POLITECNICO DI TORINO
Repository ISTITUZIONALE

La teoria dei giochi per migliorare l'interazione tra robot ed esseri umani

Original
La teoria dei giochi per migliorare l'interazione tra robot ed esseri umani / Rizzo, A., Galati, G., Simone Macrì, E.. - In: NOTIZIARIO DELL'ISTITUTO SUPERIORE DI SANITÀ. - ISSN 1827-6296. - ELETTRONICO. - 36:9(2023), pp. 7-10.

Availability:
This version is available at: 11583/2983125 since: 2023-10-19T08:14:48Z

Publisher:
Istituto Superiore di Sanità

Published
DOI:

Terms of use:

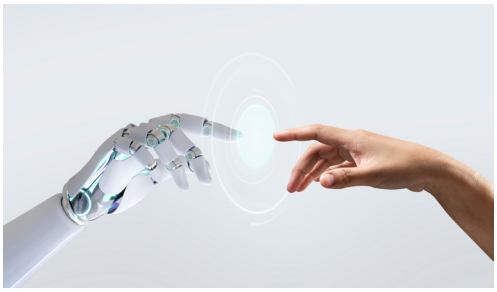
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

13 November 2026

LA TEORIA DEI GIOCHI PER MIGLIORARE L'INTERAZIONE TRA ROBOT ED ESSERI UMANI



Alessandro Rizzo[1], Giada Galati[1] e Simone Macrì[2]
[1]*Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino*
[2]*Centro di Riferimento per le Scienze Comportamentali e la Salute Mentale, ISS*

RIASSUNTO - In un mondo in cui robot e veicoli autonomi saranno sempre più presenti nelle nostre vite, è fondamentale sviluppare metodiche che, oltre a garantirne la sicurezza, ne promuovano l'accettabilità sociale. In questo studio multidisciplinare, all'intersezione tra ingegneria e psicologia, è stato sviluppato un algoritmo in grado di animare i robot, ispirandoci alle interazioni tra esseri umani (teoria dei giochi) e non a considerazioni meramente geometriche. Un ampio campione di volontari ha confermato come le traiettorie generate, a partire dalla teoria dei giochi, fossero maggiormente accettabili rispetto a quelle generate dagli algoritmi standard. Inoltre, l'algoritmo ottenuto è risultato superiore rispetto allo standard anche nell'ottimizzazione delle caratteristiche geometriche.
Parole chiave: robotica; teoria dei giochi; scienze sociali

SUMMARY (*Game theory to enhance human-robot interaction*) - In a world where robots and autonomous vehicles are pervasive, their interaction with humans becomes crucial and calls for novel techniques to enhance safety and social acceptability. We present a multidisciplinary research integrating engineering and psychology to endow autonomous robots with an innovative, human-centered motion planner that leveraged game theory. A large sample of volunteers deemed the robotic trajectories generated by our human-centered motion planner more socially acceptable than those generated by standard algorithms, which typically neglect humans in favor of the mere optimization of geometric performance indices. Surprisingly, we noted that pursuing solely a human-centered objective also improved geometric performance.
Key words: robotics; game theory; social sciences
simone.macri@iss.it

Nel film *I-Robot* (*Io, robot* 2004), ambientato nel 2035, il mondo è popolato sia da esseri umani sia da robot le cui azioni e sembianze ricalcano fedelmente le loro controparti umane. A rendere i robot il più possibile simili alle persone è Susan Calvin (l'attrice Kathryn Bridget Moynahan), una psicologa esperta di intelligenze artificiali il cui lavoro è quello di "migliorare l'antropomorfizzazione robotica" ovvero "di rendere i robot più umani". Nonostante sia stato distribuito nel 2004, questo film sembra riflettere abbastanza fedelmente quello che si potrebbe incontrare in un prossimo futuro, quando andranno condivisi gli spazi con robot sviluppati per svolgere molteplici compiti. Già oggi siamo circondati da robot a cui deleghiamo faccende più o meno complesse, dal pulire la casa, al fornirci indicazioni stradali, fino all'accoglierci in alberghi e ristoranti. Nell'ottica di una coabitazione sempre crescente, un aspetto che effettivamente sta interessando molti studiosi è proprio quello dell'antropomorfizzazione del movimento. Come deve muoversi un robot affinché sia sicuro (cioè non arrechi danno alle persone) e sia al tempo stesso accettato dagli esseri umani?

Principi matematici alla base del movimento dei robot

Da decenni gli esperti di robotica studiano tecniche di pianificazione delle traiettorie: tra le infinite possibilità per muoversi da un punto iniziale a un punto finale è infatti necessario selezionarne una che rispetti alcuni criteri fondamentali, quali l'ottimizzazione della distanza percorsa, l'elusione degli ostacoli fissi e mobili e una certa regolarità nel compiere curve e deviazioni. ▶

Accanto a questi criteri, facilmente traducibili usando funzioni matematiche, resta aperto il problema della convivenza con gli esseri umani. In tale contesto, i movimenti generati dai robot devono poter essere accettati dagli umani come familiari - un aspetto che spesso viene trascurato e che non ha un'immediata corrispondenza con formulazioni matematiche. Nello specifico, mentre i predetti criteri di ottimizzazione possono essere ottenuti tramite la soluzione di problemi geometrici (ancorché intricati), la componente che riguarda "l'interazione umana" deve essere integrata tramite modelli più complessi. In quest'ambito, la teoria dei giochi - una teoria sviluppata a partire dagli anni '50 del secolo scorso che permette di riprodurre fenomeni di interazione strategica tra agenti razionali - può essere usata con successo per promuovere l'antropomorfizzazione dei robot, emulando quei processi di negoziazione implicita che gli esseri umani mettono in atto quando si trovano a interagire condividendo uno spazio. In parole semplici, la riproduzione dell'interazione umana che utilizza la teoria dei giochi si basa su un sistema di ricompense e penalità ottenute da agenti virtuali ogni qual volta compiano delle scelte (1).

Ogni agente tende ad accumulare la maggior ricompensa e a evitare le penalità (ottenendo quindi il miglior vantaggio individuale). Questo processo viene svolto da tutti gli agenti contemporaneamente, fino a trovarsi in una condizione di equilibrio. Il più famoso di questi equilibri è quello di Nash, per cui ogni ulteriore scelta risulta peggiorativa per almeno uno degli agenti.

Psicologia e teoria dei giochi per "umanizzare" il movimento dei robot

La premessa della ricerca (1) è che le persone che si muovono in uno spazio negoziano i loro movimenti con uno schema simile a quello della teoria dei giochi, e trovano compromessi nella loro interazione quando si trovano in un equilibrio di Nash. L'ipotesi generata da questa premessa è che un robot che interagisce con gli umani, secondo uno schema basato sulla teoria dei giochi e la ricerca di un equilibrio di Nash, possa essere maggiormente accettato dalle persone con cui condivide lo spazio.

Per programmare il robot a muoversi secondo quest'ultimo schema si è proceduto in due fasi. Dapprima sono stati analizzati una serie di filmati di sorveglianza e si è verificato che la teoria dei giochi fosse efficace nel riprodurre le interazioni degli esseri umani, ricavando i necessari parametri del modello (Figura A). Questi parametri sono stati usati per programmare un robot mobile. Infine, questo robot è stato usato per interagire con gli esseri umani (Figura B).

Per saggiare l'ipotesi che il modello di movimento generato tramite la teoria dei giochi fosse maggiormente "accettabile e gradito" rispetto ai modelli basati su considerazioni geometriche (l'attuale standard della navigazione robotica), è stato chiesto a 691 adulti (afferenti a vario titolo al Politecnico di Torino quali studentesse e studenti, personale di ricerca e tecnico, e corpo docente) di partecipare volontariamente alla ricerca.

Dopo aver raccolto i dati anagrafici e le informazioni concernenti il livello di scolarizzazione e il grado di familiarità con la robotica e l'utilizzo del computer, è stato chiesto ai partecipanti di trascorrere una mezz'ora del loro tempo a guardare filmati e rispondere a diverse domande. I filmati erano composti da frecce che si muovevano sullo schermo (Figura C) seguendo traiettorie con caratteristiche precise. Nello specifico, al fine di valutare esclusivamente il movimento, depurandolo da altri fattori potenzialmente confondenti, si è partiti da filmati di persone che si muovevano in un ambiente urbano e di seguito le persone sono state sostituite con delle frecce (Figura C); successivamente sono stati rimossi digitalmente indizi visivi quali automobili e palazzi (Figura C, pannello di destra). Pertanto, ogni filmato utilizzato era caratterizzato da frecce che percorrevano traiettorie "umane" a cui era aggiunta un'ulteriore freccia. Quest'ultima poteva essere costituita da una tra le seguenti tre opzioni:

- una freccia che percorreva una traiettoria "umana" (la situazione di controllo);
- una freccia che si muoveva secondo l'algoritmo standard basato su considerazioni geometriche;
- una freccia che si muoveva secondo l'algoritmo sviluppato in base alla teoria dei giochi.



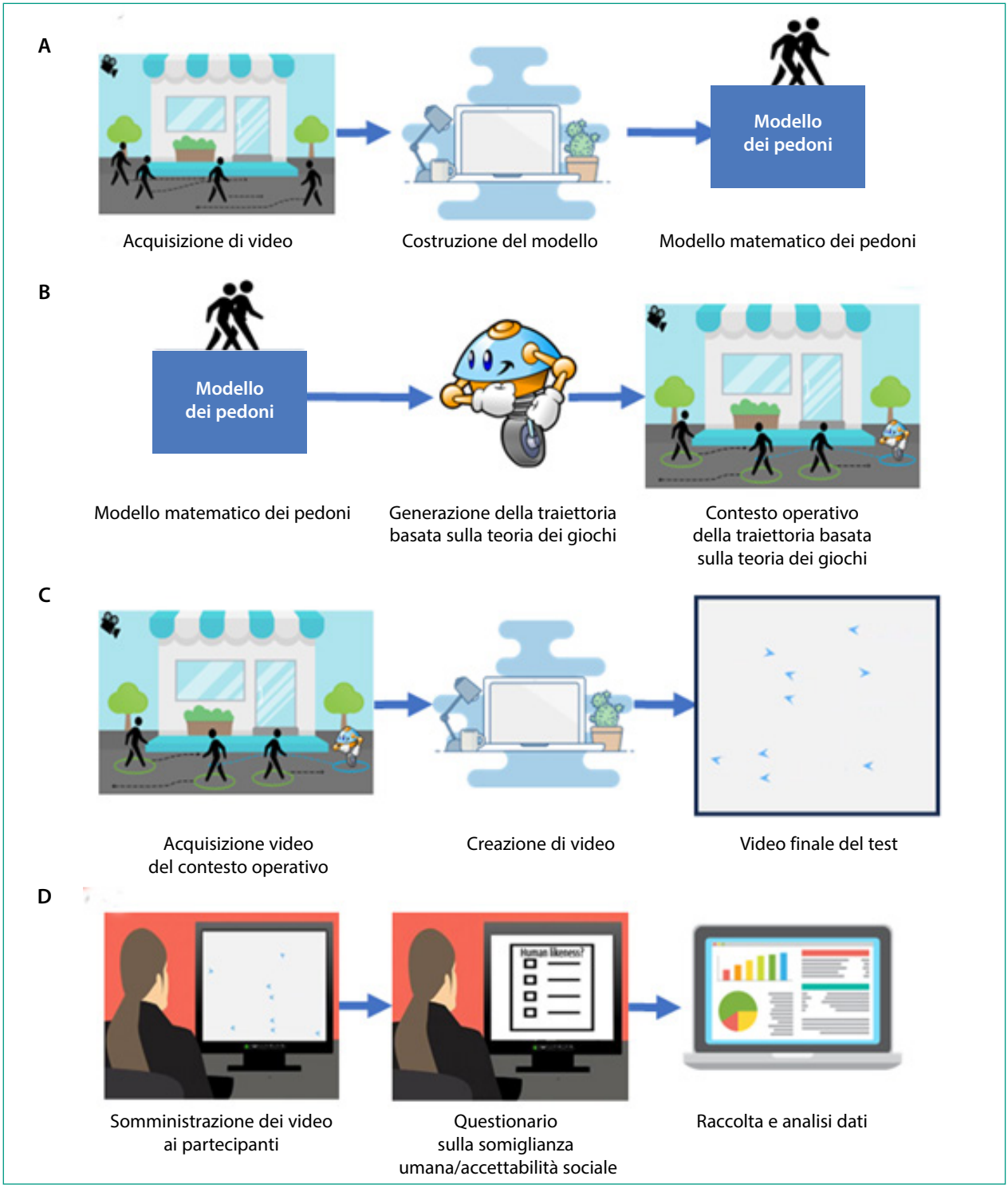
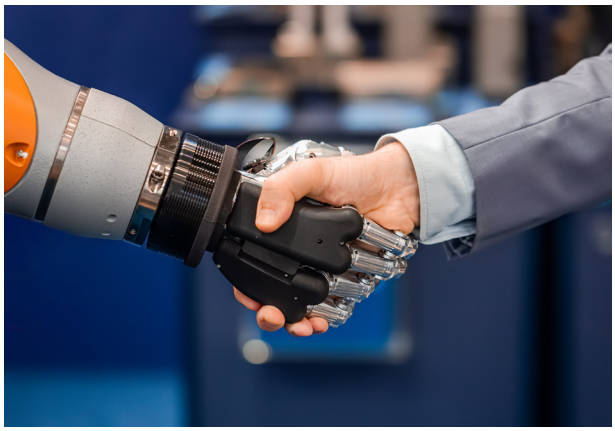


Figura A, B, C, D - Schematizzazione della procedura di sperimentazione e validazione (2) (*Fonte*: Pixabay)

Ognuno dei 691 partecipanti ha visto 21 filmati (7 per ogni tipo di condizione, presentati in ordine casuale) e ha risposto a una serie di domande (Figura D); preliminarmente, è stato chiesto ai partecipanti di valutare se, all'interno del filmato presentato, ci fosse una freccia che si muoveva in modo "strano" (in caso di risposta affermativa era chiesto ai partecipanti di indicare quale fosse la ▶



freccia che si muoveva in modo "strano" e di indicare il grado di "stranezza" secondo una scala a cinque punti). Infine, in un'ulteriore sessione di lavoro, è stata evidenziata la freccia di riferimento tramite un circoletto rosso e abbiamo chiesto ai partecipanti di valutarne il grado di naturalezza, stranezza e somiglianza del movimento rispetto a quello di un essere umano.

Coerentemente con le ipotesi di lavoro, in riferimento al grado di "stranezza" dei filmati, la maggior parte dei partecipanti ha indicato come più strani i filmati in cui era presente una freccia animata dall'algoritmo standard. Il grado di stranezza dei filmati composti esclusivamente da frecce "umane" era identico rispetto al grado di stranezza dei filmati in cui era presente una freccia animata dalla teoria dei giochi. Inoltre, quando è stato chiesto alle volontarie/i di indicare quale fosse secondo loro la freccia "strana", quest'ultima è stata identificata correttamente nel 77% dei casi nei filmati comprendenti la freccia animata dall'algoritmo standard, e solo nel 48% dei casi nei filmati comprendenti la freccia animata dalla teoria dei giochi.

Quando, nell'ultima fase del test, è stato chiesto ai partecipanti di focalizzarsi sulla freccia "robotica", sono stati ottenuti dei risultati che suffragano ulteriormente l'ipotesi di lavoro. In particolare, sebbene le frecce umane avessero un grado di naturalezza maggiore rispetto a quelle robotiche, queste ultime erano percepite come significativamente differenti in funzione dell'algoritmo che le animava. Coerentemente con le ipotesi del lavoro effettuato, le frecce animate tramite la teoria dei giochi erano percepite come più naturali rispetto a quelle animate da considerazioni geometriche.

Infine, sebbene questo non fosse lo scopo principale della ricerca presa in esame, si è avuta l'opportunità di confrontare le traiettorie stesse, indipendentemente dal grado di accettabilità, in termini di indicatori classici di prestazione quali: ottimizzazione della distanza percorsa tra il punto di partenza e il punto di arrivo; numero di curve e cambi di direzione; distanza dall'essere umano più vicino. Sebbene basato unicamente su modelli di interazione e non su modelli geometrici, l'algoritmo ottenuto è risultato migliore rispetto all'algoritmo standard anche in merito a questi parametri.

In conclusione, i risultati della ricerca suggeriscono che mettere l'essere umano e le sue aspettative al centro delle considerazioni concernenti lo sviluppo dei robot possa migliorare non solo le sempre crescenti interazioni, ma anche le prestazioni dei robot stessi.

Dichiarazione sui conflitti di interesse

Gli autori dichiarano che non esiste alcun potenziale conflitto di interesse o alcuna relazione di natura finanziaria o personale con persone o con organizzazioni, che possano influenzare in modo inappropriato lo svolgimento e i risultati di questo lavoro.

Riferimenti bibliografici

1. Nash J. Non-cooperative games. *Annals of Mathematics* 1951;54(2):286-95 (doi.org/10.2307/1969529).
2. Galati G, Primatesta S, Grammatico S, *et al*. Game theoretical trajectory planning enhances social acceptability of robots by humans. *Sci Rep* 2022;12, 21976 (https://doi.org/10.1038/s41598-022-25438-1).

TAKE HOME MESSAGES

- La convivenza tra esseri umani e robot è un problema che la sola tecnologia non riuscirà a risolvere, è necessario coinvolgere le scienze umane come la psicologia.
- L'antropomorfizzazione delle macchine è un aspetto cruciale su cui fare leva, poiché aumenterà il grado di accettabilità dei robot da parte degli esseri umani.
- La teoria dei giochi, molto usata in economia, si è dimostrata una scelta valida per migliorare il livello di antropomorfizzazione delle macchine.