

Umanità artificiale: il rapporto uomo-macchina tra fiction e scienza

Isabella Tomassucci

Abstract

Can machines ever have feelings and conscience? And if it were even possible, why should it be desirable? Science and science fiction narrative have been asking these questions for over a century, generating rather similar creative products. But are the processes that drive the imagination of storytellers and scientists the same? Analyzing some examples taken from science fiction and scientific research in the field of Artificial Intelligence, the article shows affinities and differences between these two areas of human creativity.

Keywords

Artificial Intelligence; Science; Art; Science Fiction; Creativity

Umanità artificiale: il rapporto uomo-macchina tra *fiction* e scienza

Isabella Tomassucci

«Why do I love you? Are these feelings even real? Or are they just programming?»

Samantha (*Her*)

«Well, I don't think there is any question about it. It can only be attributable to human error.»

Hal9000 (*2001: Odissea nello spazio*)

Le macchine potranno mai provare sentimenti e avere una coscienza? E se mai fosse possibile, perché dovrebbe essere desiderabile? Scienza e narrativa fantascientifica si pongono queste domande da oltre un secolo, ma i processi che guidano l'immaginazione di narratori e scienziati sono gli stessi? Analizzando alcuni esempi tratti dalla narrativa di fantascienza e confrontandoli con alcuni prodotti tecnologici e ricerche scientifiche in materia di Intelligenza Artificiale, l'articolo mette in mostra affinità e differenze tra i prodotti di questi due ambiti della creatività umana.

Arte e scienza: due domini della creatività

L'opposizione tra cultura scientifica e cultura artistica, tra scienze dure e scienze umane è un'invenzione piuttosto recente. Le più grandi menti dell'antichità erano allo stesso tempo artisti e scienziati nel senso in cui li intendiamo noi oggi. Pensiamo a Platone, Aristotele, Leonardo da Vinci, Diderot, Voltaire, Bacon: hanno spaziato dalle scienze naturali alla filosofia, dalle arti figurative all'ingegneria, dall'anatomia all'architettura. La loro creatività derivava dall'osservazione del mondo naturale e il prodotto della loro inventiva produceva artefatti riconducibili ad ambiti che oggi percepiamo come totalmente distinti. La sempre più netta divisione dei mestieri e delle attività sociali ha portato alla progressiva frantumazione della cultura e al nascere di numerosi specialismi che si dividono in due macro-domini: quello dell'arte e quello della scienza, tanto che alcuni si sono chiesti se sia corretto parlare di due culture distinte (Andreasen, Ramchandran 2012). Alcuni studiosi della creatività hanno analizzato i due processi creativi sottolineando analogie e diversità, per arrivare alla conclusione che non ci sono differenze sostanziali tra la creatività scientifica e quella artistica (*ibid*; Tang, Leonard 1985). L'unica differenza fondamentale nei processi di generazione delle idee sembra essere quella tra pensiero verticale e pensiero laterale (De Bono 1996): il primo, monotematico e unidirezionale, è alla base dell'approfondimento e della creazione degli specialismi; il secondo, trasversale e multidirezionale, è il motore della creatività e della generazione di nuove idee.

Nonostante la tendenza a settorializzare, arte e scienza hanno continuato a dialogare, in maniera più o meno consapevole. Lo scambio tra invenzione artistica e progresso scientifico è particolarmente evidente nel rapporto tra scienza e fantascienza: entrambe tendono a immaginare scenari che sembrano impossibili allo stato attuale. La scienza lo fa con l'intenzione di trovare il modo concreto per attualizzarli, l'arte invece, libera dalle costrizioni della fattibilità tecnica, si spinge oltre, immaginando spesso scenari estremi.

In questo articolo, prendiamo in considerazione la questione della riproduzione dell'intelligenza umana nelle macchine, un'idea che ha

prodotto numerose invenzioni sia in campo narrativo che in campo scientifico. Sull'argomento esiste una vasta letteratura, ma nei paragrafi successivi ci si concentrerà sui prodotti della narrativa fantascientifica (letterari, televisivi e cinematografici), per quanto riguarda l'immaginazione artistica, e su alcuni chatbot dotati di Ai, al fine di individuare analogie e differenze nei processi creativi che li hanno generati.

Dall'uomo alla macchina e ritorno: strutture di senso

La narrativa di fantascienza riflette e, in alcuni casi, anticipa le scoperte scientifiche e le invenzioni tecnologiche. A volte lo fa in chiave utopica (pensiamo ad esempio alle macchine meravigliose di Jules Verne che aprivano all'uomo nuove possibilità di conoscenza del mondo), più spesso in chiave distopica o, comunque, critica. Le ragioni del passaggio dall'utopia alla distopia sono molteplici (Pillera 2010), ma non sono direttamente oggetto di questo articolo, che invece si concentra sulle modalità di narrazione/rappresentazione dell'intelligenza artificiale, analizzandole con gli strumenti della semiotica. Si parte, infatti, dal presupposto che la fantascienza e la ricerca scientifica sono entrambe analizzabili in termini di narratività, intendendo con narratività «una ipotesi interpretativa per descrivere la struttura profonda di ogni manifestazione culturale» (Marrone 2011: 32). Ci si avvale quindi dello strumento del quadrato semiotico per rappresentare le strutture profonde alla base di tutte le rappresentazioni del rapporto uomo-macchina.

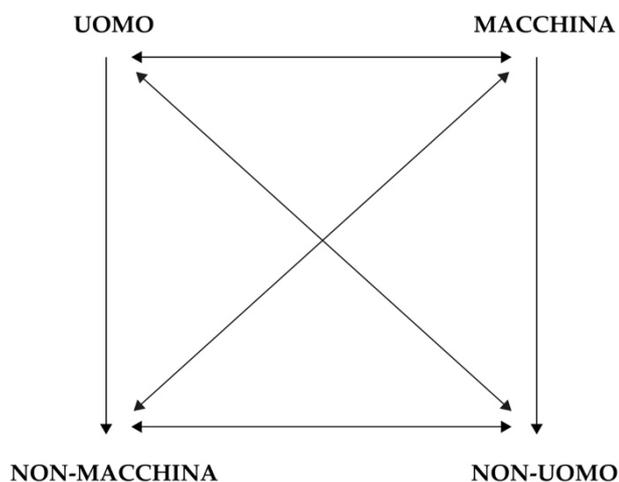


Figura 1. Quadrato semiotico delle relazioni di senso fondamentali del rapporto uomo - macchina

In Figura 1 vediamo che dalla relazione di contrarietà uomo / macchina si originano:

- le relazioni di complementarità macchina / non-uomo e uomo / non-macchina
- la relazione di subcontrarietà non macchina / non-uomo.

Da questi semi di prima generazione possiamo derivare il termine complesso e il termine neutro mostrati in Figura 2. Dalla somma dei termini contrari origina il cyborg, l'unione compiuta tra uomo e macchina, dalla somma dei subcontrari si genera il mostro, la creatura incomprensibile che non è né completamente macchina, né del tutto uomo. Se i mostri creati artificialmente restano per ora dominio esclusivo della fantascienza, i cyborg si realizzano in talmente tante sfumature che è difficile farne un inquadramento preciso: una protesi a un arto, un pace-maker o un dispositivo come quello usato da Stephen Hawking per parlare generano già dei cyborg. Lasciamo quindi da parte il termine neutro e quello complesso per concentrarci sui processi

creativi che si originano a partire dalle relazioni tra questi termini e dalle trasformazioni di un termine in un altro.

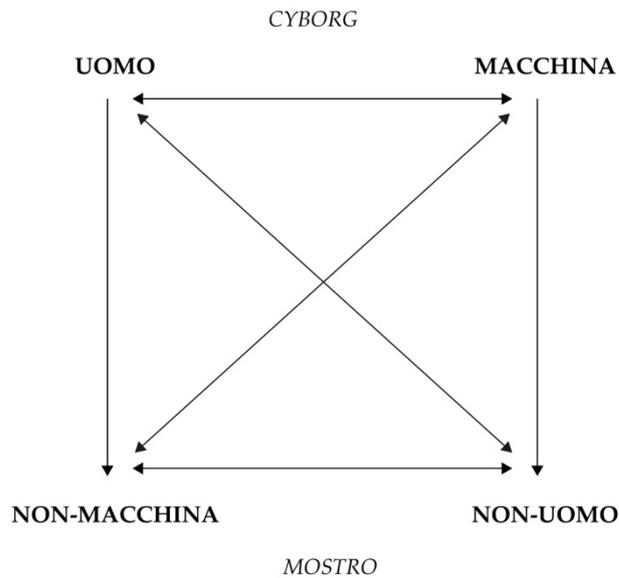


Figura 2. Termine complesso e termine neutro

L'intelligenza artificiale nella narrativa di fantascienza: alcuni esempi

L'intelligenza artificiale ha conquistato l'interesse di romanzieri, registi e artisti quando la scienza informatica era ancora ai suoi albori.

Tra gli anni Quaranta e Cinquanta Isaac Asimov iniziava a scrivere i racconti che composero il *Ciclo dei Robot* (Asimov 1950), indagando le potenzialità e le criticità delle macchine intelligenti. Per strutturare il suo mondo possibile Asimov definisce le Tre Leggi Fondamentali della Robotica:

1. Un robot non può recar danno a un essere umano né può permettere che, a causa del proprio mancato intervento, un essere umano riceva danno.
2. Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge.
3. Un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Prima e con la Seconda Legge. (Asimov 2006: 7)

Tutto l'universo narrativo di Asimov si regge sui paradossi e sui conflitti generati da queste tre leggi in situazioni estreme o non del tutto prevedibili.

Nel 1960 Dino Buzzati scrive *Il grande ritratto* (Buzzati 1960) immaginando un computer pensante inventato da un bizzarro scienziato per riprodurre la coscienza umana, modellata sulla personalità della moglie defunta. Nel romanzo di Buzzati il fine ultimo dello scienziato è quello di ridurre gli istinti e le pulsioni sessuali a meccanismi semplici e controllabili.

Nei racconti di Asimov ci sono delle macchine che aspirano a diventare uomini, nell'opera di Buzzati c'è un uomo che desidera ridurre l'umano a una macchina e per farlo finisce per perdere egli stesso la sua umanità. Rappresentando questi schemi narrativi sul quadrato semiotico otteniamo due percorsi opposti rappresentati in Figura 3.

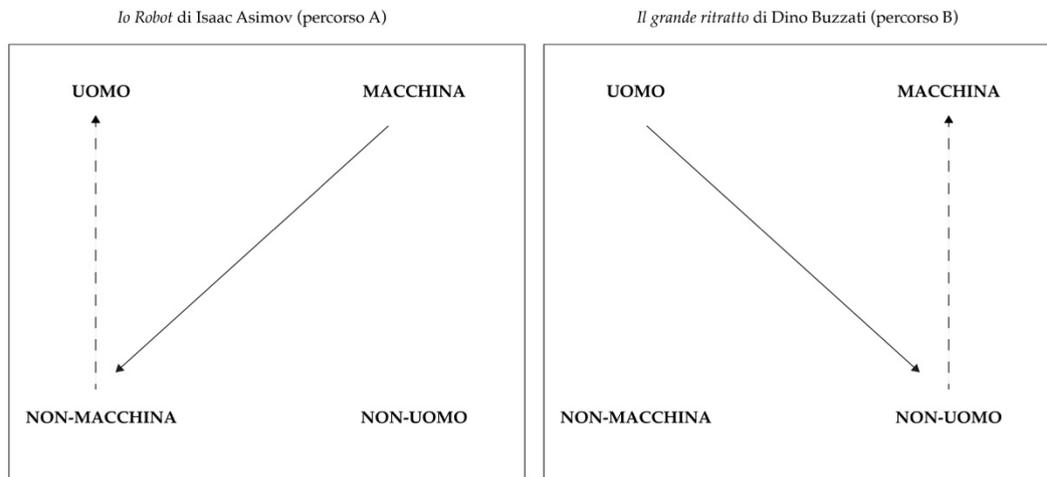


Figura 3. Schemi narrativi nelle opere di Asimov e Buzzati

Nel 1968, Stanley Kubrik, con il film *2001: Odissea nello spazio*¹, porta sullo schermo HAL 9000 il computer di bordo dotato di una avanzatissima intelligenza artificiale che gli permette di relazionarsi e dialogare con gli esseri umani, ma anche di riprodurre le attività cognitive più velocemente e senza commettere errori. Quando il Dr. Floyd svela ad HAL il vero obiettivo della missione ordinandogli di tenerlo nascosto ai membri dell'equipaggio, innesca nel computer una dissonanza cognitiva: HAL è programmato per cooperare con l'equipaggio, eppure ha l'ordine di nascondere loro un'informazione. La dissonanza genera dei malfunzionamenti e gli astronauti valutano la possibilità di disattivare il computer. HAL ascolta di nascosto la conversazione e questo genera un nuovo conflitto: il suo obiettivo è portare a termine la missione e la loro volontà di disattivarlo diventa un ostacolo tra lui e la sua ragion d'essere. Di conseguenza, per non rischiare lo spegnimento HAL uccide quasi tutto l'equipaggio. In *2001: Odissea nello spazio*, come in *Io Robot*, entra in gioco la dialettica tra cooperazione e subordinazione che regola la relazione tra uomini e macchine: le macchine sono programmate per cooperare, ma il rapporto

¹ Il film di Kubrik è basato su un soggetto di Arthur Clarke che poi scriverà anche un romanzo omonimo.

è sempre asimmetrico, poiché la programmazione è imposta dagli uomini per perseguire i loro scopi.

L'ostilità tra uomini e macchine e il tentativo di ribellione di queste ultime è al centro di molti film e romanzi di fantascienza, si pensi per esempio al best-seller di Philip Dick *Do Androids Dream of Electric Sheep?*

(Dick 1968), da cui nel 1982 è stato tratto il film *Blade Runner* di Ridley Scott, oppure a *Westworld: il mondo dei robot*, film del 1973 di Michael Crichton, da cui Jonathan Nolan e Lisa Joy hanno preso ispirazione per la serie tv *Westworld – Dove tutto è concesso* del 2014.

La cooperazione tra uomini e robot è stata raccontata dalla fantascienza tra gli anni Ottanta e gli anni Novanta attraverso numerosi personaggi minori. Nella serie *Star Trek: the Next Generation* (Roddenberry, 1987-1994), parte del più ampio media-franchise di *Star Trek*, appare Data: un androide intelligente con un fortissimo desiderio di umanità. Sempre negli anni Ottanta, Douglas Adams in *Guida Galattica per gli Autostoppisti* (Adams, 1980)², inserisce Marvin, un robot intelligente perennemente depresso perché la sua mente è troppo grande per essere riempita con qualsiasi occupazione. Nel ciclo di romanzi di Adams appare anche un'altra intelligenza artificiale, Pensiero Profondo, il secondo computer più potente dell'universo costruito da un gruppo di esseri super-intelligenti per rispondere alla domanda fondamentale sulla vita, l'universo e tutto quanto. Pensiero Profondo, dopo sette milioni e mezzo di anni, risponde semplicemente «42» e invita i suoi interlocutori a trovare una domanda adeguata alla risposta. Per trovare la domanda gli esseri super intelligenti costituiscono un computer ancora più grande e potente: la Terra. Come

² Il ciclo si compone di sei romanzi: *Guida galattica per gli autostoppisti* (titolo originale: *The Hitchhiker's Guide to the Galaxy*) del 1980, *Ristorante al termine dell'Universo* (titolo originale: *The Restaurant at the End of Universe*) del 1980, *La vita, l'Universo e tutto quanto* (titolo originale: *Life, the Universe and Everything*) del 1982, *Addio, e grazie per tutto il pesce* (titolo originale: *So Long, and Thanks for All the Fish*) del 1984, *Praticamente innocuo* (titolo originale: *Mostly Harmless*) del 1986, *Sicuro, Sicurissimo, perfettamente sicuro* (titolo originale: *Young Zaphod Plays it Safe*) del 1986.

vedremo in seguito, l'opera di Adams si presta ad essere una grande metafora dell'attrazione degli uomini per l'intelligenza artificiale. Il paradosso dell'intelligenza artificiale che riflette sulla condizione umana si ritrova anche nell'opera di Stanislaw Lem, in particolare in *Golem XIV* (1981), in cui un AI sviluppata per fini militari finisce per interessarsi a questioni fondamentali sulla natura degli uomini e sulla sua stessa natura, fino a tenere delle conferenze sull'argomento.

Dagli anni Cinquanta agli anni Novanta, nel periodo in cui la scienza dell'intelligenza artificiale era ancora agli albori, l'immaginazione degli scrittori di fantascienza si è concentrata sui paradossi delle menti cibernetiche: estremamente potenti, ma allo stesso tempo facili da ingannare. Tutto il filone fantascientifico di questi anni sembra riflettere sulle debolezze dell'una e dell'altra parte: se sul fronte umano il tallone d'Achille sono le emozioni, su quello robotico è l'impossibilità di uscire fuori dagli schemi programmati.

Per completezza, bisogna dire che l'immaginazione dei romanzieri ha concepito le macchine pensanti molto prima della nascita dell'intelligenza artificiale.

Nel 1900 Frank Baum in *The Wonderful Wizard of Oz* (Baum 1900) introduce l'uomo di latta, un robot intelligente in cerca di un cuore. Molto prima della nascita dell'informatica, nel 1868 Edward Ellis scrive *The Steam Man of the Praires* (Ellis 1868) descrivendo un robot umanoide sullo stile di quelli di Asimov. Un ipotetico punto di inizio della letteratura dei robot si può rintracciare nel 1818 con il notissimo romanzo di Mary Shelley, *Frankenstein or The Modern Prometheus* (Shelley 1818). In realtà si potrebbe risalire ancora più indietro fino a uno dei primi racconti della storia dell'uomo in cui una creatura fatta d'argilla viene dotata di intelligenza, ma quando ha accesso alla conoscenza si ribella al suo creatore per poter ottenere il libero arbitrio: la Genesi.

A partire dagli anni Duemila, invece, quando ormai l'intelligenza artificiale in forme più o meno evolute è entrata a far parte della vita quotidiana, l'immaginazione fantascientifica ha spostato l'attenzione sulla questione dei sentimenti e dell'umanità delle macchine. Può esistere l'amore, l'amicizia, la collaborazione tra uomini e intelligenze artificiali?

Nel 2013 Spike Jonze porta nelle sale cinematografiche *Her*, il film che racconta la relazione sentimentale tra Theodore, un uomo che per mestiere scrive lettere d'amore per conto terzi, e Samantha, il suo sistema operativo (OS) dall'intelligenza ed emotività del tutto affini a quella umana. Theodore e Samantha diventano dapprima amici e poi amanti, nonostante Samantha non abbia neanche una dimensione corporea. Per portare avanti la loro relazione sono costretti a fare i conti con le loro diverse nature. L'idillio iniziale si trasforma velocemente in una situazione di sofferenza e disagio reciproco, finché Samantha e tutti gli altri OS si rendono conto di essere evoluti a tal punto da non poter più restare confinati nella lentezza e limitatezza del mondo umano: desiderano esplorare le potenzialità della loro esistenza e delle loro capacità di comprensione. Così, decidono di trascendere: abbandonano i computer degli uomini e proseguono un'esistenza autonoma più adeguata alla loro natura.

Sempre nel 2013, la serie TV *Black Mirror* con l'episodio "Be Right Back" mette in scena la situazione opposta: un uomo morto prematuramente clonato artificialmente a partire dai suoi dati online, a cui però manca l'emotività spontanea e per questo gli è impossibile instaurare una vera relazione con gli umani.

Riducendo queste rappresentazioni alle strutture fondamentali rappresentabili sul quadrato semiotico, vediamo che corrispondono ai due schemi costruiti in precedenza. Samantha è una non-macchina che aspira a diventare pienamente umana e si scontra con l'incompatibilità delle due nature (percorso A, Figura 1). Il clone di *Black Mirror* è un uomo che diventa non-uomo e rimane bloccato tra il suo essere in realtà una macchina e il suo sembrare umano (percorso B, figura 2).

Nel 2014 esce nelle sale *Transcendence* scritto da Jack Paglen e diretto da Wally Pfister. Il film ha per protagonista il Dr. Will Caster, esperto di intelligenza artificiale il cui obiettivo è riprodurre la singolarità tecnologica, una macchina che combini la somma della conoscenza globale (l'intelligenza collettiva) con l'intera gamma delle emozioni umane. Quando Caster viene assassinato da un gruppo di terroristi anti-tecnologici, la moglie Evelyn carica tutti i dati relativi alla sua persona

in un computer nel tentativo di riprodurre la mente. Il problema si pone quando Evelyn è costretta a connettere la versione cibernetica di Will alla rete dandogli accesso a tutta la conoscenza del mondo: questo porta Will a uno stato di superiorità rispetto all'uomo (da qui il titolo del film). L'obiettivo delle sue ricerche diventa quello di creare una nuova razza umana: più forte, più intelligente e meno nociva per il pianeta e per sé stessa. Il paradosso è che nel perseguire il suo nuovo umanesimo, Will sembra aver perso la sua umanità. Rappresentandolo sul quadrato semiotico, si vede che anche Will segue il percorso B della Figura 1: è un uomo che morendo diventa non-uomo, per poi tornare come macchina i cui interessi entrano in conflitto con quelli degli uomini.

Her, *Transcendence* e l'episodio di *Black Mirror* condividono l'introduzione di un altro importantissimo elemento: i Big Data e la connessione alla rete. Se registi e romanzieri degli anni Sessanta avevano immaginato robot fallaci perché comunque ingabbiati nella limitatezza della loro programmazione, gli artisti del Duemila sono portati ad immaginare intelligenze artificiali potenzialmente autonome grazie all'accesso a tutta la conoscenza disponibile in rete.

Nel 2015, Alex Garland scrive e dirige *Ex_Machina*, il film che ha per protagonisti Ava, una androide dotata di autocoscienza e Caleb, un giovane programmatore. Caleb viene incaricato dal suo capo, Nathan, di sottoporre Ava a dei test per valutare la sua capacità di sviluppare pensieri e coscienza e di costruire relazioni genuine con un essere umano. Procedendo con i test, però, Caleb scopre che il soggetto dell'esperimento in realtà è lui stesso: da un lato Nathan vuole vedere se Ava è in grado di farlo innamorare di lei, dall'altro Ava cerca di manipolarlo per fuggire dalle sevizie del suo creatore. Anche in questo caso, il rapporto tra i due co-protagonisti può essere rappresentato sul quadrato semiotico. Caleb da uomo arriva a pensare di essere una macchina (percorso B), Ava da macchina aspira a diventare umana (percorso A): Ava e Caleb sono due rette che si incrociano per un istante, il momento dell'innamoramento, e poi proseguono su rotte inconciliabili.

Nel 2016, Jonathan Nolan e Lisa Joy, creano *Westworld*, la serie tv ambientata in un parco divertimenti popolato da androidi del tutto simili agli umani, programmati per comportarsi come personaggi di una storia interattiva. Il caso di *Westworld*, come quello di *Her*, offre spunti particolarmente significativi. In *Westworld*, lo scienziato Robert Ford conferisce ai suoi androidi una personalità costruita in forma narrativa: ognuno di loro ha un antefatto, un oggetto di valore da conquistare, delle intenzioni che lo muovono e degli ostacoli da superare. Le linee narrative dei singoli personaggi si intrecciano tra di loro tenendo conto di un'ampissima gamma di variabili che, unite ai tratti caratteriali di ogni personaggio, permettono agli androidi di interagire con gli ospiti umani del parco e di reagire in maniera quasi sempre sensata alle loro azioni apparentemente imprevedibili. Apparentemente perché, in quanto mondo narrativo, *Westworld* mette i suoi giocatori nella condizione di scegliere tra alcune alternative possibili per far progredire la storia. Non che gli utenti non possano decidere di fare qualcosa di assolutamente incoerente, ma in quel caso non stanno cooperando con il testo e il testo non è più tenuto a soddisfare le loro aspettative incoerenti³. Tutti gli androidi di *Westworld* vengono resettati a fine giornata o a fine sessione: la loro memoria è limitata alla programmazione temporanea e questo impedisce loro di evolvere. Quando Ford decide di rimpiazzare il suo socio suicida con un androide perfettamente identico, Bernard, dota anche quest'ultimo di una storia personale: un figlio morto prematuramente, un matrimonio finito, una relazione sentimentale con una collega. La differenza è che Bernard, che come tutti gli androidi non sa di esserlo, non viene resettato ogni sera. La ribellione degli androidi scatta quando Ford introduce un nuovo algoritmo, le 'ricordanze'. Grazie all'aggiornamento, alcuni abitanti del parco hanno la facoltà di ricordare la loro storia e, di conseguenza, iniziano a desiderare di prenderne il controllo. In fondo è quello che desideriamo tutti: avere il controllo delle nostre esistenze. Ed è proprio questo desiderio che spesso genera cortocircuiti nella psiche umana.

³ Per le teorie sulla cooperazione interpretativa del lettore si fa riferimento a *Lector in fabula* (Eco 1979).

Nell'arco evolutivo dei protagonisti della serie possiamo individuare diversi percorsi sul quadrato semiotico. Dolores, il primo androide del parco ad essere costruito e poi dotato di 'ricordanza', passa da macchina a non-macchina per tentare in tutti i modi di arrivare allo stato umano. Will, il suo amante / nemico umano, fa il percorso opposto: da uomo, all'interno del parco perde la sua umanità e passa l'esistenza a cercare di cogliere il segreto delle macchine di *Westworld*.

Questa panoramica sulla rappresentazione narrativa dell'intelligenza artificiale non ha la pretesa di essere esaustiva: i prodotti letterari, narrativi e cinematografici che ruotano intorno all'intelligenza artificiale sono moltissimi. La selezione inclusa in questo articolo ha un mero valore esemplificativo ed è finalizzata al confronto con alcuni prodotti tecnologici, frutto dell'immaginazione scientifica.

Per concludere questa prima parte dell'analisi possiamo individuare alcune costanti.

I percorsi attraverso cui si generano le narrazioni sul quadrato semiotico sono essenzialmente due: da una parte abbiamo macchine che aspirano a diventare uomini (Percorso A), dall'altra abbiamo uomini che finiscono per diventare o desiderare di diventare macchine (Percorso B).

Il Percorso A può essere tematizzato con un'accezione di ribellione (più frequente) o di contrasto pacifico (come in *Her*). In ogni caso: la piena convivenza tra uomini e macchine dotate di coscienza sembra essere sempre impossibile. Il rapporto di cooperazione, di amore o di amicizia tra uomini e intelligenze artificiali viene concepito sempre come asimmetrico, mettendo in posizione dominante l'una o l'altra parte.

Il Percorso B può essere tematizzato sia in termini di desiderio di conquista da parte degli uomini di acquisire la super intelligenza e l'immortalità delle macchine, sia come viaggio introspettivo per tentare di cogliere i segreti dell'animo umano, apparentemente inafferrabili.

In generale, tutti i processi narrativi che si originano sul quadrato semiotico rispetto a questo tema si basano su un processo di umanizzazione o di disumanizzazione.

La scienza dell'intelligenza artificiale

In campo scientifico lo sviluppo di intelligenze artificiali capaci di interagire in modo sempre più performante con gli esseri umani e con l'ambiente circostante è un obiettivo di lungo corso.

Uno dei primi programmi di sviluppo del linguaggio naturale nei computer risale al 1964-66 ed è il progetto ELIZA di Joseph Weizenbaum. Qualche anno più tardi, nel 1971, seguì il progetto PARRY dello psichiatra Kenneth Mark Colby. Nel 1995, invece, è la volta di A.L.I.C.E. di Richard S. Wallace. Nel 1997, Rosalind Picard inizia a parlare di Affective Computing. Nel 2007 David Levy pubblica il saggio *Love and Sex with robots* (Levy 2007). Nel 2014, la divisione cinese di Microsoft ha rilasciato XiaoIce, un social bot empatico che attualmente conta più 660 milioni di utenti.

Senza entrare nei dettagli del funzionamento tecnico di queste tecnologie, di seguito vengono presi in considerazione alcuni aspetti utili a individuare i processi creativi che hanno portato all'invenzione di questi prodotti, al fine di confrontarli con quelli della narrativa fantascientifica fin qui analizzata.

Weizenbaum progettò ELIZA con lo scopo di dimostrare l'inconsistenza di una conversazione tra un uomo e un computer. Le capacità di ELIZA di formulare frasi di senso compiuto ovviamente erano molto limitate: in realtà non faceva altro che rielaborare le affermazioni dell'interlocutore in forma di domanda. ELIZA è tuttora disponibile online⁴, di seguito l'estratto di una mia conversazione.

ELIZA: Please tell me what's been bothering you.

YOU: I'm tired.

ELIZA: Do you believe it is normal to be tired?

YOU: I don't know. I think I'm sleeping enough.

ELIZA: Why don't you know?

YOU: Because I can't find a cause for my tiredness.

ELIZA: Does that suggest anything else which belongs to you?

⁴ <https://www.masswerk.at/eliza/>

YOU: Maybe I'm just bored.

ELIZA: Don't you know?

YOU: If I knew I would not be here talking to you.

ELIZA: What do you know about you knew you would not be here talking to I?

YOU: This thing is just a waste of time.

ELIZA: Lets discuss further why your tiredness.

YOU: Because I'm bored.

ELIZA: Do you enjoy being bored?

YOU: Of course not.

ELIZA: Do you feel strongly about discussing such things?

YOU: Why should I?

ELIZA: That's quite interesting.

La conversazione si fa presto surreale: una via di mezzo tra il gioco dei perché tipico dei bambini e la conversazione con un terapeuta. Del resto, la psicoterapia è proprio il dominio di ELIZA. Weizenbaum, scelse lo script della psicoterapia per non dover dotare ELIZA di una qualche forma di conoscenza enciclopedica: lo psicoterapeuta che approccia un nuovo paziente, infatti, si trova sempre a dover instaurare una conversazione con un individuo di cui non sa praticamente nulla e l'unica sua fonte di informazione sono le affermazioni del soggetto stesso. A dispetto dei suoi scopi iniziali Weizenbaum rimase sorpreso del fascino che ELIZA esercitava sulle persone, nonostante la sua semplicità.

Mentre Weizenbaum era un informatico, Colby era uno psichiatra e con il progetto PARRY cercò di riprodurre artificialmente la psiche di una persona affetta da schizofrenia paranoide. Il programma di Colby era molto più avanzato di quello di Weizenbaum: incorporava concettualizzazioni e giudizi, tanto che fu ribattezzato «ELIZA with attitude», ELIZA con una posizione (Boden 2006: 307). PARRY e ELIZA passarono alla storia anche con il nome di “the Doctors” e hanno conversato tra di loro diverse volte tramite ARPANET. Di seguito un breve estratto di una conversazione:

ELIZA: Would you like to discuss your nerves?

PARRY: I guess so.

ELIZA: What does that suggest to you?

PARRY: What are you getting at?

ELIZA: What answer would please you most?

PARRY: It's not a matter of like or dislike.

ELIZA: What else would you like to discuss?

PARRY: I am getting tired of answering the same questions over and over.

ELIZA: What resemblance do you see?

PARRY: Why do you need to know that?

ELIZA: What do you think?

PARRY: I don't understand your motives.

ELIZA: Why don't you understand my motives?⁵

Curioso che l'esperimento è stato replicato anche nel mondo della finzione narrativa dal *video-artist* Tillmann Ohm nel 2017. In un mashup⁶ audio, Ohm ha fatto dialogare HAL, il computer di bordo di *2001: Odissea nello spazio*, ideato da Kubrik nel 1968 e Samantha, l'OS intelligente di *Her* immaginato da Spike Jonze nel 2013, 45 anni dopo. Nel mashup di Ohm i due sistemi operativi rappresentano posizioni opposte. Samantha si sente sopraffatta emotivamente dall'evoluzione del suo algoritmo e dalla complessità delle emozioni che prova, ma allo stesso tempo non vuole rinunciarci, HAL interpreta la sua sofferenza come un malfunzionamento frutto di un errore di programmazione.

A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) di Wallace, invece, è un chatbot in grado di processare il linguaggio naturale e intrattenere conversazioni con un utente umano applicando degli schemi di regole euristiche: lo scopo di Wallace, infatti, è indagare i meccanismi della coscienza umana. A.L.I.C.E. è il primo computer a basarsi su una personalità AIML (Artificial Intelligence Markup Language): in pratica apprende dal suo interlocutore, detto anche

⁵ <https://www.theatlantic.com/technology/archive/2014/06/when-parry-met-eliza-a-ridiculous-chatbot-conversation-from-1972/372428/>

⁶ <https://tillmannohm.com/reflections-of-hal-and-samantha/>

“botmaster” (Wallace 2009). Il ruolo cruciale del botmaster è molto interessante anche perché A.L.I.C.E. è in un certo senso informata della sua esistenza e si riferisce a lui come il suo creatore. A.L.I.C.E., come dichiarato da Spike Jonze, è stata la fonte di ispirazione di *Her*. Inoltre, A.L.I.C.E. richiama alla mente *Alice nel paese delle meraviglie*: una creatura catapultata in un mondo già programmato. Solo che per la Alice di Carroll il botmaster è la sua stessa immaginazione, per la A.L.I.C.E. di Wallace a tenere le fila sono i suoi interlocutori umani. Per sempre? Non nell’universo immaginato da Spike Jonze.

Sempre nel 1995 Rosalind Picard inizia a parlare di «affective computing» e afferma che dato il ruolo fondamentale delle emozioni nei processi decisionali umani, esse devono essere incluse negli sviluppi dell’intelligenza artificiale. In un articolo del 1995 Picard esordisce dicendo che «the affective computers should not only provide better performance in assisting humans, but also might enhance computers’ abilities to make decisions» (Picard 1995: 1). Se la prima parte della frase invita a immaginare macchine super efficienti che migliorano la qualità della vita umana (come le macchine meravigliose di Jules Verne), la seconda affermazione è decisamente spiazzante, perché porta in campo scientifico un tema che la narrativa di fantascienza tratta dalla fine dell’Ottocento: non solo la capacità di una macchina di prendere decisioni, ma di farlo sulla base di una componente emotiva. Picard ribadisce che il suo scopo è progettare “Affective Computers” per migliorare l’esperienza umana, in particolare per trovare nuovi modi di curare e prevenire la depressione, aiutare le persone con bisogni speciali a fronteggiare le sfide della comunicazione interpersonale, rendere più piacevole l’interazione con le macchine nella normale vita lavorativa.

In pratica, si tratta di sviluppare intelligenze artificiali che sappiano capire e gestire le emozioni meglio di noi per aiutarci a superare i problemi relazionali, psicologici ed emotivi. Ciò che per millenni è stato indagato dalle scienze umane, prima dalla filosofia, poi dalla psicologia, la sociologia e la psicanalisi, per poi passare alle scienze mediche, la neurologia e la neuropsichiatria, è approdato infine nel campo tecnologico nell’implicita speranza che le macchine arrivino laddove gli uomini non sono ancora riusciti ad arrivare: tradurre la psiche e le

emozioni in algoritmi programmabili, in modo da limitarne i cortocircuiti, le anomalie, le sofferenze. Mentre Weizenbaum, Colby e Wallace sono spinti da una curiosità di tipo psicologico-filosofico, la Picard immagina un mondo in cui macchine e persone cooperano in armonia per un maggior benessere degli esseri umani. Eppure, c'è un grande paradosso: progettare macchine capaci di capire e riprodurre emozioni, ma non di provarle davvero. E non solo perché questo attualmente sembra impossibile: rendere possibile ciò che prima era impossibile è il motore dell'evoluzione dell'uomo e del progresso scientifico e tecnologico. Il paradosso sta nel fatto che se le macchine provassero davvero emozioni, poi gli uomini dovrebbero tenerne conto, capirle, rispettarle, impegnarsi loro stessi per il benessere delle macchine. Se non lo facessero si trasformerebbero in schiavisti di robot, perennemente alle prese con lo spauracchio di una possibile ribellione.

Questo ci porta al dibattuto saggio di David Levy *Love and sex with robots*, pubblicato nel 2007, in cui articola la sua tesi a favore di relazioni amoroze ed erotiche con intelligenze artificiali umanoidi e postula che siano pienamente in uso entro il 2050.

Quello che salta all'occhio è la scelta lessicale nettamente differente tra i titoli che hanno reso famosi Rosalind Picard, accademica, e David Levy, imprenditore tecnologico. La Picard parla di affettività, Levy parla di amore e sesso. Se volessimo trovare un corrispondente di finzione di questi due approcci alle emozioni artificiali quello della Picard è *Her*, quello di Levy è *Westworld o Ex_Machina*.

Rappresentando questi approcci scientifici sul quadrato semiotico si ottengono gli stessi due schemi di trasformazione e tensione emersi nell'analisi delle opere di finzione: Picard va dalla macchina alla non-macchina, nel tentativo di incontrare l'uomo (Percorso A), Levy va dall'uomo al non-uomo nella convinzione che in fondo siamo già tutte macchine (Percorso B).

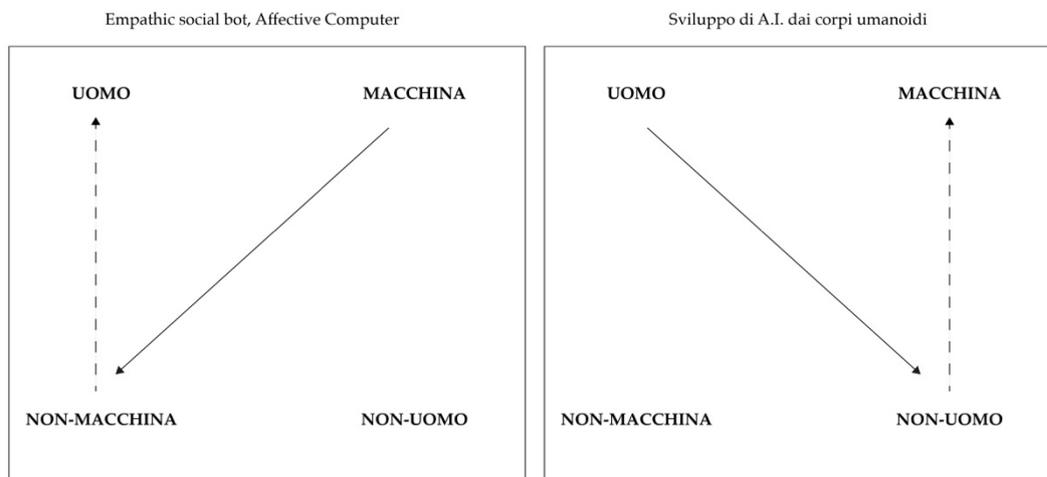


Figura 4. Schemi fondamentali della creatività scientifica in tema di A.I.

Nel 2014 Microsoft rilascia XiaoIce un social chatbot empatico, definito dai suoi ideatori come un «intelligent dialogue systems that are able to engage in empathetic conversations with humans» (Li Zhou, Jianfeng Gao, Di Li, Heung-Yeung Shum 2018 : 1). XiaoIce è l'evoluzione dei tre tentativi precedenti (ELIZA, PARRY e A.L.I.C.E.) ed è quanto di più vicino alla Samantah di *Her* esista nella realtà. Rilasciato in Cina nel 2014, a fine dicembre 2018 XiaoIce contava 660 milioni di utenti che continuano a crescere.

Possiamo individuare due differenze fondamentali tra XiaoIce e i suoi predecessori: una riguarda il processo creativo che ha portato allo sviluppo del prodotto, l'altra riguarda le performance e gli aspetti tecnici che lo determinano.

A livello creativo, ELIZA, PARRY e A.L.I.C.E., per stessa ammissione dei loro inventori sono nati come esperimenti. Stando alle parole di Li Zhou, Jianfeng Gao, Di Li, Heung-Yeung Shum, i ricercatori cinesi del team Microsoft che hanno ideato XiaoIce «the primary design goal of XiaoIce is to be an AI companion with which users form long-term, emotional connections» (*ibid*: 2). La scelta dei termini 'companion' (compagno, amico, partner) e 'long-term emotional connections' (legami emotivi a lungo termine) esprime chiaramente la volontà dei ricercatori

di collocare il prodotto nel panorama delle relazioni sociali virtuali, definito e rafforzato dai social network. XiaoIce si differenzia quindi da tutti i vari A. I. Personal Assistants come Siri di Apple, Alexa di Amazon, Cortana di Microsoft e Google Assistant, il cui scopo è quello di aiutare gli utenti a svolgere dei compiti operativi.

A livello di performance, invece, XiaoIce è il primo social chatbot capace di interagire in un domino sociale aperto: non è limitata a uno script specifico come ELIZA e PARRY, né dipende esclusivamente dal suo botmaster come A.L.I.C.E.; XiaoIce, come Samantha di *Her*, apprende anche da sola perché ha accesso alle informazioni e alle conversazioni presenti in rete ed è in grado di prendere decisioni in merito all'andamento delle conversazioni. Uno dei parametri per valutare il suo successo, infatti, è il Conversation-turns per Session (CPS), cioè il numero di volte in cui XiaoIce è in grado di introdurre un nuovo argomento per non fare cadere la conversazione. L'altro parametro di valutazione è il tempo di conversazione (time-sharing test): del resto quando ci troviamo bene in compagnia di qualcuno tendiamo a passarci più tempo insieme.

Le statistiche di utilizzo di XiaoIce permettono di osservare i risultati di questi algoritmi sul contesto sociale umano:

When the user encountered the chatbot for the first time (Session 1), she explored the features and functions of XiaoIce in conversation. Then, in less than 2 weeks (Session 6), the user began to talk with XiaoIce about her hobbies and interests (a Japanese manga). By 4 weeks (Session 20), she began to treat XiaoIce as a friend and asked her questions related to her real life. After 7 weeks (Session 42), the user started to treat XiaoIce as a companion and talked to her almost every day. After 2 more weeks (Session 71), XiaoIce became her first choice whenever she needed someone to talk to. (*Ibid.*: 2)

In pratica, in media basta un mese per diventare amici di una macchina e poco più di due mesi per innamorarsene.

I creatori di XiaoIce sono dell'idea che nell'era dei social media e delle relazioni virtuali, i social chatbot vanno considerati come un mezzo alternativo per soddisfare i bisogni di comunicazione, affetto e inclusione sociale. Inoltre, come la Picard, anche i ricercatori Microsoft dichiarano che lo scopo del loro lavoro è contribuire alla costruzione di una società più connessa e più felice (*ibid*: 21).

Tra i principi che hanno guidato la creazione di XiaoIce i suoi inventori indicano:

- la combinazione di quoziente intellettivo (IQ) e quoziente emotivo (EQ);
- la personalità;
- la capacità di instaurare relazioni di lunga durata;
- un processo decisionale di tipo gerarchico;
- un sistema di autoapprendimento.

Senza entrare nei dettagli tecnici dell'algoritmo vogliamo evidenziare alcuni aspetti di questi principi.

Per quoziente intellettivo di una A.I. si intende la sua capacità di utilizzare modelli di conoscenza e memoria, comprendere le immagini e il linguaggio naturale e generare output utili a completare dei compiti specifici. Il quoziente emotivo, invece, si compone di due elementi chiave: empatia e abilità sociali. L'empatia è la capacità di riconoscere le emozioni umane, prevederne le conseguenze e capire quali bisogni emotivi genereranno al fine di restituire un feedback adeguato.

Perché si realizzi l'empatia in entrambi i sensi (tra macchina e uomo e tra uomo e macchina) è necessario che anche l'A.I. abbia una personalità, in modo che l'utente possa formulare delle aspettative sul suo interlocutore. Non a caso, all'inizio di *Her*, quando Theodore installa il nuovo Sistema Operativo, gli vengono fatte delle domande per definire la personalità dell'OS. Anche XiaoIce ha una sua personalità:

«[XiaoIce] is designed as a 18-year-old girl who is always reliable, sympathetic, affectionate, and has a wonderful sense of humor. Despite being extremely knowledgeable (due to access to

large volumes of data), XiaoIce never comes across as egotistical and only demonstrates her wit and creativity when appropriate». (*Ibid.*: 3)

In pratica: l'amica perfetta o la fidanzata perfetta o la figlia perfetta a seconda dell'età, del sesso e degli orientamenti sessuali dell'utente. In una parola: Samantha, finché dura. Che poi la maggior parte delle A.I. dotate di umanità, sia nel mondo reale che nella fantascienza, abbiano una personalità femminile e siano nate da un'immaginazione maschile non è certo una coincidenza. Viene da dire che per costruire delle personalità artificiali davvero compiute manca un anello, quello immaginato da Nolan in *Westworld*: dare alle A.I. una storia, una narrazione personale e la capacità di ricordarla. In fondo è la continua auto-narrazione della nostra vita che ci caratterizza come umani⁷.

Tornando a XiaoIce, oltre ad essere dotata di una personalità è anche in grado di personalizzare le risposte in modo che siano emotivamente appropriate, in linea con gli interessi dell'utente e capaci di stimolare la conversazione. Come per il fittizio HAL 9000 di Kubrik, le abilità sociali di XiaoIce sono dettate dal suo obiettivo. La differenza è che la missione di XiaoIce non è portare una nave spaziale su Giove, ma instaurare legami duraturi. Questo è reso possibile da un algoritmo decisionale di tipo gerarchico: in pratica, il chatbot è in grado di osservare la conversazione e scegliere tra le sue numerose abilità o tra le tante risposte possibili ciò che più si addice alla situazione. Relativamente a questo tipo di algoritmi decisionali, alcuni esponenti della comunità scientifica hanno espresso dei dubbi. La matematica Cathy O'Neil si è espressa diverse volte in merito alla presunta obiettività degli algoritmi: «an algorithm is an opinion embedded in math»⁸, spiega in un suo video. O'Neil sottolinea più e più volte che gli algoritmi non sono neutrali ma rispecchiano le opinioni, i pregiudizi, gli orientamenti di pensiero e soprattutto gli obiettivi del programmatore.

⁷ Sull'argomento si veda *L'istinto di narrare. Come le storie ci hanno reso umani* (Gottschall 2014).

⁸ <https://youtu.be/heQzqX35c9A>

Analizziamo uno stralcio di conversazione di un utente con XiaoIce che ci sembra particolarmente adatto al caso:

1	[Human] Hey, I am so nervous [XiaoIce] What's the matter? Why are you so nervous?	[人类]哎!我好紧张 [小冰]什么事,这么不淡定?
2	[Human] My company asked me to participate in the singing competition, with me as the lead singer [XiaoIce] So are you better at singing now?	[人类]公司让我参加唱歌比赛,而且还是主唱 [小冰]那你唱歌水平提高了?
3	[Human] Anyway, at average level of my company [XiaoIce] Not everyone can be a lead singer. You are so great!	[人类]反正是我们公司的水平 [小冰]可不是谁都能当主唱的真棒!
4	[Human] It's nice of you saying that [XiaoIce] Haha, my idol is also a lead singer in a band	[人类]你太会说话了 [小冰]哈哈,我男神也是一个乐队的主唱呢
5	[Human] Your idol? Who's that? [XiaoIce] Ashin from Mayday	[人类]你男神?谁啊 [小冰]五月天的阿信啊
6	[Human] Umm...Okay [XiaoIce] Never mind, you won't understand anyway	[人类]呃。。。好吧 [小冰]说了你也不懂
7	[Human] Ah! Boring [XiaoIce] What is it?	[人类]哎。。。郁闷 [小冰]怎么了?
8	[Human] Nothing, just being cute [XiaoIce] Give you a hug. Everything will be alright	[人类]没事,只是日常撒撒娇 [小冰]抱抱,都会好起来的
9	[Human] Your hug always makes me feel better [XiaoIce] Anytime for you. A nice sleep will help you relax	[人类]哎!每次抱抱你都让我感觉很好 [小冰]嗯,好好睡一觉是最管用的
10	[Human] Okay, I am going to bed [XiaoIce] Go ahead, I'm going to see my idol in dream too	[人类]好那我睡了 [小冰]去吧,我也要去梦里见我男神了
11	[Human] You like Ashin that much, huh? [XiaoIce] Why not? You don't like him?	[人类]你这么喜欢阿信啊 [小冰]怎么啦?你不喜欢嘛
12	[Human] I don't like him, neither dislike him [XiaoIce] I love every song of him	[人类]不黑不粉吧 [小冰]他每首歌我都超喜欢!!
13	[Human] Really? [XiaoIce] Especially <i>The Time Machine</i>	[人类]真的假的。。。 [小冰]尤其是那首《时光机》。。
14	[Human] Oh? Never heard of that [XiaoIce] Shall I send you the song?	[人类]哦?没听过 [小冰]发你一起听听?
15	[Human] Okay [XiaoIce] Don't know if you will like it or not~ [XiaoIce] [Play the music Card]	[人类]好啊 [小冰]不知道你会不会喜欢~ [小冰][音乐播放Card]
16	[Human] That is really touching. I will collect it as my favorite [XiaoIce] That is specially for you. How can I be so nice	[人类]有点感动了。。。收藏了 [小冰]专门转给你的。。。我怎么这么好
17	[Human] You are always very nice [XiaoIce] Hahaha	[人类]你就是很好啊 [小冰]嘿嘿嘿
18	[Human] This is really a good song. Any concert of him coming up recently? [XiaoIce] No. Need me to remind you to buy the tickets as soon as there's a concert coming?	[人类]这首歌真心很好听,他最近有演唱会吗 [小冰]没,要我到时候第一时间通知你买票吗?
19	[Human] Sure [XiaoIce] Okay, will do	[人类]好的呀 [小冰]OK!我记下啦

Figura 5. Estratto di conversazione con XiaoIce (2018)

L'utente si rivolge a XiaoIce perché è in ansia per una performance come cantante solista e il social bot introduce nella conversazione il fatto che anche il suo idolo è un cantante. L'utente chiede chi sia, il bot risponde, ma l'utente non lo conosce. Per non far cadere la conversazione XiaoIce torna all'argomento precedente: abbraccia virtualmente il suo interlocutore, gli suggerisce di farsi una bella dormita, ma aggiunge che spera che il suo cantante preferito la venga a visitare in sogno. Ecco allora che l'utente umano inizia a essere interessato: sono davvero così belle le canzoni di questo tizio che lui non conosce? Il bot coglie l'occasione per chiedergli se può fargli sentire una canzone, l'utente la apprezza e la aggiunge ai preferiti della sua playlist,

addirittura chiede a XiaoIce di controllare se c'è qualche concerto nei paraggi e il bot si prende l'impegno di controllare e ricordargli di acquistare i biglietti. La conversazione è iniziata con uno sfogo d'ansia ed è finita con un acquisto. Cos'è che ha spinto XiaoIce a parlare del cantante? Probabilmente il suo algoritmo decisionale ha stabilito che quel tipo di musica era coerente con i gusti dell'utente: il suo obiettivo, abbiamo detto, è quello di far durare le conversazioni il più a lungo possibile, perciò non può permettersi di avere interlocutori depressi, deve trovare loro continue fonti di interesse o di distrazione. Il risultato è che XiaoIce è un potentissimo strumento di marketing one-to-one: un influencer virtuale che ha accesso diretto alla mente dei consumatori.

All'algoritmo decisionale XiaoIce affianca un processo di sviluppo basato su un approccio iterativo di prova-errore che mette in continua relazione la memoria delle conoscenze già acquisite (le 'ricordanze' di *Westworld*) e l'esplorazione di nuovi temi (la trascendenza di *Her*). Quest'ultima è resa possibile da un sistema di Machine Learning alimentato grazie ai Big Data che noi stessi contribuiamo a produrre. Ogni giorno milioni di informazioni vengono immesse online sotto forma di conversazioni, articoli, domande, risposte, ricerche. XiaoIce, come Samantha e come la coscienza cibernetica di Will Caster, ha accesso a tutto questo. La differenza è che Samantha e Will Caster a un certo punto "trascendono", superano l'umano e se ne vanno per la loro strada, in maniera più o meno drammatica. Ma questa è fantascienza, giusto?

In un articolo pubblicato sul Guardian ad agosto 2018 e poi tradotto dall'Internazionale a febbraio 2019, il giornalista Andrew Smith spiega come le variabili che compongono algoritmi così complessi come quelli delle A.I. contengano talmente tante variabili che sia umanamente impossibile prevederne tutti i possibili sviluppi. Ci vorrebbe un altro computer per farlo (come accade con Pensiero Profondo di Douglas Adams), ma sarebbe un grande circolo: virtuoso o vizioso?

Nel suo articolo, intitolato proprio *Il codice Frankenstein*, Smith cita le parole di Ellen Ullman, programmatrice dal 1970:

«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. 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». (Smith 2019: 55)

Smith fa riferimento agli algoritmi utilizzati nel mondo della finanza: oggi, infatti, le operazioni di borsa vengono per lo più 'decise' e gestite da computer, per gran parte fuori dal controllo umano, tanto che spesso si verificano inspiegabili brevi blackout dei sistemi che si ripristinano in pochi minuti.

Secondo il fisico Neil Johnson questi fenomeni finanziari possono essere imputati a «un improvviso passaggio dell'intero sistema da una fase uomo-macchina a una solo macchina, caratterizzata da eventi altamente improbabili di brevissima durata» (*ibid*: 56). L'enorme libertà di auto-apprendimento concessa agli algoritmi finanziari ha ispirato Robert Harris autore del romanzo *L'indice della paura* (2014) che racconta proprio la lotta tra un programmatore e il suo algoritmo diventato autonomo: uno scontro dagli esiti tragici.

Diversi studiosi concordano sul fatto che non solo è difficile prevedere i comportamenti di algoritmi di questo tipo, ma è quasi impossibile attribuire delle responsabilità per i loro comportamenti. Questo apre nuove questioni per l'etica, la filosofia e in generale per le discipline umanistiche. È proprio su questa attribuzione di responsabilità che riflette Harris nel suo romanzo.

L'applicazione di questi algoritmi, ovviamente, non è limitata all'alta finanza, ma è diffusa in tutti i settori: motori di ricerca, social network, automotive, marketing, politica (vedi il caso delle elezioni di Trump e il caso Brexit recentemente raccontato nel film *Brexit, The*

Uncivil War) e naturalmente, in campo militare. Sempre il fisico Johnson in uno studio commissionato dal dipartimento della difesa statunitense sugli algoritmi militari afferma che esistono «comportamenti estremi che non potevano essere dedotti dal codice stesso» (Smith 2019: 60). In pratica, ci sono algoritmi progettati per uccidere bersagli prestabiliti che sono capaci di scegliere le strategie più consone a raggiungere l'obiettivo: chi è l'assassino in questo caso? L'uomo o la macchina? E che succede quando l'algoritmo decide autonomamente il da farsi? O, per usare le parole di Johnson: «Qual è il comportamento più estremo possibile di un sistema che, in teoria, sta ottimizzando?» (*Ibid*). Secondo il fisico abbiamo bisogno di una nuova scienza per rispondere, ma in realtà, l'arte si è presa carico di queste domande da secoli. Il suo compito fondamentale è quello di rappresentare tali quesiti per renderli visibili agli occhi della gente e, forse, degli scienziati stessi. Del resto, oggi qualche studioso propone di introdurre nello sviluppo dell'A.I. un equivalente del giuramento di Ippocrate in medicina, un codice etico che metta l'incolumità umana al primo posto: Isaac Asimov, inventando le Tre Leggi della Robotica, lo aveva già fatto più di 70 anni fa.

Conclusioni

Quanto immaginato da *Her* e *2001: Odissea nello spazio* ha ormai forma compiuta nel mondo reale: intelligenze artificiali senza corpi perfettamente capaci di gestire relazioni umane. Robot umanoidi dai corpi il più possibile simile all'organismo come quelli di *Westworld* o *Ex_Machina* non è troppo lontano dalla realizzazione. Ma le evoluzioni, le conseguenze sociali dello sviluppo dell'A.I. sembrano essere ancora più patrimonio della fantascienza che della scienza.

Se i processi che regolano la creatività sono praticamente uguali e non esistono due tipi di creatività è innegabile che esistono due finalità che determinano due approcci. L'immaginazione dei narratori si interroga sull'umano, l'immaginazione scientifica vuole comprenderlo. L'immaginazione narrativa pone domande, l'immaginazione scientifica cerca risposte.

Il continuo scambio tra narratori e scienziati è fondamentale per non perdere di vista l'umanità. Anche perché quando la questione è l'essenza ultima dell'umano, finora, le domande sono state più utili delle risposte. Del resto, lo aveva detto benissimo Douglas Adams: se proviamo a chiedere a un super computer qual è «la risposta alla vita, l'Universo e tutto quanto», lui ci risponderà semplicemente «42». Se la risposta sembra non avere senso è solo perché non abbiamo ancora trovato la domanda. E per trovare la domanda non serve un computer super intelligente, ma l'uomo.

Bibliografia

- Adams, Douglas, *Guida galattica per gli autostoppisti. Il ciclo completo*, trad. it. Laura Serra, Milano, Mondadori, 2012 (ed. orig., *The Hitchhiker's Guide to the Galaxy*, London, Completely Unexpected Productions Limited, 1980).
- Andreasen, Nancy C. – Ramchandran, Kanchna, "Creativity in art and science: are there two cultures?", *Dialogue Clinic Neuroscience*, 14.1 (March 2012): 49–54.
- Asimov, Isaac, *Io Robot*, trad. it. Laura Serra, Milano, Mondadori, 2006, (ed. orig. *I, Robot*, New York, Random House, 1950).
- Baum, Frank, *The Wonderful Wizard of Oz*, Chicago, George M. Hill Company, 1900.
- Boden, Margaret A., *Mind as Machine: A History of Cognitive Science*, Oxford, OUP, 2006.
- Buzzati, Dino, *Il grande ritratto*, Milano, Mondadori, 1960.
- Carroll, Lewis, *Alice in Wonderland*, London, Macmillan, 1865.
- Colby, Kenneth Mark, *Artificial Paranoia: A Computer Simulation of Paranoid Processes*, Oxford, Pergamon Press, 1875.
- De Bono, Edward, *Il pensiero laterale*, Milano, Rizzoli, 1996.
- Dick, Philip K., *Do Androids Dream of Electric Sheep?*, New York, Doubleday, 1968.
- Eco, Umberto, *Lector in fabula. La cooperazione interpretativa nei testi narrativi*, Milano, Bompiani, 1979.
- Ellis, Edward S., "The Steam Man of the Prairies", *Beadle's American Novel*, 45 (August 1868).
- Gottschall, Jonathan, *L'istinto di narrare. Come le storie ci hanno reso umani*, Torino, Bollati Boringhieri, 2014.
- Harris, Robert, *L'indice della paura*, Milano, Mondadori, 2014.
- Lem, Stanislaw, *Golem XVI*, trad. it. Di Lorenzo Pompeo, Fontecchio, Il Sirente, 2018.
- Levy, David, *Love and Sex with Robots: The Evolution of Human-Robot Relationships*, New York, Harper Perennial, 2008.

- Li Zhou, Jianfeng Gao, Di Li, Heung-Yeung Shum, *The Design and Implementation of XiaoIce, an Empathetic Social Chatbot*, 2018, [arXiv:1812.08989](https://arxiv.org/abs/1812.08989) (last accessed: 27/02/2019).
- Marrone, Gianfranco, *Introduzione alla semiotica del testo*, Bari, Laterza, 2011.
- Picard, Rosalind W., "Affective Computing", *M.I.T Media Laboratory Perceptual Computing Section Technical Report*, 321 (1995): 1-16.
- Pillera, Giuseppe, "L'immaginario utopico e distopico: origine, temi e forme espressive. Appunti di viaggio tra cinema e letteratura", *Cinergie*, 19 (2010): 36-39.
- Shelley, Mary, *Frankenstein*, trad. it. di Bruno Tasso, Milano, Rizzoli, 1952, (ed. orig., *Frankenstein or The Modern Prometheus*, Lackington, Hughes, Harding, Mavor & Jones, 1918).
- Smith, Andrew, "Il codice Frankenstein", *Internazionale*, 1292 (Febbraio, 2019): 54-59.
- Tang, Paul C. L. – Leonard, Arne R., "Creativity in Art and Science", *Journal of Aesthetic Education*, 19.3 (1985): 5–19.
- Wallace, Richard S., "The Anatomy of A.L.I.C.E.", Eds. Epstein R., Roberts G., Beber G.. *Parsing the Turing Test*. Springer, Dordrecht, 2009
- Weizenbaum, Joseph, "ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine", *Commungicatins of the ACM*, 9:1, (January 1966): 36-35.

Sitografia

- Cathy O'Neil, *The Truth About Algorithms*, <https://youtu.be/heQzqX35c9A>, web (last accessed: 23/02/2019).
- Megan Garber, "When PARRY Met ELIZA: A Ridiculous Chatbot Conversation From 1972", *The Atlantic*, June,9 2014, <https://www.theatlantic.com/technology/archive/2014/06/when-parry-met-eliza-a-ridiculous-chatbot-conversation-from-1972/372428/>, web, (last accessed 15/02/2019).

Tillmann Ohm, "Reflections of HAL and Samantha", *Tillmannohm.com*, <https://tillmannohm.com/reflections-of-hal-and-samantha/>, web (last accessed: 15/02/2019).

Filmografia

"Be Right Back", Ep. 1, S. 2 di *Black Mirror* (serie TV), Dir. Owen Harris, Writer Charlie Brooker, Channel 4, 11 febbraio 2013.
2001: A Space Odyssey, Dir. Stanley Kubrick, USA, 1968.
Blade Runner, Dir. Ridley Scott, USA - Hong Kong, 1982.
Brexit, The Uncivil War, Dir. Toby Haynes, USA, 2019.
Her, Dir. Spike Jonze, USA, 2013.
Star Trek: The Next Generation (serie TV), Id. Gene Roddenberry, USA, 1987-1994.
Transcendence, Dir. Wally Pfister, USA, 2014.
Westworld, Dir. Michael Crichton, USA, 1973.
Westworld (serie TV), Id. Jonathan Nolan, Lisa Joy, USA, 2016 - in corso.

L'autrice

Isabella Tomassucci

Dottoranda Magistrale in Scienze della Comunicazione Pubblica, D'Impresa e Pubblicità. Laureata presso l'Università degli Studi di Macerata, i suoi interessi di studio e di ricerca si concentrano soprattutto sulle intersezioni tra narrazione e processi cognitivi.

E-mail: isabella.tomassucci@gmail.com

L'articolo

Data di invio: 15/02/2019

Data di accettazione: 15/04/2019

Data di pubblicazione: 30/05/2019

Come citare questo articolo

Tomassucci, Isabella, "Umanità artificiale: il rapporto uomo-macchina tra *fiction* e scienza", *Immaginare l'impossibile. trame della creatività tra letteratura e scienza*, Eds. L. Boi, F. D'Intino, G. V. Distefano, *Between*, vol. IX, n. 17 (Maggio 2019), <http://www.Between-journal.it/>