propria, è perché siamo stati spinti a vederli in questo modo. Grandi aziende come Facebook e Google hanno venduto e difeso i loro algoritmi promettendo l'oggettività, la capacità di soppesare un insieme di condizioni con distacco matematico e senza nessun coinvolgimento emotivo. Non c'è da meravigliarsi quindi se questo uso degli algoritmi per prendere decisioni si è esteso alla concessione di mutui, cauzioni, indennità, posti nelle università.

Solo che ormai non accettiamo più così docilmente le strategie pubblicitarie per venderci questo tipo di algoritmi. Nel suo libro uscito nel 2016, *Armi di distruzione matematica*, Cathy O'Neil, un'ex ragazza prodigio della matematica che ha lasciato Wall street per insegnare e oggi cura l'ottimo blog Mathbabe, ha dimostrato che gli algoritmi potrebbero ingigantire e rendere ancora più radicati i pregiudizi umani. D'altronde i software sono scritti in prevalenza da uomini ricchi bianchi e asiatici, e riflettono la loro mentalità. Perché un pregiudizio produca dei danni non è necessaria la malafede e, a differenza di quanto facciamo con le persone, non possiamo chiedere a un algoritmo di spiegarci perché ha preso una certa decisione. O'Neil vorrebbe una "revisione degli algoritmi" di qualsiasi sistema influisca direttamente sul pubblico. Contro questa richiesta sensata, però, l'industria tecnologica lotterà con le unghie e con i denti, perché vende algoritmi.

La buona notizia è che la battaglia è cominciata. Quella cattiva è che sembra già inadeguata. Ci siamo concentrati sulle lontane promesse e sui pericoli dell'intelligenza artificiale e quasi nessuno si è accorto che stiamo entrando in una nuova fase della rivoluzione degli algoritmi che rischia di essere altrettanto pericolosa e disorientante.

Gli algoritmi messi sotto accusa da O'Neil e da altri esperti sono poco trasparenti ma prevedibili: fanno quello per cui sono stati programmati. In teoria, un bravo programmatore può esaminarli e mettere in discussione i princìpi su cui si basano. Ma mentre da una parte ci sono gli algoritmi che potremmo definire "stupidi", nel senso che fanno il loro lavoro in base a parametri stabiliti dagli esseri umani, dall'altra c'è il sogno più o meno lontano di un'intelligenza artificiale simile a quella umana. Una macchina davvero intelligente dovrebbe essere in grado di mettere in discussione la correttezza dei propri calcoli in base a qualcosa di simile a quello che per noi è l'intuito. Per mettere quest'ipotesi in prospettiva, la divisione Deep mind di Google è stata giustamente elogiata per aver creato un programma capace di diventare un campione delle sale giochi partendo semplicemente dall'indicazione di mirare al punteggio più alto possibile. Questa tecnologia, chiamata "apprendimento per rinforzo", funziona, perché per imparare come si accumulano i punti un computer può fare milioni di partite in pochissimo tempo.

Qualcuno chiama questo tipo di capacità "intelligenza artificiale ristretta", ma il termine "intelligenza" è impiegato più o meno come Facebook usa la parola "amico", per fa credere che qualcosa sia meno minaccioso e apparentemente più comprensibile di quello che è in realtà. Questo perché la macchina agisce fuori da qualunque contesto e non sa fare altro. E, soprattutto, non può trasferire quello che ha imparato da un gioco all'altro (cioè "apprendere per trasferimento"), quindi la sua intelligenza generale è inferiore a quella di un bambino piccolo o perfino di una seppia. In questo senso potremmo definire "intelligente" anche una torre di trivellazione o un pidocchio delle piante. I computer sono già molto più bravi di noi a svolgere certi compiti specializzati, ma il giorno in cui le loro capacità generali saranno pari alle nostre è ancora lontano, se mai arriverà.

Tra gli algoritmi "stupidi" e la vera intelligenza artificiale, però, c'è una via di mezzo piena di problemi che abbiamo già imboccato senza rifletterci troppo e quasi senza discuterne, e soprattutto senza accordarci sugli obiettivi, l'etica, la sicurezza e il metodo da seguire. Se gli algoritmi che usiamo non sono ancora intelligenti, cioè capaci di stabilire da soli che "c'è qualcosa che non va in un certo calcolo o in una certa decisione e quindi lo rifaccio", stanno comunque cominciando a imparare dal loro ambiente. E una volta che un algoritmo comincia ad apprendere, non sappiamo più con certezza quali saranno le sue regole e i suoi parametri. Quindi non possiamo essere sicuri di come interagirà con altri algoritmi, con il mondo materiale o con noi. Mentre in linea di principio gli algoritmi "stupidi" sono prevedibili e interrogabili, questi altri non lo sono. Dopo un po' che imparano, non sappiamo più cosa sono: diventano imprevedibili. Ispirandoci al romanzo di Mary Shelley, saremmo tentati di chiamarli "algoritmi Frankenstein".

I computer sono già più bravi di noi a svolgere certi compiti, ma il giorno in cui le loro capacità generali saranno pari alle nostre è ancora lontano

Lente balene

Questi algoritmi non sono nuovi in sé. Me ne sono occupato per la prima volta cinque anni fa, mentre facevo una ricerca per un articolo del Guardian sulle operazioni di borsa eseguite da computer superveloci. Avevo scoperto una cosa straordinaria: le borse erano diventate un ecosistema digitale creato dall'essere umano e organizzato in pile di scatole nere acquattate come ninja in costosi centri di elaborazione dati. Mentre prima le operazioni di borsa avvenivano in uno spazio fisico, ora tutto era affidato a un server centrale, in cui agili algoritmi predatori si nutrivano dei pachidermici algoritmi di investitori istituzionali e li spingevano a vendere a un prezzo più basso e a comprare a uno più alto, ingannandoli sulle vere condizioni del mercato. Gli operatori umani (anche se nessun essere umano operava più direttamente) chiamavano "balene" queste entità più grandi e lente, che in buona parte erano costituite da fondi comuni e fondi pensione. Ormai le balene erano la principale fonte di profitto. In pratica, gli algoritmi cercavano di raggirarsi a vicenda, combattevano battaglie invisibili alla velocità della luce, piazzando e annullando lo stesso ordine diecimila volte al secondo o inserendone nel sistema talmente tanti da scuotere l'intero mercato, tutto al di fuori del controllo umano.

Non c'è da sorprendersi che questa situazione fosse instabile. Nel 2010 c'era stato un *flash crash*, un crollo improvviso, durante il quale il mercato era andato in caduta libera per cinque traumatici minuti e poi si era ripreso senza nessun motivo apparente. All'epoca andai a Chicago a parlare con Eric Hunsader, un abile programmatore in grado di analizzare i dati del mercato in modo molto più dettagliato delle autorità di vigilanza. Hunsader mi dimostrò che entro il 2014 ci sarebbero stati "mini crolli improvvisi" ogni settimana, anche se non riusciva a spiegarmi esattamente perché. Lui e i suoi collaboratori avevano cominciato a dare un nome ad alcuni degli "algoritmi" che vedevano. Come i cacciatori di cerchi nel grano britannici battezzavano le forme che trovavano nei campi, li avevano chiamati Wild thing, Zuma, Click e Disruptor.

Neil Johnson, un fisico specializzato in sistemi complessi della George Washington university, alla fine del 2013 ha realizzato uno studio sulla volatilità dei mercati. "È affascinante", mi ha detto. "Da anni si parla dell'ecologia dei sistemi informatici in modo vago, in termini di virus e così via. Invece è un sistema reale che possiamo studiare. Il problema più grande è che non sappiamo come funziona o cosa potrebbe provocare. E l'atteggiamento di tutti sembra essere del tipo: 'Occhio non vede, cuore non duole'".

È significativo che l'articolo di Johnson sia stato pubblicato sulla rivista scientifica Nature e che descriva quei crolli in borsa come "un improvviso passaggio dell'intero sistema da una fase uomo-macchina a una solo macchina, caratterizzata da eventi altamente improbabili di brevissima durata, i cosiddetti 'cigni neri'". Secondo lo storico della scienza George Dyson, questo scenario è stato complicato ulteriormente dal fatto che alcuni operatori di borsa permettono agli algoritmi d'imparare, "lasciando sempli-

Milton Keynes, Regno Unito. Un robot per le consegne a domicilio

LEON NEAL (GETTY IMAGES)

cemente che le scatole nere provino a fare operazioni diverse con piccole somme di denaro e, se funziona, rinforzano quelle regole. Sappiamo che l'hanno fatto. In pratica ci sono casi in cui nessuno sa quali siano le regole: gli algoritmi le creano da soli, e li si lascia evolvere come gli organismi in natura". Gli osservatori esterni al settore della finanza hanno cominciato a ipotizzare che ci sarebbe stato un catastrofico *flash crash* globale (definito *splash crash*) e che il settore in più rapida crescita nella finanza sarebbe stato costituito da strumenti che avrebbero tratto vantaggio dalla volatilità. Nel suo romanzo uscito nel 2011, *L'indice della paura*, Robert Harris immagina la nascita di un'intelligenza generale artificiale proprio da questo flusso digitale. Con mia grande sorpresa, nessuno degli scienziati con cui ho parlato ha escluso questa possibilità.

Potremmo liquidare il fenomeno come uno dei misteri dell'alta finanza, se non fosse che questo tipo di tecnologia è stato adottato per la prima volta dall'industria del porno e poi da tutti gli altri. E visto che la finanza è la pornografia del ventunesimo secolo, quando mi è sembrato di vedere che gli algoritmi usati in borsa stavano causando problemi altrove, ho chiamato di nuovo Neil Johnson. "Ha proprio ragione", mi ha detto: nel mondo circolano nuovi algoritmi che hanno "la capacità di riscrivere parti del loro codice" e quindi diventano una sorta di "algoritmi genetici".

Johnson pensa di averne trovato la prova facendo qualche incursione per verificare i fatti su Facebook. Se è come dice lui, significa che gli algoritmi si stanno adattando. "Dopotutto, Facebook non è altro che un grande algoritmo", dice Johnson. "E penso che questo sia proprio il suo problema. Può riconoscere la mia faccia in una foto sulla pagina di qualcun altro, prendere i dati del mio profilo e metterci in collegamento. Questo richiede un algoritmo molto semplice e concreto. Ma la questione è: quale sarà l'effetto con miliardi di questi algoritmi che agiscono tutti insieme? Da regole microscopiche non è possibile prevedere i comportamenti che impareranno ad avere a livello macro. Facebook sostiene di sapere esattamente quello che succede in situazioni specifiche, e probabilmente è vero. Ma cosa succede nell'insieme? Questo è il problema".

Sul tema Johnson e un gruppo di studiosi dell'università di Miami e di quella di Notre Dame hanno prodotto un saggio con cui si propongono di dimostrare che i tentativi di collegare tra loro le persone sui social network inevitabilmente polarizzano la società nel suo complesso. Secondo lui, Facebook e altre aziende dovrebbero costruire modelli dei possibili effetti dei loro algoritmi come gli scienziati creano modelli del cambiamento climatico o degli schemi meteo.

O'Neil dice di aver escluso volutamente questi algoritmi capaci di adattarsi dal suo *Armi di distruzione matematica*. In un ambiente algoritmico intricato in cui non c'è niente di chiaro, è molto difficile attribuire responsabilità a particolari segmenti di codice. È più facile ignorarli o negarne l'importanza, perché sia quei segmenti sia i loro effetti sono difficili da individuare, spiega. Quindi mi suggerisce che, se voglio vederli all'opera, dovrei chiedere come sarebbe un *flash crash* di Amazon.

"Anch'io cercavo questi algoritmi", dice, "ma di recente un amico che vende libri su Amazon mi ha detto che per quelli come lui la questione dei prezzi è ormai una follia. Ogni tanto trovi qualcuno che twitta 'Potete comprare filati di lusso su Amazon a quarantamila dollari'. E ogni

volta che sento cose di questo tipo, penso: 'Ok, questo deve essere l'equivalente di un *flash crash*!'".

Esistono molte notizie di eventi anomali su Amazon, spesso sotto forma di commenti divertiti di venditori, e uno studio del 2016, in cui si legge: "Stanno emergendo casi in cui algoritmi creati per fissare i prezzi in concorrenza tra loro hanno interagito in modo inaspettato e prodotto prezzi imprevedibili, e casi in cui gli algoritmi stessi sono stati intenzionalmente progettati per imporre certi prezzi". Anche qui il problema è come individuare le responsabilità in un ambiente algoritmico caotico in cui è quasi impossibile applicare o individuare un semplice rapporto di causa-effetto.

Pericoli reali

Quand'è in gioco la sicurezza, queste cose diventano importanti. Quando un automobilista alla guida di una Toyota Camry è morto uscendo di strada dopo aver fortemente accelerato senza un motivo apparente, gli esperti della Nasa hanno passato sei mesi a esaminare i milioni di righe di codice del sistema operativo dell'auto, senza trovare alcuna prova a sostegno delle accuse mosse dai familiari del guidatore, cioè che la macchina avesse accelerato di propria iniziativa. Solo dopo aver passato venti mesi a scavare nel codice, due esperti di software hanno dimostrato che la famiglia aveva ragione, scoprendo un ammasso intricato, quello che i programmatori chiamano *spaghetti code*, pieno di algoritmi in conflitto tra loro che generavano risultati anomali e imprevedibili. I veicoli che si guidano da soli possono contenere cento milioni di righe di codice e, dato che nessun programmatore è in grado di prevedere tutto quello che può succedere su una strada, devono poter imparare ed essere continuamente aggiornati. In un ambiente così fluido, com'è possibile evitare incidenti, senza contare che gli algoritmi possono essere attaccati dagli hacker?

Vent'anni fa George Dyson aveva previsto quasi tutto quello che sta succedendo oggi nel suo libro *L'evoluzione delle macchine. Da Darwin all'intelligenza globale*, diventato un classico. Il problema, mi ha detto, è che stiamo costruendo sistemi che vanno oltre la nostra capacità di controllo. Pensiamo che se è deterministico (cioè agisce in base a regole fisse, come fa un algoritmo), un sistema sia prevedibile, e tutto ciò che è prevedibile sia anche controllabile. Ma entrambe queste affermazioni sono sbagliate. "In realtà questo sistema procede in modo autonomo e frammentario", dice Dyson. "Quello che mi ossessionava vent'anni fa e che ormai ha invaso il mondo sono gli organismi pluricellulari digitali, simili a quelli biologici: tutti quei frammenti di codice che sono nei telefoni delle persone, e che collettivamente agiscono come un unico organismo pluricellulare. C'è una vecchia regola, detta legge di Ashby, secondo cui un sistema di controllo dev'essere complesso quanto quello che deve controllare, e ora ci siamo dentro fino al collo, soprattutto a causa dello sviluppo dei veicoli autonomi, il cui software deve avere un modello completo di tutto, che quasi per definizione non saremo mai in grado di capire. Un auto potrebbe schiantarsi contro un camion dei vigili del fuoco semplicemente perché ci siamo dimenticati di inserire il camion dei vigili del fuoco tra i possibili ostacoli".

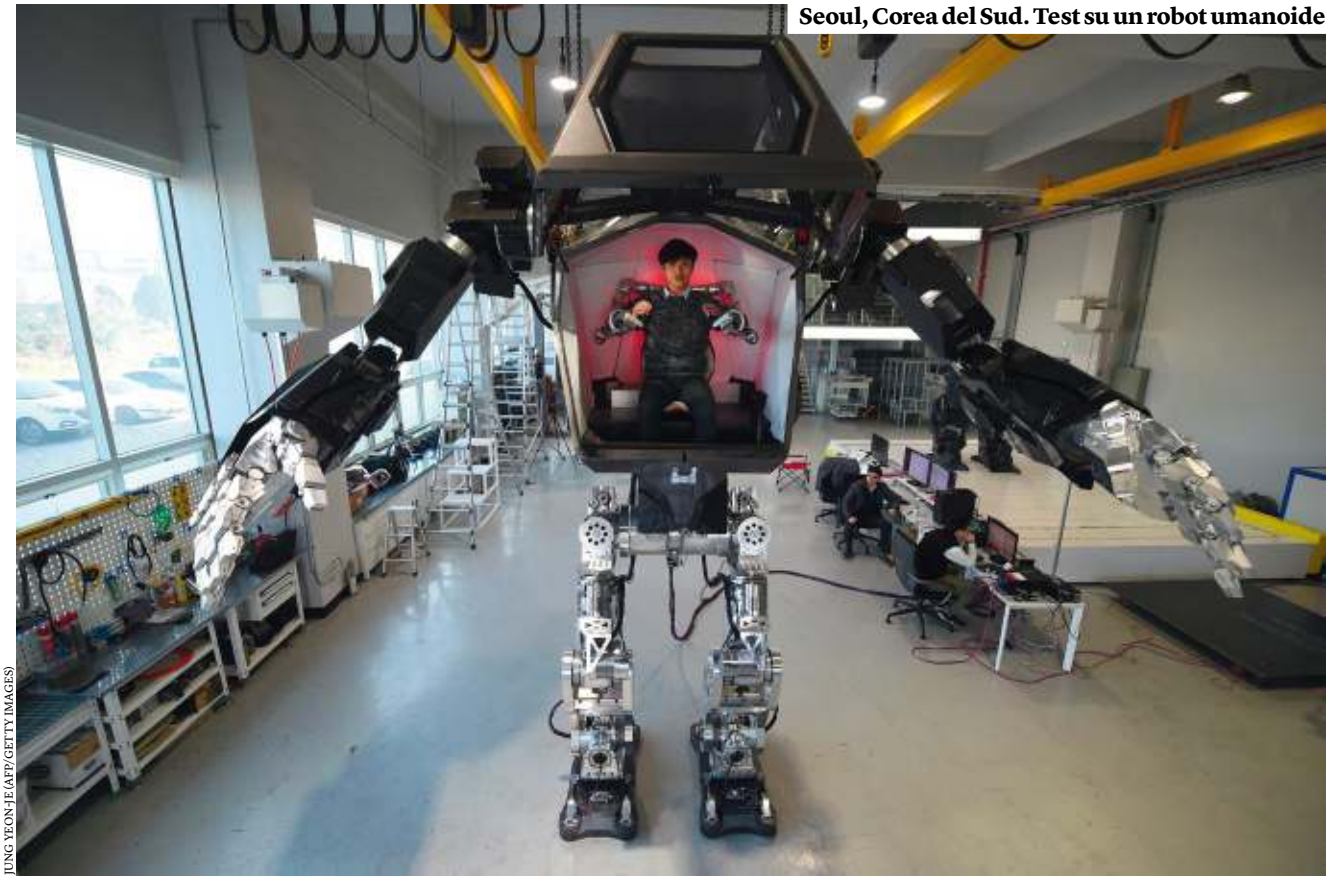
Le leggi – basate sul principio che per attribuire una responsabilità è necessario dimostrare un'intenzione o un atto di negligenza – vanno riviste

A differenza dei vecchi sistemi elettromeccanici, questi nuovi algoritmi sono anche impossibili da testare in modo esaustivo. Finché non avremo, se mai le avremo, macchine superintelligenti in grado di farlo per noi, continueremo a rischiare. Dyson dubita che avremo mai veicoli autonomi liberi di circolare nelle strade delle città, Toby Walsh, un docente di intelligenza artificiale all'università del New South Wales, in Australia, che ha scritto il suo primo programma a tredici anni e prima dei venti aveva già una sua azienda informatica, mi spiega perché dal punto di vista tecnico. "Nessuno sa scrivere un segmento di codice che insegna al computer a riconoscere un segnale di stop. Abbiamo passato anni a provarci, e non ci siamo riusciti. Evidentemente ci siamo bloccati perché non siamo capaci di scomporre il problema. Quando programmi, scopri che devi imparare a scomporre i problemi in parti sufficientemente semplici da corrispondere a un'unica istruzione per la macchina. E non lo sappiamo fare quando abbiamo davanti un problema complesso come identificare un segnale di stop o tradurre una frase dall'inglese al russo, va oltre le nostre capacità. Sappiamo solo scrivere un algoritmo più generale in grado di imparare a farlo in base a un certo numero di esempi". Questo è il motivo dell'attuale interesse per l'apprendimento automatico. Ora sappiamo che Herzberg, la donna investita in Arizona, è morta perché l'algoritmo non è riuscito a inserirla nella categoria giusta. È dipeso da un errore di programmazione, da un addestramento insufficiente o dall'arrogante rifiuto di ammettere i limiti della nostra tecnologia? Forse non lo sapremo mai. " E prima o poi smetteremo di scrivere algoritmi", prosegue Walsh, "perché le macchine saranno in grado di farlo molto meglio di noi. In questo senso, il mestiere di programmatore forse sta per sparire".

Secondo Walsh è per questo che oggi è ancora più importante imparare a programmare, perché più ci allontaniamo da questo lavoro e più ci sembrerà qualcosa di magico su cui non possiamo intervenire. Quando ha letto la definizione di algoritmo che ho dato in questo articolo, l'ha trovata incompleta e ha commentato: "Direi che ora il termine 'algoritmo' indica qualsiasi grande sistema complesso in grado di prendere decisioni e l'ambiente più ampio in cui è inserito, che lo rende ancora più imprevedibile". Una prospettiva agghiacciante. Secondo Walsh, la nuova frontiera della tecnologia sarà l'etica, con "una nuova età dell'oro per la filosofia".

È d'accordo con lui Eugene Spafford, esperto di sicurezza cibernetica della Purdue university. "Quando si devono fare delle scelte, entra in gioco l'etica. Vogliamo avere qualcuno a cui chiedere spiegazioni o attribuire colpe, e non possiamo farlo se abbiamo davanti un algoritmo. Una delle critiche principali rivolte a questi sistemi è che non è possibile tornare indietro e capire esattamente perché è stata presa una certa decisione. Il numero di scelte intermedie è così vasto che non possiamo sapere come siamo arrivati a quel punto e dimostrare senz'ombra di dubbio che c'è un responsabile".

L'argomentazione opposta è che se un programma ha sbagliato può essere riscritto o modificato per evitare che sbagli di nuovo. Nonostante questo, e anche se



Seoul, Corea del Sud. Test su un robot umanoide

JUNG YEON-JE (AFP/GETTY IMAGES)

alla lunga l'automazione diventerà sempre più sicura, le nostre leggi – basate sul principio che per attribuire una responsabilità è necessario dimostrare che c'è stata un'intenzione o un atto di negligenza – dovranno essere riviste. Un cane non può essere ritenuto legalmente responsabile di aver morso qualcuno, il suo padrone probabilmente sì, ma solo se il comportamento dell'animale è considerato prevedibile. Nel mondo degli algoritmi, forse, molti risultati inaspettati non erano prevedibili; ma questa potrebbe diventare un'arma a doppio taglio e rendere più facile nascondere le responsabilità.

Gli obiettivi militari

In futuro, il commercio, i social network, la finanza e i trasporti potrebbero diventare il male minore. Anche se non sono più il volano dell'innovazione come in passato, gli eserciti sono ancora i principali utilizzatori delle nuove tecnologie. Non c'è da meravigliarsi, quindi, se la notizia che le armi autonome stanno invadendo silenziosamente i campi di battaglia è stata accolta con preoccupazione dagli scienziati e dagli esperti di tecnologia. Attualmente, a guardia della zona smilitarizzata tra la Corea del Nord e la Corea del Sud c'è un cecchino robotizzato, e anche se l'azienda che lo ha fabbricato, la Samsung, nega che sia totalmente autonomo, nessuno le crede. La Russia, la Cina e gli Stati Uniti sostengono di essere a vari stadi dello sviluppo di sciami di droni armati e coordinati. Washington vuole costruire dei missili capaci di sorvolare un campo di battaglia per giorni prima di scegliere gli obiettivi da colpire. Un gruppo di dipendenti di Google si è dimesso e migliaia di altri hanno contestato la scelta del colosso tecnologico di fornire software per l'apprendimento automatico al Project Maven, il programma di "algoritmi di guerra" del Pentagono. In seguito alle proteste, l'azienda alla fine ha deciso di non rinnovare il contratto e di pubblicare un codice etico per l'uso dei suoi algoritmi.

Come altre aziende tecnologiche, Google aveva difeso l'etica del suo software, dichiarando che avrebbe contribuito a scegliere meglio i bersagli e quindi a risparmiare vite umane. Il problema è come fanno i tecnici informatici a sapere quello che faranno i loro algoritmi una volta sul campo di battaglia, soprattutto visto che le parti svilupperanno controsistemi algoritmici progettati per confondere le idee alle armi nemiche. Come nel caso del mercato azionario, probabilmente l'imprevedibilità sarà considerata un vantaggio più che uno svantaggio, perché offre alle armi maggiori possibilità di resistere ai tentativi di sabotaggio. Per questo e per altri motivi, corriamo il rischio di trasformare completamente le nostre macchine, intrappolando il mondo in una rete di *spaghetti code*.

Lucy Suchman, della Lancaster university, nel Regno Unito, è tra i ricercatori che hanno scritto una lettera aperta a Google, chiedendo all'azienda di riflettere sull'uso a fini militari del suo lavoro. I motivi delle aziende tecnologiche sono facili da comprendere, dice Suchman: i contratti con l'esercito fanno guadagnare molto. E per quanto riguarda il Pentagono, ha ormai una rete di sensori e sistemi di sorveglianza molto più estesa della sua capacità di usare i dati raccolti. "Sono sopraffatti dai dati, perché ora hanno nuovi mezzi per raccoglierli e conservarli, ma non sono in grado di elaborarli. Perciò i dati risultano praticamente inutili, a meno che succeda un miracolo. Secondo me, la continua acquisizione di aziende per la raccolta dei dati è dovuta a una specie di pensiero magico, si spera sempre di tro-

vare ‘una tecnologia miracolosa che darà un senso a tutto questo’”.

Suchman fornisce anche statistiche che gettano una luce agghiacciante sul progetto Maven. Secondo le analisi degli attacchi con i droni effettuati in Pakistan dal 2003 al 2013, meno del 2 per cento delle persone uccise erano obiettivi di alto livello e costituivano una seria minaccia per gli Stati Uniti. Si ritiene che circa il 20 per cento fossero non combattenti, mentre del restante 75 per cento non si sa praticamente niente. Anche se i bersagli validi fossero il doppio, il triplo o il quadruplo, costringerebbero comunque qualsiasi persona di buon senso a riflettere. “Abbiamo una tecnologia d’identificazione piuttosto rozza e il progetto Maven si propone di automatizzarla, rendendola ancora meno affidabile e più discutibile. È una pessima idea”, dice Suchman.

Per la maggior parte dei problemi di cui abbiamo parlato, le soluzioni già esistono o si possono trovare, ma non senza incentivare le aziende tecnologiche a mettere gli interessi della società sullo stesso piano dei loro. Una riflessione più a lungo termine è che forse gli attuali metodi di programmazione non sono più all’altezza della situazione, date le dimensioni, la complessità e l’interdipendenza degli algoritmi su cui facciamo sempre più affidamento. Una soluzione, adottata dalla Federal aviation authority statunitense per quanto riguarda l’aviazione civile, è registrare e valutare il contenuto di tutti i programmi e i successivi aggiornamenti in modo così dettagliato da poter anticipare le interazioni tra gli algoritmi, ma questo non è praticabile su vasta scala. Una parte dell’industria aerospaziale usa un metodo relativamente nuovo chiamato programmazione basata su modelli, in cui le macchine fanno la maggior parte del lavoro di codifica e sono in grado di testarlo in corso d’opera.

Le soluzioni si possono trovare, ma incentivando le aziende tecnologiche a mettere gli interessi della società sullo stesso piano dei loro

Lilly Irani, una collega di Suchman dell’università della California a San Diego, ricorda che nei sistemi algoritmici le informazioni viaggiano alla velocità della luce, fuori dal controllo di qualsiasi essere umano. E le discussioni tecniche sono spesso usate come cortina di fumo per evitare ogni responsabilità.

“Quando parliamo di algoritmi, spesso parliamo di burocrazia. Le scelte dei programmatori e dei politici sono presentate come neutre, mentre in passato qualcuno avrebbe dovuto assumersene la responsabilità. Le aziende tecnologiche affermano che stanno lavorando per rendere più preciso il progetto Maven, cioè per fare in modo che vengano uccise le persone giuste. Ma nessuno mette in discussione il presupposto che si possa uccidere chiunque viva dall’altra parte del mondo, e che sia l’esercito statunitense a stabilire cos’è sospetto. Quindi le discussioni tecnologiche sono usate per coprire questioni politiche. Anche la scelta di usare gli algoritmi per automatizzare certi tipi di decisioni è politica”.

Le convenzioni legali della guerra moderna danno per scontata la responsabilità umana per le decisioni che vengono prese. Come minimo, la guerra degli algoritmi confonde le acque in un modo di cui un giorno potremmo pentirci.

Ma questo tipo di programmazione potrebbe non essere la panacea che alcuni sperano. Non solo allontana ancora di più gli esseri umani dal processo ma, com’è emerso in uno studio realizzato dal fisico Johnson per il dipartimento della difesa statunitense, c’è anche il rischio di “comportamenti estremi che non potevano essere dedotti dal codice stesso”, perfino in grandi sistemi complessi costruiti usando questa tecnica. Attualmente si sta cercando soprattutto di trovare il modo per far risalire certi comportamenti inaspettati degli algoritmi alle righe di codice che li hanno prodotti. Nessuno sa se c’è una soluzione, ma è probabile che davanti ad algoritmi aggressivi progettati per entrare in conflitto tra loro o adattarsi nessuna soluzione funzionerà.

Qualche precauzione

Mentre aspettiamo di capirlo, dobbiamo prendere qualche precauzione. Paul Wilmott, un esperto britannico di analisi quantitativa fortemente critico nei confronti delle operazioni di borsa affidate ai computer superveloci, suggerisce ironicamente di “imparare a sparare, fare la marmellata e lavorare a maglia”. Spafford, l’esperto di sicurezza informatica, consiglia d’imporre alle aziende tecnologiche di assumersi la responsabilità del comportamento dei loro prodotti, anche se non è possibile individuare specifiche righe di codice “canaglia” o dimostrare che c’è stata negligenza. Spafford osserva che la venerabile Association for computing machinery ha già modificato il suo codice etico, aggiungendo alle norme da rispettare qualcosa di simile al giuramento di Ippocrate dei medici, per cui i professionisti del settore devono impegnarsi a non danneggiare nessuno e a prevedere le possibili conseguenze del loro lavoro.

Da parte sua, Johnson ritiene che il nostro disagio nei confronti degli algoritmi è almeno in parte teorico, è solo la febbre di crescita di un nuovo campo dell’esperienza umana. Ride ricordando la prima volta che abbiamo parlato di questo argomento qualche anno fa. All’epoca le mie domande riguardavano timori di nicchia, a cui si interessavano poche persone che studiavano in modo maniacale quello che succedeva in borsa. “E ora siamo arrivati al punto che gli algoritmi influiscono anche sulle elezioni. Che sta succedendo? Secondo me, dal punto di vista scientifico il problema è che i programmatori di software sono abituati a scrivere programmi che “ottimizzano” le cose. E giustamente, perché spesso così si può migliorare la distribuzione del peso su un aereo o calcolare la velocità che consente di consumare meno carburante. In circostanze normali, ottimizzare è sensato. Ma in circostanze insolite non lo è, e dobbiamo chiederci: ‘Qual è la cosa peggiore che potrebbe succedere quando questo algoritmo comincia a interagire con gli altri?’. Il problema è che non abbiamo neanche una parola per questo concetto, e meno che mai gli strumenti scientifici per studiarlo”.

Si ferma un attimo, come per riflettere meglio sul problema. “Il fatto è che ottimizzare significa massimizzare o minimizzare qualcosa, che poi in termini informatici è la stessa cosa. Quindi qual è l’opposto dell’ottimizzazione, cioè qual è la soluzione meno ottimale, e come la identifichiamo e misuriamo? La domanda che dobbiamo farci, e non facciamo mai, è: ‘Qual è il comportamento più estremo possibile di un sistema che, in teoria, sto ottimizzando?’”.

E alla fine di un altro breve silenzio, dice con un leggero tono di sorpresa: “Fondamentalmente, abbiamo bisogno di una nuova scienza”. ◆ *bt*