

QuaderniCIRD



n. 27 (2023)

Numero ordinario

ISSN: 2039-8646

Homepage: <<https://www.openstarts.units.it/dspace/handle/10077/3845>>

QuaderniCIRD

Rivista del Centro Interdipartimentale
per la Ricerca Didattica dell'Università di Trieste

*Journal of the Interdepartmental Center
for Educational Research of the University of Trieste*

n. 27 (2023)

Direttore responsabile

Michele Stoppa, Dipartimento di Matematica, Informatica e Geoscienze

Condirettore

Luciana Zuccheri, Dipartimento di Matematica, Informatica e Geoscienze

Comitato editoriale

Silvia Battistella, Dipartimento di Scienze della Vita

Furio Finocchiaro, Dipartimento di Matematica, Informatica e Geoscienze

Helena Lozano Miralles, Dipartimento di Studi Umanistici

Tiziana Piras, Dipartimento di Studi Umanistici

Monica Randaccio, Dipartimento di Scienze Giuridiche, del Linguaggio, dell'Interpretazione e della Traduzione

Paolo Sorzio, Dipartimento di Studi Umanistici

© copyright Edizioni Università di Trieste, Trieste 2023.

Proprietà letteraria riservata.

I diritti di traduzione, memorizzazione elettronica, di riproduzione e di adattamento totale e parziale di questa pubblicazione, con qualsiasi mezzo (compresi i microfilm, le fotocopie e altro) sono riservati per tutti i paesi.

EUT - Edizioni Università di Trieste
Via E. Weiss, 21 - 34128 Trieste

[HTTP://EUT.UNITS.IT](http://EUT.UNITS.IT)



QuaderniCIRD

n. 27 (2023)

Sommario

4 Michele Stoppa
Presentazione

Prima Parte

9 Franco C. Grossi
Ergonomia Medica: Intelligenza Artificiale e Gemelli Digitali in Medicina

Dossier “Matematica e ...”

40 Luciana Zuccheri
Introduzione al Dossier “Matematica e ...”

44 Andrea Sgarro, Laura Franzoi
Distanze, lingue e parole

55 Elisabetta Podda, Alessandro Logar
Sophie Germain e l’ultimo Teorema di Fermat

80 Agnese Valentini
Come performa ChatGPT in matematica? Prime sfide su alcuni problemi elementari di Analisi 1

101 Fabio Pasticci
La crittografia classica come occasione di ragionamento matematico

Seconda Parte

Note e aggiornamenti critici

122 Michele Stoppa
Note e aggiornamenti critici. ALLEGRI R., GARBARINO O., 2018, Vernazza. Le trasformazioni nel tempo dell’insediamento e del territorio, Novara, Libreria Geografica, 174 pp.

Questo numero della rivista è stato curato da:

Michele Stoppa, Luciana Zuccheri.

Revisione dei sunti in Spagnolo: Helena Lozano Miralles.

Presentazione*

1. IL FOCUS

Con la pubblicazione, nel 2021, del *Programma Strategico per l'Intelligenza Artificiale IA 2022-2024*¹, a cura del Ministero dell'Università e della Ricerca, del Ministero dello Sviluppo Economico e del Ministro per l'Innovazione tecnologica e la Transizione Digitale, il Governo Italiano ha inteso promuovere lo sviluppo di un mirabolante “ecosistema nazionale di ricerca” d'eccellenza, da armonizzare con un più ampio sistema europeo, precipuamente dedicato all'*Intelligenza Artificiale* e alle sue molteplici applicazioni nei campi più disparati, con l'intento di favorire l'innovazione del tessuto economico e sociale del Paese.

L'ambizioso entusiastico documento, che delinea *obiettivi, settori prioritari e aree di intervento*, è certamente destinato a far discutere e sta indubbiamente stimolando l'interesse e un vivace dibattito nella comunità scientifica e nel variegato universo degli stakeholders. Anche la nostra rivista ha inteso raccogliere la sfida e offre, pertanto, in questo numero, due contributi dedicati ad altrettanti settori prioritari di applicazione dell'Intelligenza Artificiale, ossia l'ambito *salute e benessere*, con un corposo contributo pubblicato nella prima parte, e l'ambito *sistema educativo*, con un ulteriore contributo che compare, invece, nel Dossier disciplinare “Matematica e ...”. Come spesso avviene, lo sviluppo – in questo caso tumultuoso – delle conoscenze scientifiche e delle nuove tecnologie pone inevitabilmente preoccupazioni di natura etica (e non solo), pertanto il loro utilizzo deve *sempre* essere sottoposto a scrutinio critico nonché alla prudente responsabilità etica dell'uomo ultimamente ispirata a *principi non negoziabili*, potendo l'introduzione pervasiva di tali innovazioni offrire, di volta in volta, certamente importanti benefici alla comunità umana ma anche, se non

* Title: Presentation.

¹ Cfr. <<https://assets.innovazione.gov.it/1637937177-programma-strategico-iaweb-2.pdf>>.

adeguatamente governate, scoperchiare il mitico “vaso di Pandora”, determinando il profilarsi di scenari distopici.

2. LA PRIMA PARTE

Nella prima parte del numero, **Franco C. Grossi** riprende il tema dell'*ergonomia medica* – già introdotto nel n. 26 (2023) della rivista, a cui si rinviano i lettori interessati – e illustra ora, in un ampio contributo, le lusinghiere opportunità che l'intelligenza artificiale offre già oggi alla pratica medica (e un domani potrà ulteriormente offrire con il suo sviluppo e perfezionamento...) sia sul piano diagnostico che sul piano terapeutico e, sinanco, nel caso dell'intelligenza artificiale generativa, sul piano previsionale.

Dopo un'utile ricognizione degli “strumenti a disposizione”, l'autore esamina in modo sistematico il supporto offerto dall'intelligenza artificiale alla diagnostica (test di laboratorio, metodiche di immagine e test funzionali, soffermandosi in particolare sugli elettrocardiogrammi, sulla risonanza magnetica e sulla tomografia a emissione di protoni) nonché in ambito chirurgico (robot chirurgici e gemelli digitali) consentendo di perseguire una medicina di precisione.

2. IL DOSSIER DISCIPLINARE “MATEMATICA E ...”

Il numero propone, quindi, un *Dossier disciplinare* di carattere monografico dedicato a quattro originali temi trasversali di indubbio interesse per la matematica, come si precisa nell'*Introduzione al Dossier* a firma della curatrice **Luciana Zuccheri**, a cui si rinvia per ulteriori precisazioni e approfondimenti.

Il contributo di **Andrea Sgarro** e **Laura Franzoi** esplora il concetto di *distanza*, tra l'altro divenuto oggetto di un'originale monografia di respiro interdisciplinare, l'*Encyclopedia of Distances*. Gli autori si soffermano innanzitutto sulla distanza metrica in matematica e sugli assiomi che la caratterizzano, per passare quindi in rassegna alcuni esempi di distanze non metriche (distanza di Muljačić, distanza di Levenštejn e distanze di Longobardi). Si tratta di distanze di interesse per la linguistica che consentono di studiare le affinità linguistiche tra le diverse lingue.

Elisabetta Podda e **Alessandro Logar** propongono un contributo dedicato a Sophie Germain, una figura di indubbio interesse per la storia della matematica. Dopo un opportuno gustoso inquadramento biografico della pervicace studiosa francese nel quadro del vivace panorama culturale e scientifico del suo tempo e, dopo aver delineato il suo contributo originale a interessanti sviluppi nel campo della teoria dei numeri e della fisica matematica, gli autori si soffermano sui risultati conseguiti dalla studiosa nonché sulle tecniche dimostrative utilizzate dalla Germain nell'affrontare con acribia la congettura di Fermat.

Il contributo di **Agnese Valentini** affronta la questione dell'utilizzo dell'Intelligenza Artificiale a supporto dell'insegnamento universitario dell'Analisi matematica. Il contributo, dopo una sintetica introduzione all'architettura di ChatGPT, presenta una serie di esperimenti cui è stato sottoposto il software, a cui sono stati rivolti quesiti sia di natura computazionale (studi di funzione nonché calcolo di limiti e integrali) sia di tipo dimostrativo tratti dalle prove somministrate agli esami di profitto degli insegnamenti di Analisi Matematica 1 erogati per le Lauree triennali in Fisica e in Matematica dell'Università di Trieste.

Dall'analisi dei risultati emergono anche significativi punti di debolezza di un utilizzo "disinvolto" dell'intelligenza artificiale nella didattica della matematica: il contributo offre alcuni utili suggerimenti didattici per minimizzare gli errori e valorizzare le potenzialità dello strumento che, almeno allo stato dell'arte, richiede un attento intervento di supporto formativo da parte della docenza.

Il contributo di **Fabio Pasticci** è dedicato alla crittografia, presentata come originale opportunità per la didattica della matematica da attuare nella Scuola primaria, volta a far percepire agli alunni l'utilità pratica della disciplina nella risoluzione di problemi concreti. L'autore si sofferma sul crittosistema di Cesare e, in particolare, indugia sulle tecniche di crittoanalisi, offrendo esemplificazioni di simpatiche esercitazioni volte a suscitare la curiosità e l'interesse degli alunni.

3. LA SECONDA PARTE

Nell'ambito della rubrica *Note e aggiornamenti critici*, l'ultimo contributo del numero, a firma dello scrivente, tratteggia un'esemplificazione di come si possa procedere con efficacia allo *studio interdisciplinare* di un territorio paradigmatico.

Ciò è reso possibile grazie alla valorizzazione delle pregevoli suggestioni offerte dalla monografia di R. Allegri e O. Garbarino che, inserendosi a pieno titolo nella consolidata tradizione degli *studi regionali*, consente di delineare un esaustivo quadro sui processi di umanizzazione che si sono susseguiti, interessando nel corso delle successive fasi storiche un peculiare lembo del Levante ligure.

MICHELE STOPPA
Direttore, "QuaderniCIRD"
Dipartimento MIGe
Università di Trieste
mstoppa@units.it

Prima parte

*Ergonomia Medica: Intelligenza Artificiale e Gemelli Digitali in Medicina**

FRANCO C. GROSSI**

Universidad Nacional del Nordeste
Corrientes (República Argentina)
grossi@med.unne.edu.ar

ABSTRACT

The doctor listens to signals coming from different parts of the patient's body and identifies signs, symptoms and other connotative clinical-anamnestic elements, necessary for a personalized diagnosis. Today his senses are extended through ICT allowing diagnoses, prognoses, and therapies to be aided by the use of artificial intelligence. For instance, through imaging diagnostic, he is able to visualize all the organs of the human body. The interpretation of data from these highly complex devices can be supported by generative artificial intelligence, which compares the gathered data with that from millions of medical records, even predicting the evolution or onset of pathologies. Using digital twins, surgeon can simulate three-dimensional scenarios with augmented reality before commencing surgical procedures, thereby increasing the likelihood of success. The contribution of Ergonomics is fundamental in all of this, as it studies the interactions between healthcare workers, patients, medical equipment and the hospital environment, proposing adequate anthropocentric interfaces.

PAROLE CHIAVE

ERGONOMIA MEDICA / MEDICAL ERGONOMICS / ERGONOMÍA MÉDICA; INTELLIGENZA ARTIFICIALE IN MEDICINA / ARTIFICIAL INTELLIGENCE IN MEDICINE / INTELIGENCIA ARTIFICIAL EN MEDICINA; REFERTAZIONE MEDICA CON L'INTELLIGENZA ARTIFICIALE GENERATIVA / MEDICAL REPORTING WITH GENERATIVE ARTIFICIAL INTELLIGENCE / INFORME MÉDICO CON INTELIGENCIA ARTIFICIAL GENERATIVA; GEMELLI DIGITALI IN CHIRURGIA / DIGITAL TWINS IN SURGERY / GEMELOS DIGITALES EN CIRUGÍA.

1. PRAFAZIONE

Le Tecnologie per l'Informazione e la Comunicazione (Acronimo: ICT - Information and Communication Technologies) sono in grado di estendere i nostri sensi all'infinito, dotandoli

* Title: *Medical Ergonomics: Artificial Intelligence and Digital Twins in Medicine* / Titulo: *Ergonomía Médica: Inteligencia Artificial y Gemelos Digitales en Medicina.*

** Emeritus Professor of Ergonomics and academic delegate of the National University of the Northeast (Argentina) for relations with Italian universities. He taught *Information & Communication Technology* and *Applied Ergonomics* at the University of Trieste, 1997 to 2013.

altresì di un'enorme memoria e di incredibili velocità di elaborazione e trasmissive¹. Ad esempio, in medicina vengono utilizzati gli ultrasuoni sia in *diagnosi*² sia in *terapia*³, i campi elettromagnetici nella *diagnostica per immagini*, le *radiazioni ottiche coerenti* nella chirurgia laser e moltissime altre tecnologie consentono al medico di praticare la sua professione con maggiore precisione e sicurezza.

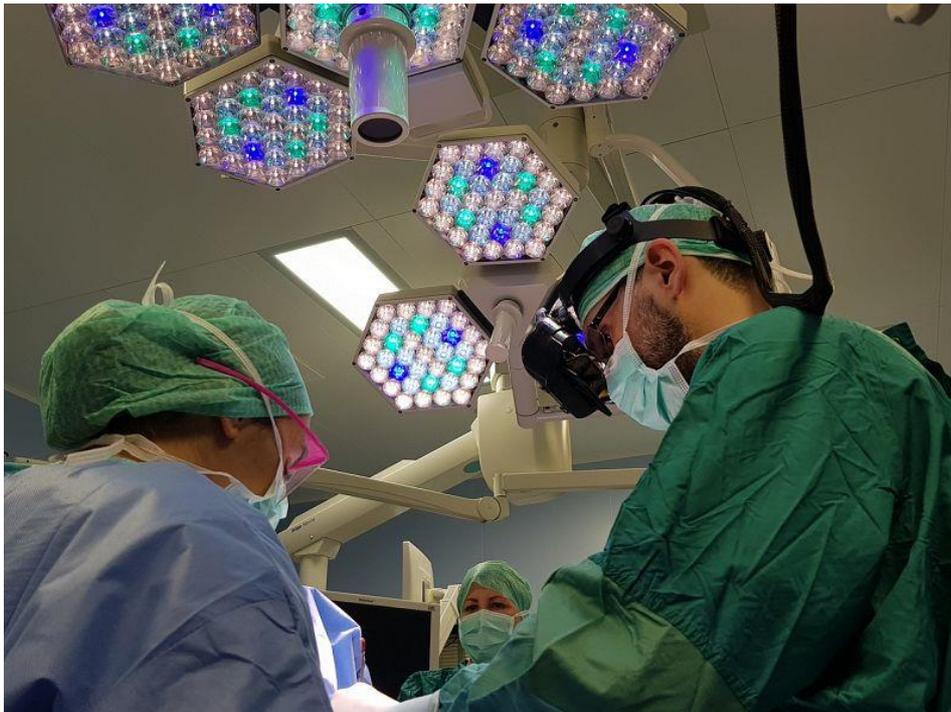


Figura 1. Realtà Aumentata utilizzata in chirurgia⁴.

Tutte le limitazioni dei suoi organi di senso vengono superate e potenziate e, tramite le applicazioni dell'Ergonomia, si riescono a realizzare adatte interfacce anche mediante l'impiego dell'*Intelligenza Artificiale*, che provvede a identificare tutte le informazioni necessarie per formulare adeguati percorsi terapeutici e di cura. L'utilizzo dell'*Intelligenza Artificiale Generativa*, poi, consente al medico non solo di diagnosticare in maniera efficace determinate patologie, ma anche di prevedere il loro insorgere nel tempo. L'uso dei *Gemelli Digitali* in chirurgia permette, tramite la *Realtà Aumentata* e il *Metaverso*, di

¹ Cfr. GROSSI 1999.

² Ad esempio nell'*ecografia*.

³ Ad esempio nella *litotrissia* extracorporea per il trattamento della calcolosi urinaria.

⁴ Fonte: <<https://www.quinewspisa.it/bologna-chirurgia-operazione-guidata-da-realta-aumentata.htm>>.

sperimentare gli effetti di interventi operatori prima ancora di eseguirli e di sperimentare nuove soluzioni di ricerca addirittura sino a livello cellulare. Dopo una sommaria descrizione delle principali tecnologie avanzate in campo medico, verranno rappresentate alcune applicazioni sul campo.

2. ERGONOMIA E COMUNICAZIONE MEDICA

Come abbiamo visto, la *comunicazione*, intesa nella sua accezione più ampia, quale condivisione di conoscenze in un'interazione simbolica codificata, assume oggi un nuovo valore, grazie ai *new media* e alla realtà virtuale, che consentono l'espansione dei nostri sensi estesi nel *villaggio globale mcluhaniano*⁵.

Quando, nella metà degli anni Novanta, partecipammo alla fondazione del Corso di Laurea in Scienze della Comunicazione, presso l'Università di Trieste⁶, eravamo consapevoli del ruolo fondamentale e determinante che questa disciplina avrebbe avuto negli anni a venire, sia per quanto riguarda le interrelazioni uomo-uomo, sia per quelle uomo-macchina-ambiente. E l'*Ergonomia*, proprio per la sua specificità di approccio multidisciplinare antropocentrico, in questo caso applicata alla Comunicazione, avrebbe dotato gli studiosi di importantissimi strumenti per sviluppare inattese e innovative conoscenze in questo ambito di ricerca.



Figura 2. L'importanza dell'Empatia in Sanità⁷.

⁵ Cfr. MCLUHAN, POWERS 1992.

⁶ È stato il più importante in Italia con più di 7.000 studenti iscritti.

⁷ Fonte: <<https://blog.medicalgps.com/the-importance-of-empathy-in-healthcare/>>.

Ad esempio, in campo medico, lo studio ergonomico dell'*empatia*⁸ tra i pazienti e gli operatori sanitari è diventato essenziale per un proficuo svolgersi delle interazioni verbali e non verbali, così come quello atto a risolvere i problemi di utilizzo e di refertazione dei dati provenienti dalle apparecchiature medicali per la diagnostica, che divengono ogni giorno più performanti e difficili da utilizzare.

Ma vediamo di illustrare i principali modelli che descrivono il processo comunicativo. Uno dei primi modelli venne ideato, nel 1948, da Claude Shannon⁹ e da Warren Weaver¹⁰, i quali indagarono anche sui problemi della comunicazione, che distinsero in *tecnici*¹¹, *semantici*¹² e *di efficacia*. In questo modello il messaggio del *mittente* viene codificato al fine di essere inviato tramite un mezzo trasmissivo e, una volta giunto a destinazione, decodificato per essere compreso dal *ricevente*. Si tratta di un modello che descrive il processo comunicativo in modalità unidirezionale. Un ulteriore passo avanti venne effettuato da Paul Watzlawick¹³, il quale introdusse il concetto di *feedback*¹⁴, fondamentale per conoscere se il ricevente ha compreso, o meno, il messaggio inviatogli dal mittente.

In Ergonomia si studia il feedback in maniera aprioristica, in modo da anticipare il gradimento e l'accondiscendenza da parte dell'interlocutore umano. Ad esempio l'operatore sanitario si avvarrà della propria esperienza per valutare l'interlocutore prima dell'approccio comunicazionale, mediante il possibile utilizzo dei dettami della

⁸ S'intende per *empatia* la capacità di porsi nella situazione di un'altra persona o, più esattamente, di comprendere immediatamente i processi psichici dell'altro.

⁹ Cfr. SHANNON 1948.

¹⁰ Il *modello Shannon-Weaver* suddivide il *processo comunicativo* in cinque parti: la sorgente che invia il messaggio, un trasduttore che lo codifica e lo inserisce, quale segnale, in un mezzo trasmissivo, un decodificatore che provvede a tradurre il segnale in un messaggio comprensibile per il destinatario. Il modello prevede anche la presenza di un eventuale rumore di fondo che permea il canale trasmissivo.

¹¹ La loro attività di ricerca si concentrò soprattutto sul livello tecnico, riguardante il problema di come utilizzare un segnale per riprodurre accuratamente un messaggio da un luogo a un altro. La difficoltà a questo proposito è che il rumore indotto può distorcere il segnale. Proposero una soluzione per superare tale problema, l'utilizzo della *ridondanza*, infatti se il messaggio originale è ridondante allora si possono individuare le distorsioni, il che rende possibile ricostruire il messaggio originario.

¹² La *semantica* è quella sezione della linguistica che studia il significato delle parole, degli insiemi delle singole lettere, delle frasi e dei testi.

¹³ Paul Watzlawick formulò anche i *cinque assiomi della comunicazione*: non si può non comunicare, ogni comunicazione ha un contenuto, la comunicazione dispone di elementi di punteggiatura, la comunicazione implica modalità digitali e analogiche, la comunicazione può essere simmetrica o complementare.

¹⁴ Cfr. WATZLAWICK, BAVELAS, JACKSON 2014.

*cinesica*¹⁵, della *prosemica*¹⁶, della *somatotipia*¹⁷, della *fisiognomica*¹⁸ e della *mimica facciale*¹⁹. In questo modo egli sarà in grado di esprimersi entrando in condizioni di empatia²⁰ con il proprio interlocutore. Questo per quanto riguarda la comunicazione umana²¹.

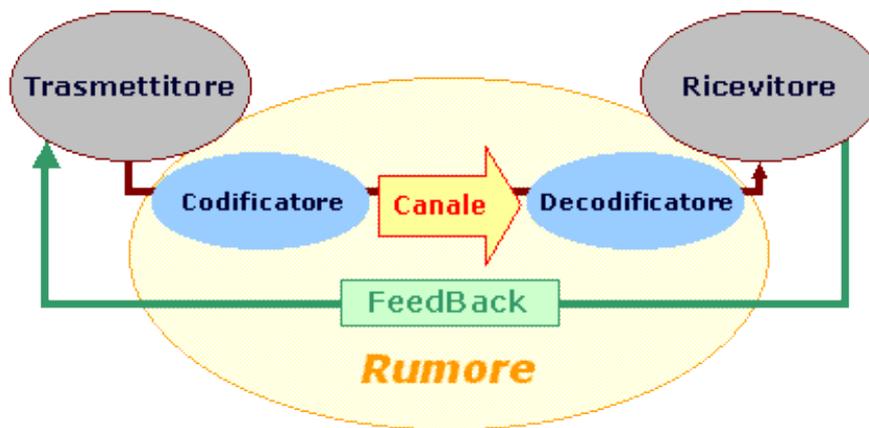


Figura 3. Il feedback applicato da Watzlawick al modello di Shannon²².

Un altro aspetto riguarda, invece, la comunicazione/interazione dell'uomo con il sistema apparecchiature mediche e ambiente ospedaliero. Ci viene allora incontro l'*Ergonomia Medica*²³, che prevede un'interfaccia di tipo *antropocentrico*. Ed anche in questo caso è stato studiato un *modello di comunicazione*²⁴ applicabile in Ergonomia e denominato AUAF²⁵ (acronimo: *Active Users in Advance Feedback*), il quale ipotizza il recepimento di feedback aprioristici, prima di attuare l'interrelazione.

¹⁵ La *cinesica* è la disciplina che studia il linguaggio del corpo.

¹⁶ La *prosemica* è la disciplina semiologica che studia i gesti, il comportamento, lo spazio e le distanze all'interno di una comunicazione, sia verbale sia non verbale.

¹⁷ I *somatotipi* rappresentano una *classificazione biotipologica* identificata attorno al 1940 dallo psicologo e medico statunitense William Herbert Sheldon. Negli anni Cinquanta e Sessanta altri studiosi ne perfezionarono e approfondirono i principi, creando dei metodi di misurazione antropometrica per stabilire l'appartenenza di un soggetto a un determinato somatotipo.

¹⁸ La *fisiognomica* è una disciplina pseudoscientifica cognitiva, che tenta di dedurre i caratteri psicologici e morali di una persona dai lineamenti e dalle espressioni del volto.

¹⁹ La *mimica facciale* è una parte della *cinesica* che riguarda il modo in cui si altera il volto delle persone. Gli esseri umani lasciano trasparire anche in questo modo il loro pensiero e le loro emozioni, in quanto la mimica facciale è difficile da controllare spontaneamente.

²⁰ L'*empatia* è la capacità di comprendere o sentire ciò che un'altra persona sta vivendo, cioè concretamente la capacità di "mettersi nei panni di un altro".

²¹ Cfr. GROSSI 2014a.

²² Fonte: <<https://sites.units.it/ergocom/docs/it/1.0.html>>.

²³ Cfr. GROSSI 2023a.

²⁴ Cfr. GROSSI 2000.

²⁵ Cfr. GROSSI 2014b, GROSSI in Siti web.

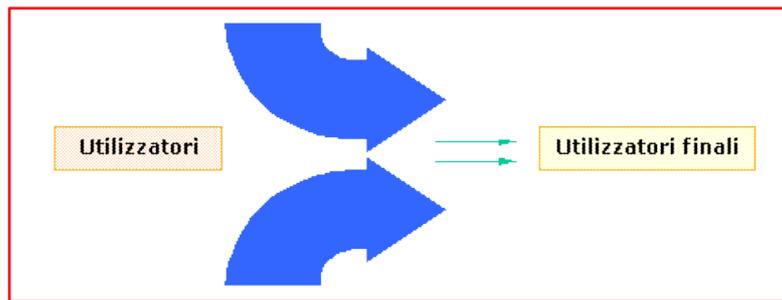


Figura 4. Il modello AUAF per la comunicazione aprioristica²⁶.

Questo modello è utilizzabile nelle interazioni uomo/macchina/ambiente, in quanto consente di anticipare le risposte del ricevente, al fine di produrre una comunicazione efficiente, efficace e di gradimento. Ad esempio nella comunicazione uomo/uomo l'operatore sanitario acquisisce dati sul paziente, sia dalla sua cartella sanitaria, sia dalle caratteristiche fisiche del suo interlocutore ed è in grado di interagire in maniera proattiva.

Nel caso di refertazioni mediante l'utilizzo di apparecchiature diagnostiche, il medico potrà comunicare dati anticipatori al paziente mediante, come vedremo, l'utilizzo dell'intelligenza artificiale, in grado di riconoscere anche, in modalità prodromica, l'eventuale comparsa di una *sintomatologia* caratteristica.

3. LA REALTÀ VIRTUALE E IL METAVERSO

Per *realtà virtuale*²⁷ s'intende la simulazione all'elaboratore di una situazione reale, con la quale il soggetto umano può interagire per mezzo di interfacce non convenzionali estremamente sofisticate, quali, ad esempio, occhiali e caschi ove viene rappresentata la scena virtuale e vengono riprodotti i suoni, nonché guanti dotati di sensori per simulare stimoli tattili e per tradurre i movimenti in istruzioni per il software²⁸.

Il fine della realtà virtuale è quello di simulare un ambiente reale per mezzo di tecnologie elettroniche, sino a fornire a chi la sperimenta l'impressione di trovarsi realmente immerso in quell'ambiente.

²⁶ Fonte: <<https://sites.units.it/ergocom/docs/it/1.0.html>>.

²⁷ Cfr. GROSSI 2023b.

²⁸ Cfr. CAROBENE in Siti web.

In particolare, la speculazione sulla realtà virtuale intrapresa nel Novecento²⁹, evocava mondi fantastici ed è stata oggetto di attenzione soprattutto da parte di scrittori, registi e artisti. All'alba del terzo millennio, il senso di questa tecnologia si consolida e ci porta a una posizione epistemologica più scientifica che nasce dalla considerazione di come la percezione dell'ambiente virtuale viene acquisita dalla mente attraverso l'interfaccia del "corpo esteso"³⁰.

Nell'ottica di questo passaggio esterno-interno, il soggetto utilizza artefatti tecnologici per aumentare le sue possibilità di azione nell'ambiente virtuale, combinando anche la percezione spaziale con l'azione psicomotoria in questo specifico contesto.

L'ambiente virtuale, divenuto trasparente al sistema cognitivo, viene percepito grazie a una continua "mappatura" che si fonde con l'interattività favorita dalle "affordances"³¹ degli artefatti tecnologici.



Figura 5. La Realtà Virtuale Semi-immersiva³².

Attualmente le tecnologie della *realtà virtuale* vengono suddivise in *semi-immersive*³³, *immersiva*³⁴, *aumentate*³⁵ e *collaborative*³⁶; in particolare, nella realtà virtuale immersiva

²⁹ Di *realtà virtuale* si inizia a parlare a cavallo tra gli anni Trenta e gli anni Quaranta del Novecento, quando lo scrittore Stanley Weinbaum pubblica il racconto breve *The Pygmalion's Spectacles*, in cui si fa esplicito riferimento a visori VR basati su registrazioni olografiche di esperienze in grado di stimolare anche il senso del tatto e dell'olfatto, oltre a quello della vista e dell'udito.

³⁰ Cfr. MCLUHAN 2001.

³¹ Con il termine *affordance* (trad.: invito all'uso) si definisce la qualità fisica di un oggetto che suggerisce a un essere umano le azioni appropriate per manipolarlo.

³² Fonte: <<https://www.digitalforbusiness.com/realtà-virtuale-medicina/>>.

³³ Realtà virtuale semi-immersiva (*Semi-Immersive Virtual Reality*).

³⁴ Realtà virtuale completamente immersiva (*Fully Immersive Virtual Reality*), che consente l'*embodiment*, ovvero "l'incarnazione" nell'ambiente virtuale.

³⁵ Realtà aumentata (AR, *Augmented Reality*).

³⁶ Realtà virtuale collaborativa (*Collaborative Virtual Reality*).

vengono utilizzate interfacce molto complesse, che permettono di percepire un “mondo” virtuale, dove la “presenza” spaziale dovrebbe coincidere il più possibile con quella fisica. Un *ambiente virtuale* è la simulazione di un *ambiente reale*, esplorabile tridimensionalmente in tempo reale, in cui l’utente è in grado di interagire con artefatti e altri utenti, nel rispetto del concetto di *embodiment*, che esprime la rappresentazione mentale dell’incarnazione nell’ambiente virtuale.



Figura 6. La Realtà Virtuale in sala operatoria³⁷.



Figura 7. Immagine a sinistra: realtà Virtuale Immersiva in sala operatoria³⁸. Immagine a destra: la *Realtà Aumentata* nella chirurgia spinale consente agli operatori di vedere la colonna vertebrale in 3D attraverso la pelle durante un intervento di impianto spinale mini-invasivo³⁹.

Ad esempio, in medicina, l’insegnamento tradizionale dell’anatomia prevede solitamente l’uso di un atlante anatomico, di un periodo di tempo trascorso in aula settoria e di tutta una serie di procedimenti accademici. Oggi, grazie alla realtà virtuale, è possibile offrire un migliore apprezzamento delle strutture del corpo umano (organi e apparati) nello spazio virtuale o reale e per facilitare la transizione dall’ambiente di apprendimento a quello clinico.

³⁷ Fonte: <<https://trends.medicaexpo.it/project-436674.html>>.

³⁸ Fonte: <<https://bcmj.org/blog/immersive-virtual-reality-medical-education>>.

³⁹ Fonte: <<https://health.uconn.edu/spine/services/minimally-invasive-spine-surgery/augmented-reality-spine-surgery/>>.

Lo stesso dicasi nello studio dell'istologia, ove normalmente viene utilizzato il microscopio per analizzare la struttura microscopica e ultramicroscopica dei tessuti e degli organi, dal punto di vista morfologico, istochimico e delle attività funzionali da essi esplicate. Anche in questo caso l'utilizzo della realtà virtuale consente di amplificare la visione delle componenti più piccole del tessuto muscolare o nervoso, sino alle cellule⁴⁰. Quindi, l'utilizzo della realtà virtuale in una combinazione tra citometria 3D tomografica e realtà virtuale potrà essere decisivo nei futuri scenari della diagnostica medica sul '*single cell imaging*'.

Per quanto riguarda la progettazione della realtà virtuale e dell'ambiente virtuale abbiamo a disposizione tutta una serie di tipologie progettuali che si riferiscono "primariamente" alla trasduzione sensoriale e alle sinestesie che avvengono tra i cinque sensi. Il percorso progettuale prosegue poi attraverso la costruzione di adeguati modelli visivi e cognitivi e, per una corretta impostazione del progetto in ambiente virtuale e della relativa interfaccia, vengono utilizzate alcune metodologie ergonomiche di tipo oggettivo e soggettivo. L'*usabilità* è un metodo oggettivo e permette di quantificare l'efficacia e l'efficienza rispetto alla soddisfazione dell'utente. La *gradevolezza*, infine, permette di individuare i desideri soggettivi degli utenti che possono solo essere qualificati e difficilmente quantificabili.

Il *Metaverso*⁴¹ è un'applicazione della realtà virtuale. Ha la caratteristica di estendere il mondo fisico utilizzando le tecnologie della realtà virtuale e della realtà aumentata, consentendo agli utenti di interagire senza problemi all'interno di ambienti simulati, utilizzando avatar e ologrammi.

⁴⁰ Il CNR e l'Università degli Studi di Napoli Federico II, hanno sviluppato un sistema di visualizzazione tridimensionale della cellula all'interno del metaverso, che hanno denominato *citometria tomografica 3D*, che consente altresì di evitare l'utilizzo di marcatori fluorescenti (label, tossici per le cellule stesse) e di restituire mappe tridimensionali di ciascuna cellula in flusso. Il metodo sviluppato, denominato *Generalized Computational Segmentation based on Statistical Inference (Generalized CSSI)*, permette di visualizzare e ottenere parametri quantitativi di una cellula partendo dall'immagine ottenuta attraverso il microscopio tomografico, ovvero un microscopio in grado di generare un'immagine 3D dei suoi organelli interni. Lo studio è stato pubblicato sulla rivista *Small Methods* (cfr. BIANCO *et al.* 2023).

⁴¹ Il termine *Metaverso*, neologismo composto da "meta-" (dal greco che significa dopo, oltre) e "universo", nasce nel romanzo *cyberpunk* di fantascienza di Neal Stephenson del 1992 *Snow Crash*, per indicare un mondo virtuale tridimensionale popolato con repliche umane digitali. Un archetipo del Metaverso è stata la piattaforma di mondi virtuali *Second Life*, nata nel 2003. Quindi non è un concetto nuovo nel mondo digitale quello di una realtà basata su un mondo virtuale.

Nel corso del prossimo decennio, il Metaverso rivoluzionerà quasi ogni aspetto della vita sociale e aziendale, della formazione e del mondo degli affari, consentendo la collaborazione in spazi virtuali, luoghi fisici aumentati e una combinazione di entrambi. Creerà inoltre nuovi processi comunicativi e trasformerà anche le interazioni tra pazienti, personale sanitario e realtà ospedaliera.



Figura 8. Il Metaverso e il suo uso in Medicina⁴².



Figura 9. L'utilizzo del Metaverso in Medicina⁴³.

Il Metaverso definisce, in sintesi, una zona di convergenza di spazi virtuali interattivi, localizzati nel cyberspazio e accessibili dagli utenti attraverso un avatar con funzione di rappresentante dell'identità individuale. Qui ognuno può creare ciò che vuole in 3D, anche la rappresentazione completa del corpo umano con i suoi organi e le proprie funzioni e può testare interventi e materiali prima che vengano realizzati.

⁴² Fonte: <<https://ie.binus.ac.id/2022/03/28/metaverse-will-the-future-of-health-industry-be-in-it/>>.

⁴³ Fonte: <<https://www.cognihab.com/blog/metaverse-in-the-healthcare-industry-how-to-make-the-most-out-of-it/>>.

4. L'INTELLIGENZA ARTIFICIALE TRADIZIONALE E QUELLA GENERATIVA⁴⁴

L'*intelligenza artificiale* (IA) è diventata un tema predominante in ogni settore della vita sociale, formativa e scientifica dell'ultimo decennio, anche perché ha prodotto notevoli progressi nella tecnologia e nell'efficienza operativa⁴⁵. È però necessario effettuare una distinzione tra la IA, denominata *tradizionale* o *debole* o *ristretta* e quella cosiddetta *generativa*.



Figura 10. Rappresentazione virtuale dell'Intelligenza Artificiale⁴⁶.

L'*Intelligenza Artificiale tradizionale* si concentra soprattutto nell'esecuzione intelligente di un compito specifico e si riferisce a sistemi progettati per rispondere a un particolare insieme di input. Questi sistemi hanno la capacità di apprendere dai dati e di prendere decisioni o fare previsioni basate su tali dati. Per esempio, analizzando le immagini fornite da un Risonanza Magnetica, l'operatore, coadiuvato dalla IA tradizionale potrà effettuare una refertazione molto più precisa rispetto a quella che avrebbe potuto diagnosticare mediante le proprie personali capacità ed esperienze. Con l'ausilio dell'*Intelligenza Artificiale di tipo generativo*, invece, sarà in grado di prevedere l'insorgere di patologie o disfunzioni. Quindi, l'Intelligenza Artificiale Generativa può essere considerata come la prossima evoluzione dell'intelligenza

⁴⁴ Cfr. MARR 2023.

⁴⁵ Due delle tecnologie basate sull'intelligenza artificiale più diffuse sono la guida autonoma delle automobili e gli avatar annunciatori su Internet.

⁴⁶ Fonte: <https://www.freepik.com/premium-photo/modern-artificial-intelligence-assistant-business-great-design-all-purposes_43984717.htm>.

artificiale, la quale viene “addestrata” su un enorme numero di dati al fine di apprendere gli schemi sottostanti e generare nuovi dati in linea con l’addestramento ricevuto. Consideriamo, ad esempio GPT-4⁴⁷, il modello di previsione del linguaggio di OpenAI⁴⁸, come un eccellente esempio di intelligenza artificiale generativa.



Figura 11. Rappresentazione virtuale dell’Intelligenza Artificiale Generativa in ambiente medico⁴⁹.

La principale differenza tra l’IA tradizionale e l’IA generativa risiede nelle loro capacità e applicazioni. I sistemi di intelligenza artificiale tradizionale vengono principalmente impiegati per analizzare dati ed effettuare refertazioni, mentre l’intelligenza artificiale generativa va oltre, creando nuovi dati per effettuare *previsioni*. In altre parole, l’intelligenza artificiale tradizionale eccelle nel riconoscimento dei *modelli*, mentre l’intelligenza artificiale generativa eccelle nella creazione di *nuovi modelli*. L’intelligenza artificiale tradizionale può analizzare i dati e dire cosa vede, ma l’intelligenza artificiale generativa può utilizzare gli stessi dati per creare qualcosa di completamente nuovo.

⁴⁷ GPT-4 – *Generative Pre-trainer Transformer 4*, è un modello linguistico multimodale di grandi dimensioni creato da OpenAI. È il quarto della serie di modelli di base GPT. È stato inizialmente rilasciato il 14 marzo 2023 ed è stato reso disponibile al pubblico tramite il prodotto chatbot a pagamento *ChatGPT Plus* e tramite l’API (Acronimo: *Application Programming Interface* – Interfaccia di Programmazione dell’Applicazione) di OpenAI.

⁴⁸ OpenAI è un’organizzazione di ricerca statunitense sull’intelligenza artificiale (AI) fondata nel dicembre 2015, la cui attività è dedicata alla ricerca nell’ambito dell’intelligenza artificiale con l’intenzione dichiarata di sviluppare una IA generale «sicura e vantaggiosa», che definisce come «sistemi altamente autonomi che superano gli esseri umani nel lavoro economicamente più prezioso».

⁴⁹ Fonte: <<https://www.linkedin.com/pulse/chatgpt-generative-ai-healthcare-revolutionizing-prabhu-stanislaus>>.

5. I BIG DATA E I DATA LAKE

L'Enciclopedia Treccani identifica i *Big Data* come «l'ingente insieme di dati digitali che possono essere rapidamente processati da banche dati centralizzate»⁵⁰. In altre parole si tratta dell'immensa quantità di dati che circolano su Internet in tempo reale, abbinati alle tecnologie per l'accumulo, l'analisi e l'elaborazione di dati strutturati o non strutturati⁵¹.



Figura 12. Rappresentazione virtuale dei *Big Data*⁵².

In particolare quelli che definiamo *Big Data* sono caratterizzati da tre fattori concomitanti, il loro *volume*, la loro *velocità di elaborazione*⁵³ e la loro *varietà di essenza*⁵⁴. Il volume rappresenta la quantità di dati e la loro potenza di ricezione, la velocità rappresenta la velocità di elaborazione e la frequenza degli aggiornamenti che

⁵⁰ Cfr. TRECCANI in Siti web.

⁵¹ I *dati strutturati* sono dati che sono stati predefiniti e formattati in una struttura impostata prima di essere inseriti nell'archivio dati, invece i *dati non strutturati* sono dati archiviati nel loro formato nativo e non elaborati fino a quando non vengono utilizzati.

⁵² Fonte: <<https://images.app.goo.gl/KdQFLXrmDknW4wWu5>>.

⁵³ Per archiviare, salvare, catalogare, riutilizzare tutti questi dati è necessario disporre di una grandissima velocità, altrimenti sarebbe inutile disporre di così tante informazioni che poi non possano essere immediatamente utilizzate. Inoltre qualunque flusso di dati, anche se di enormi dimensioni, deve essere studiato e analizzato in maniera molto rapida per rilevare la presenza o meno di nuovi codici maligni o di nuove minacce informatiche generate dagli hacker.

⁵⁴ Ogni dato, infatti, è molto diverso dall'altro, quindi le informazioni archiviate nei *Big Data* sono molto variegata e ognuna ha una provenienza specifica. Questa diversificazione crea dei problemi di gestione e unione dei vari dati per ottenere delle informazioni più complete. Sono necessari complessi strumenti di analisi dei dati per riuscire a capire in maniera chiara i *Big Data* e il modello di *Data Analytics* rappresenta il processo per ricavare valore da questa enorme mole di informazioni.

possono essere eseguiti, mentre la varietà rappresenta una molteplicità di dati non strutturati. I *Big Data*, trattati con l'Intelligenza Artificiale, aprono grandissime possibilità, infatti l'analisi dei dati raccolti con l'IA risulta essere di gran lunga più efficace dell'analisi umana. Il rapido sviluppo della ricerca e dello sviluppo dell'intelligenza artificiale negli ultimi anni è stato fortemente influenzato dalle scoperte tecnologiche dovute al così denominato “*apprendimento profondo*”.

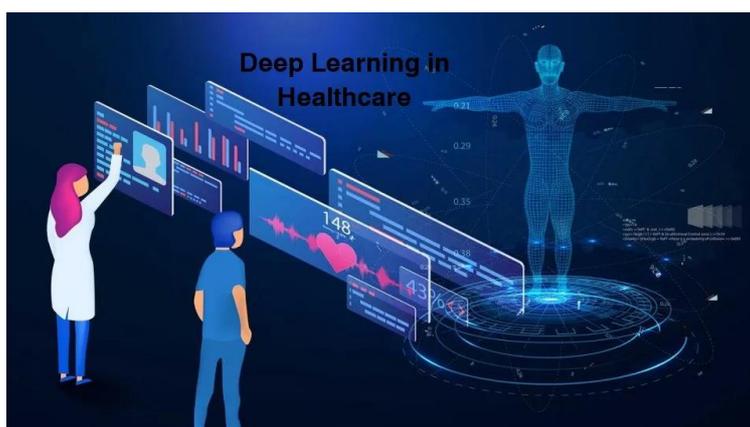


Figura 13. Rappresentazione schematica dell'Apprendimento Profondo⁵⁵.

L'apprendimento profondo è uno dei sistemi di apprendimento automatico mediante il quale le macchine estraggono e imparano automaticamente i dati necessari utilizzando algoritmi strutturati a più livelli.

In altre parole, per apprendimento profondo s'intende un insieme di tecniche basate su *reti neurali artificiali*⁵⁶ organizzate in diversi strati, dove ogni strato calcola i valori per quello successivo affinché l'informazione venga elaborata in maniera sempre più completa. L'apprendimento profondo richiede molti dati grezzi prima di estrarre i dati necessari e proprio i *Big Data* sono essenziali a questo processo.

Un'altra tecnologia in grado di far progredire il rapporto con i *Big Data* è l'*edge computing*, che è una tecnologia atta a separare ciò che può essere elaborato localmente, o sul campo, invece di inviare tutto al *cloud* e suddividere, quindi, l'elaborazione tra il

⁵⁵ Fonte: <<https://medium.com/@usama.6832/introduction-to-deep-learning-in-healthcare-519336d651b1>>.

⁵⁶ Nel campo dell'apprendimento automatico, una *rete neurale artificiale* è un modello computazionale composto da "neuroni" artificiali, ispirato a una rete neurale biologica.

cloud e il sito *edge*⁵⁷, garantendo un'elaborazione più veloce. In altri termini l'*edge computing* è un modello di calcolo distribuito nel quale l'elaborazione dei dati avviene il più vicino possibile a dove i dati vengono generati, migliorando i tempi di risposta e risparmiando sulla larghezza di banda⁵⁸.

L'elaborazione dei dati in prossimità del luogo in cui vengono generati reca a considerevoli vantaggi in termini di latenza⁵⁹ di elaborazione, riduzione di traffico dati e maggior resilienza in caso di interruzione nella connessione dati. In altri termini, invece di inviare tutti i dati direttamente al *cloud*, è possibile ottimizzarli e accumularli in *edge*.

Dopo aver estratto i dati necessari prelevandoli dai *Big Data* è possibile inserirli in un *Data Lake*⁶⁰, che è un archivio digitale informatico centralizzato (*repository*) destinato alla memorizzazione, analisi e correlazione di dati strutturati e non strutturati, in formato nativo. La sua peculiarità è quella di consentire il recupero e l'organizzazione dei dati secondo il tipo di analisi che s'intende effettuare⁶¹. Siccome i *Big Data* sono caratterizzati dal loro enorme volume, dall'incredibile velocità con cui producono e memorizzano le informazioni e dalla varietà e provenienza dei dati trattata, sono difficilmente inquadrabili con le tradizionali tecniche di organizzazione dei database⁶². Ed è proprio per queste considerazioni che vengono utilizzati i *Data Lake*, i quali acquisiscono e conservano dati non ancora elaborati per uno scopo specifico.

⁵⁷ Un po' come accadde nella transizione dall'architettura di rete centralizzata (*mainframe* con terminali "stupidi" dotati solo di schermo e tastiera) e quella distribuita con terminali intelligenti provvisti di capacità di memoria e di elaborazione.

⁵⁸ La *larghezza di banda* è la quantità massima di dati che è possibile inviare attraverso il percorso di comunicazione in un secondo. Nel caso di una connessione internet indica la velocità e la quantità di informazioni che è in grado di trasportare e si esprime in bit/s.

⁵⁹ La *latenza di rete* misura il ritardo nella comunicazione di rete, ovvero l'intervallo di tempo, normalmente espresso in millisecondi, che intercorre fra il momento in cui viene inviato il segnale al sistema e il momento in cui è disponibile la sua risposta. Le reti con un ritardo più lungo hanno una latenza elevata, mentre quelle con tempi di risposta rapidi hanno una latenza bassa.

⁶⁰ Termine coniato da James Dixon, all'epoca Direttore Tecnico dell'azienda americana di software Pentaho, che nel 2010 coniò il termine di *Data Lake* con l'obiettivo di creare un'unica fonte per tutti i dati interessanti per un'azienda.

⁶¹ Lo scopo del *data lake*, pertanto, è quello di rendere disponibile una visione dei dati non necessariamente raffinata a supporto delle attività di *data discovery*, caratteristica che lo rende adatto a utenti esperti. Al contrario, il *data warehouse* (DWH) ha l'obiettivo di rendere disponibile, attraverso tool di *business* e *big data analytics*, una visione di dati elaborati per uno scopo e/o un processo di business ben preciso (cfr. SILEO in Siti web).

⁶² In informatica, un *database* (base di dati o banca dati) è un insieme di dati organizzati immagazzinati e accessibili per via elettronica.



Figura 14. Rappresentazione visiva di un *Data Lake*⁶³.

Un *dato* in un *Data Lake* non viene definito fino al momento in cui non viene eseguita una *query*⁶⁴ che lo coinvolga. La grande novità risiede nella possibilità di archiviare dati con formati molto differenti senza necessità di doverli uniformare e “normalizzare”. Questo permette di estrarre dati da qualunque fonte informativa senza che questa sia organizzata come un *data base* relazionale, organizzato in tabelle con strutture e caratteristiche definite.

6. L'INTELLIGENZA ARTIFICIALE IN DIAGNOSTICA E CHIRURGIA

Dopo aver analizzato gli strumenti a disposizione nella nostra “cassetta degli attrezzi” andiamo a utilizzarli nella pratica medica. La prima applicazione pratica che andremo a esaminare riguarda la *diagnostica*⁶⁵, settore che si è andato sviluppando velocemente negli ultimi anni, soprattutto per l'incessante progredire delle tecnologie anche in questo ambito. Normalmente, al fine di pervenire a una *diagnosi* si procede a reperire

⁶³ Fonte: <<https://www.linkedin.com/pulse/big-data-vs-lake-mustafa-qizilbash>>.

⁶⁴ Una *query* in informatica consente di reperire e utilizzare dati. Una *query* può fornire una risposta a una semplice domanda, eseguire calcoli, combinare dati di tabelle diverse, aggiungere, modificare o eliminare dati da un *database*.

⁶⁵ La *diagnostica* indica il complesso di atti - sia conoscitivi sia valutativi - necessari a pervenire alla *diagnosi*, ossia a identificare il quadro clinico presentato da un paziente.

dati del paziente mediante l'*anamnesi*⁶⁶, l'*esame obiettivo*⁶⁷ e gli *esami strumentali*. All'interno della grande categoria degli esami strumentali (*test diagnostici*) si raggruppano gli esami atti a incrementare le informazioni sulle condizioni del malato e sottoporre a verifica le ipotesi diagnostiche formulate con l'anamnesi e l'esame obiettivo, aiutando così il medico a orientarsi con maggior precisione verso la diagnosi definitiva.

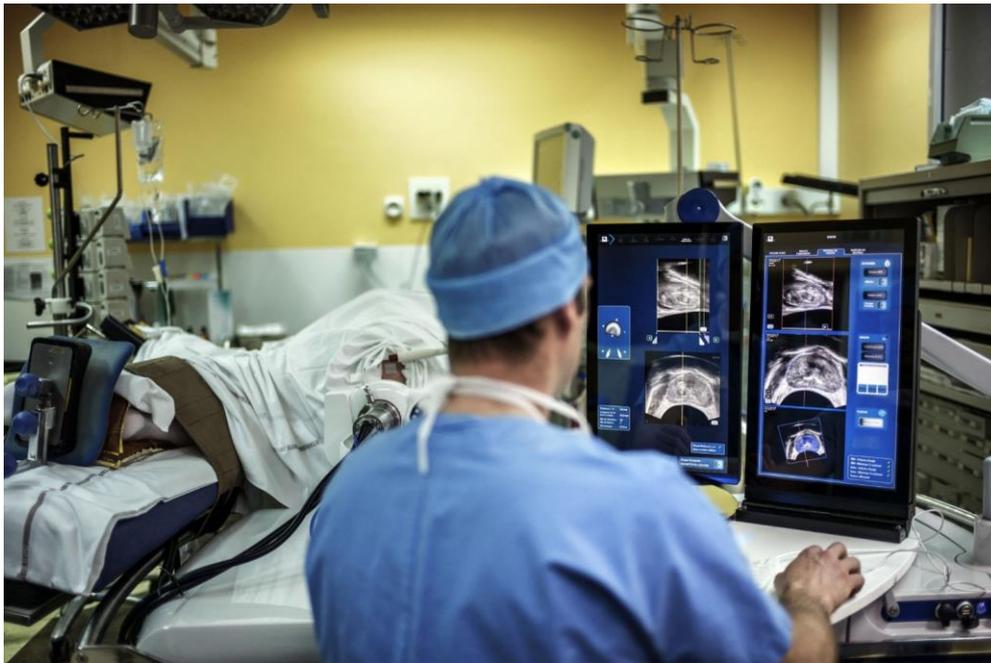


Figura 15. L'Intelligenza Artificiale applicata alla Diagnostica⁶⁸.

Le *indagini diagnostiche* si suddividono in *test di laboratorio*, *metodiche di immagine* e *test funzionali*. Insomma si tratta di reperire un grande numero di dati e confrontarli con quelli di cartelle cliniche similari e di possibili patologie. Ed è qui che facciamo intervenire l'Intelligenza Artificiale (IA), che dispone di una memoria enorme, di una incredibile velocità di elaborazione e di un'illimitata capacità di reperimento di informazioni tramite Internet (*Big Data*).

⁶⁶S'intende per *anamnesi* la storia clinica di un soggetto in esame, raccolta dal medico direttamente o indirettamente come elemento fondamentale per la formulazione della diagnosi; comprende anche le notizie sui precedenti ereditari e sullo stato di salute dei familiari.

⁶⁷ Per *esame obiettivo* (generalmente abbreviato nelle cartelle cliniche come "e.o.") s'intende l'insieme di manovre diagnostiche effettuate dal medico per verificare la presenza o assenza, nel paziente, dei segni (o *sintomi obiettivi*) indicativi di una deviazione dalla condizione di normalità fisiologica.

⁶⁸ Fonte: <<https://www.bloomberg.com/opinion/articles/2023-04-15/how-artificial-intelligence-like-chatgpt-is-influencing-medical-diagnoses>>.

L'IA memorizza tutti i dati del paziente (anamnesi ed esame obiettivo), analizza quelli provenienti dai test diagnostici e li confronta con quelli provenienti dai *Big Data*, dopo averne estratto quelli pertinenti e averli depositati in un apposito *Data Lake*.

L'IA utilizza algoritmi avanzati di analisi e apprendimento automatico per estrarre approfondimenti e generare risultati rilevanti e il tutto con una tempistica da record. Insomma, l'IA offre al medico i risultati degli esami in modalità sintetica, corredati da una serie di "linee diagnostiche guida" e, nella variante dell'IA generativa, anche previsioni riguardanti il possibile insorgere di disfunzioni o patologie.

Abbiamo visto che lo scopo principale dell'intelligenza artificiale è quello di replicare e persino superare le capacità cognitive umane in vari ambiti, inclusa l'analisi delle immagini nell'ambito della diagnostica, ove l'IA include attività quali⁶⁹: la classificazione delle immagini, ovvero l'assegnazione di un'etichetta o di una categoria a un'intera immagine, il rilevamento di oggetti, ovvero l'identificazione e la localizzazione di un insieme di oggetti all'interno di un'immagine, la segmentazione semantica, ovvero la partizione di un'immagine in più parti, assegnando un'etichetta di classificazione a ciascun pixel. L'intelligenza artificiale in genere coinvolge, poi, l'*apprendimento automatico*, un sottocampo che si concentra sulla capacità dei computer di apprendere dai dati e fare previsioni o assumere decisioni senza essere esplicitamente programmati. Purtroppo i metodi attualmente in uso di apprendimento automatico possono spesso incorrere in problemi, vista l'elevata dimensione dei dati pixel grezzi e questo limita la loro efficacia in presenza di immagini complesse. Come abbiamo visto, il *deep learning* è un sottocampo dell'apprendimento automatico che si occupa di algoritmi ispirati alla struttura e alla funzione del cervello, in particolare alle reti neurali a più strati. Nell'analisi delle immagini si è rivelata particolarmente utile un'architettura speciale, denominata *reti neurali convoluzionali* (CNN)⁷⁰. Come altri modelli di *deep learning*, le CNN

⁶⁹ Cfr. LINDGREN BELAL *et al.* 2024.

⁷⁰ Nell'apprendimento automatico, una *rete neurale convoluzionale* (CNN o ConvNet dall'inglese *convolutional neural network*) è un tipo di rete neurale artificiale *feed-forward* in cui il pattern di connettività tra i neuroni è ispirato dall'organizzazione della corteccia visiva animale, i cui neuroni individuali sono disposti in maniera tale da rispondere alle regioni di sovrapposizione che tassellano il campo visivo.

possono apprendere rappresentazioni gerarchiche delle funzionalità direttamente dai dati grezzi dei pixel, alleviando così la necessità di funzionalità manuali.

La *Radiomica*⁷¹ è un gruppo di metodi strettamente correlati all'apprendimento automatico classico, in quanto prevede l'estrazione di un gran numero di caratteristiche quantitative da immagini mediche, come caratteristiche di forma, struttura e intensità. Le tecniche di apprendimento automatico vengono quindi impiegate per costruire *modelli predittivi* utilizzando queste funzionalità radiomiche.

Nel prossimo futuro, si auspica che i dati raccolti dagli esami di *imaging* radiologico siano convertiti in dati quantitativi e che questi dati siano sfruttati come supporto decisionale alla pratica clinica per migliorare l'accuratezza diagnostica e il potere prognostico.

6.1 ELETTROCARDIOGRAMMA E INTELLIGENZA ARTIFICIALE

Prendiamo ad esempio l'*elettrocardiogramma* (ECG), un test oramai standardizzato, che mette a disposizione preziose informazioni sulla salute e sulle malattie cardiache e non cardiache⁷².

Ebbene la sua interpretazione richiede una notevole esperienza professionale. I metodi avanzati di intelligenza artificiale coadiuvano l'operatore, consentendo un'interpretazione rapida e precisa, rendendo così disponibile questa tecnologia

⁷¹ Per *Radiomica* s'intende l'analisi delle immagini mediche volta a ottenere, tramite opportuni metodi matematici e l'uso dei computer, informazioni di tipo *quantitativo*, assolutamente non rilevabili tramite la loro semplice osservazione visiva da parte dell'operatore. La "medicina di precisione" è un obiettivo chiave della medicina oncologica moderna, i trattamenti vengono personalizzati sulla base delle caratteristiche specifiche del paziente e della sua malattia e la Radiomica sta rapidamente emergendo in questo ambito ed è uno dei campi di ricerca di maggior interesse. Attraverso la Radiomica le immagini mediche da noi conosciute, ottenute dagli esami TC, RM o PET, vengono convertite in informazioni numeriche. Per meglio comprendere cosa sia la Radiomica è necessario premettere che alcuni tumori sono caratterizzati da alterazioni molecolari, come ad esempio quelle genomiche. Perché sia possibile definire queste alterazioni, generalmente, è necessario avere un campione del tessuto neoplastico, che si ottiene solo con biopsie o interventi chirurgici invasivi. Oggi però la diagnostica per immagini può permettere di caratterizzare i tessuti in modo non invasivo e, in alcuni casi, di visualizzare le profonde differenze fenotipiche dei diversi tumori. Poiché i tumori sono eterogenei nel loro volume e si modificano nel tempo, le immagini diagnostiche possono fornire una visione completa dell'intero tumore e possono essere ripetute nel tempo in modo non invasivo per monitorare le modificazioni indotte anche dalle terapie. Tali informazioni numeriche descrivono l'immagine stessa. Il numero finale di informazioni estrapolate dalle immagini è talvolta elevato e richiede metodiche dedicate ai cosiddetti "*Big Data*". Questo enorme patrimonio di dati numerici, che va ben oltre la semplice osservazione visiva, definisce molte caratteristiche del tumore (forma, volume e struttura tissutale) e dell'ambiente circostante. Con tali tecniche, è possibile studiare l'eventuale associazione fra i dati ottenuti dalle immagini e le caratteristiche molecolari e genomiche del tumore, con l'obiettivo finale di estrarre direttamente dalle immagini indicazioni sull'aggressività della malattia, sulle terapie più indicate e sulla risposta alle cure (cfr. ISTITUTO EUROPEO DI ONCOLOGIA in Siti web, alla voce "Radiomica").

⁷² Cfr. SIONTIS 2021.

anche a livello diffuso, in quanto si constata un rapido aumento della disponibilità di tecnologie ECG mobili e indossabili. Ciò risulta essenziale per il rilevamento di malattie cardiovascolari nelle popolazioni a rischio. Inoltre, tutti quei segnali e pattern⁷³ in gran parte irriconoscibili per gli operatori umani possono essere rilevati con precisione da reti di intelligenza artificiale multistrato, rendendo così l'ECG un potente biomarcatore non invasivo. Il rilevamento dell'aritmia ventricolare tramite AI ha raggiunto un'accuratezza, sensibilità e specificità del 99,2–98,8%⁷⁴.



Figura 16. Intelligenza Artificiale applicata all'ECG⁷⁵.

Una recente ricerca, che ha utilizzato 8,5 milioni di ECG di 2,1 milioni di pazienti raccolti in un periodo di quattro decenni e che si basa su un approccio di apprendimento simile a quello di ChatGPT, si è avvalsa di modelli basati su immagini per l'analisi dell'ECG, consentendo potenzialmente una valutazione più efficace della funzionalità cardiaca e la diagnosi preventiva di infarti e di cardiomiopatie ipertrofiche⁷⁶.

6.2 RISONANZA MAGNETICA E INTELLIGENZA ARTIFICIALE

Nell'ambito della diagnostica per immagini la *Risonanza Magnetica* ha un posto di eccellenza. Si tratta di una modalità di *imaging* che consente la visualizzazione non

⁷³ I pattern ECG sono generalmente classificati in tre categorie a seconda che il segnale provenga dal nodo SA (cioè ritmi sinusali), dagli atri (cioè ritmi atriali) o dai ventricoli (cioè ritmi ventricolari).

⁷⁴ Cfr. MARTINEZ-SELLES, MARINA-BREYSSE 2023.

⁷⁵ Fonte: <<https://www.tctmd.com/news/artificial-intelligence-ecgs-may-id-asymptomatic-lv-dysfunction>>.

⁷⁶ Cfr. VAID *et al.* 2023.

invasiva delle strutture e organi interni del corpo e delle loro funzioni, ed è ampiamente utilizzata per varie applicazioni cliniche in settori quali la neurologia, l'oncologia, la cardiologia, l'ortopedia, la pediatria⁷⁷.

La risonanza magnetica ha diversi vantaggi rispetto ad altre modalità di *imaging*, tra cui la mancanza di radiazioni ionizzanti, la capacità di generare un eccellente contrasto dei tessuti molli e la capacità di acquisire immagini in qualsiasi piano, orientamento e profondità. Tuttavia, la RM è caratterizzata da alcune limitazioni sostanziali, soprattutto relativamente al lungo tempo di scansione, che si traduce in un costo elevato e una maggiore sensibilità ai movimenti del paziente nel corso della prova.

Recentemente, le tecniche basate sull'intelligenza artificiale hanno consentito notevoli progressi nell'affrontare queste limitazioni, fornendo un'acquisizione accelerata, una migliore sintesi delle immagini e quantificazione dei parametri, una segmentazione automatizzata e una ottimizzazione nella pianificazione della scansione. Un'altra peculiarità in cui l'intelligenza artificiale ha apportato notevoli vantaggi nella diagnostica RM per immagini sono le diagnosi e le prognosi automatizzate.

6.3 TOMOGRAFIA A EMISSIONE DI POSITRONI E INTELLIGENZA ARTIFICIALE

La *tomografia a emissione di positroni* (o PET, *Positron Emission Tomography*) è una tecnica diagnostica medica di medicina nucleare utilizzata per la produzione di bioimmagini (immagini del corpo). A differenza della tomografia computerizzata (TC) e della risonanza magnetica nucleare (RM), che forniscono informazioni di tipo morfologico, la PET dà informazioni di tipo fisiologico permettendo così di ottenere mappe dei processi funzionali all'interno del corpo. Maggiori informazioni possono pervenire da una combinazione di PET e TAC o RM⁷⁸.

Le modalità di *imaging* della medicina nucleare, come la PET, presentano però elevati livelli di rumore e bassa risoluzione spaziale, rendendo necessario il miglioramento delle

⁷⁷ Cfr. SHIMRON, PERLMAN 2023.

⁷⁸ La PET-TC è un'indagine innovativa della diagnostica per immagini. Gli studi vengono eseguiti mediante macchine "ibride" cioè tomografi PET-TC o PET-RM nei quali il sistema di rilevazione PET viene combinato con la TAC o con la Risonanza Magnetica (RM).

immagini post-ricostruzione, al fine di migliorarne la qualità e l'accuratezza quantitativa. I modelli di intelligenza artificiale, come le *reti neurali convoluzionali*, le *reti generative avversarie*⁷⁹ e le *U-net*⁸⁰, hanno prodotto risultati molto promettenti nel miglioramento di tali immagini⁸¹. Ad esempio, i modelli di *deep learning* hanno dimostrato una grande potenzialità nel ridurre la dose di radiotracciante⁸² immesso e i relativi tempi di scansione, senza sacrificare la qualità dell'immagine e l'accuratezza diagnostica.



Figura 17. Chirurgia robotizzata⁸³.

⁷⁹ Una *rete generativa avversaria* (in inglese *generative adversarial network* o GAN) è una classe di metodi di apprendimento automatico in cui due reti neurali vengono addestrate in maniera competitiva nel contesto di un gioco a somma zero. Ciò consente alla rete neurale di apprendere come generare nuovi dati aventi la stessa distribuzione dei dati usati in fase di addestramento. Ad esempio, è possibile ottenere una rete neurale in grado di generare volti umani iperrealistici.

⁸⁰ La U-Net è un tipo di architettura di *deep learning* progettata specificamente per la segmentazione delle immagini mediche. È costituita da una rete di decodificatori che campiona l'immagine di input per apprendere una rappresentazione compressa e da una rete di decodificatore, che campiona la rappresentazione compressa per generare una mappa di segmentazione.

⁸¹ Cfr. BALAJI *et al.* 2024.

⁸² La PET impiega sostanze radiomarcate con radioisotopi che emettono positroni. La dose di irradiazione è equivalente circa a quella che si assume quando si esegue una TAC.

⁸³ Fonte: <<https://www.smithsonianmag.com/innovation/the-past-present-and-future-of-robotic-surgery-180980763/>>.

6.4 CHIRURGIA E INTELLIGENZA ARTIFICIALE

Abbiamo potuto constatare come l'intelligenza artificiale (IA), basata su di algoritmi che consentono alle macchine di svolgere funzioni cognitive (come la risoluzione di problemi e il processo decisionale), ha cambiato ormai da tempo il volto dell'assistenza sanitaria attraverso l'apprendimento automatico (ML, *Machine Learning*) e l'elaborazione del linguaggio naturale (NLP, *Natural Language Processing*).

Così, anche i *robot chirurgici*, appositamente guidati dall'IA, divengono dispositivi governati dal computer per lo svolgimento di ben precise funzioni in sala operatoria, consentendo ai chirurghi "fisici" di concentrarsi sugli aspetti complessivi di un intervento. Il loro utilizzo riduce altresì i possibili "ondeggiamenti" osteomuscolari degli operatori fisici durante l'intervento chirurgico e li aiuta a migliorare le proprie capacità e a ottenere risultati migliori⁸⁴. Con l'aiuto dell'Intelligenza Artificiale, i robot chirurgici riescono a identificare informazioni critiche e pratiche all'avanguardia, esaminando e valutando un enorme numero di dati.

7. I GEMELLI DIGITALI IN CHIRURGIA, RICERCA E FORMAZIONE

S'intende per *gemello digitale* un modello digitalizzato di un prodotto, sistema o processo fisico, previsto o reale, del mondo reale (che ne è il suo *gemello fisico*), il quale funge da controparte digitale effettivamente indistinguibile e ciò per scopi pratici, come simulazione, integrazione, test e monitoraggio⁸⁵.

I gemelli digitali vengono generati combinando dati provenienti da sensori e altre fonti, al fine di creare un *modello virtuale informatizzato* di un'entità⁸⁶. Essi sono repliche virtuali di entità fisiche che consentono di comprenderne le dinamiche e le specifiche funzionalità degli organismi umani. In medicina, l'applicazione più ambiziosa dei gemelli digitali mira a raggiungere una rappresentazione digitale completa del paziente fisico, al fine di migliorare la prevenzione, la diagnosi e il trattamento delle malattie.

⁸⁴ Cfr. Kwo 2021.

⁸⁵ Cfr. GROSSI 2023c.

⁸⁶ Cfr. SHU *et al.* 2023.

Ad esempio, la creazione di gemelli digitali in chirurgia, consente la pianificazione operatoria degli interventi in anticipo, verificandoli con delle simulazioni, nel contesto delle caratteristiche anatomiche specifiche di ciascun paziente e delle relative variazioni fisiologiche, combinando un resoconto realistico delle prestazioni effettuate con la capacità di misurare i parametri intraoperatori, onde evitare danni all'organismo e riducendo così al minimo i rischi per il paziente⁸⁷.

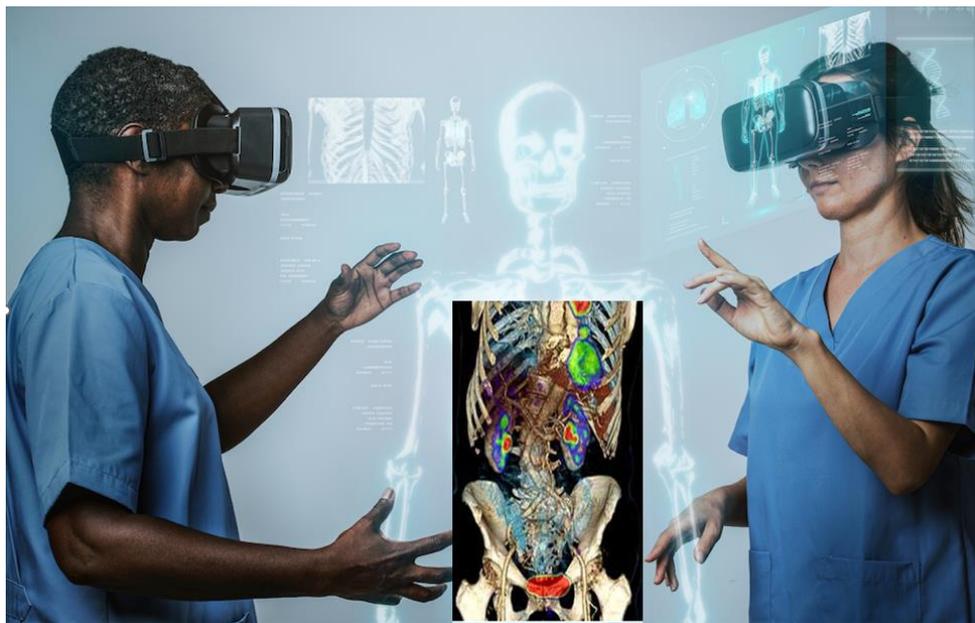


Figura 18. Esempio di utilizzo dei Gemelli Digitali in configurazione preoperatoria con la Realtà Aumentata⁸⁸.

Se utilizzati intraoperatoriamente, i gemelli digitali possono fornire una guida in tempo reale ai chirurghi per una consapevolezza situazionale complementare e, a loro volta, facilitare il loro processo decisionale. Infine, i gemelli digitali consentono di digitalizzare completamente le procedure per la tenuta dei registri, la valutazione postoperatoria e la generazione di set di dati per gli algoritmi di apprendimento automatico⁸⁹.

Inoltre, recenti studi hanno comprovato che i pazienti esposti alla realtà virtuale preoperatoria dimostrano una maggiore soddisfazione nell'effettuare l'intervento

⁸⁷ Cfr. KAPLAN *et al.* 2023.

⁸⁸ Fonte: GROSSI 2023c.

⁸⁹ L'apprendimento automatico (*machine learning*) nasce dalla teoria che i computer possono imparare a eseguire compiti specifici senza essere programmati per farlo, grazie al riconoscimento di schemi tra i dati. Il *machine learning* utilizza algoritmi che imparano proprio dai dati in modo iterativo.

chirurgico⁹⁰. Sfruttando la potenza di questa tecnologia, gli ospedali possono creare un ambiente coinvolgente, mediante le prassi proprie dell'Ergonomia, tendente a ridurre al minimo lo stress e migliorando così l'esperienza perioperatoria.

Nella telechirurgia le immagini dei gemelli digitali sono visualizzate con appositi visori di realtà aumentata (AR) e di realtà virtuale (VR), sovrapponendo le scansioni o i piani preoperatori all'anatomia del paziente per la guida intraoperatoria.

Un'ulteriore applicazione trova ulteriore spazio anche nella formazione accademica e di ricerca, ove il modello digitale dinamico del paziente, collocato in adeguate piattaforme di realtà virtuale, potrebbe anche dare origine a innovativi risultati clinici, tramite sperimentazioni in cui vengono verificati nuovi strumenti, tecniche o terapie e poi testati, in prima istanza, tramite i gemelli digitali.

8. IL METAVERSO IN MEDICINA

Abbiamo visto che l'utilizzo del *Metaverso* può facilitare lo sviluppo, la prototipazione, la valutazione, la regolamentazione, e la ristrutturazione della pratica medica basata sull'Intelligenza Artificiale e, in particolare, facilitare le diagnosi e le terapie guidate dall'*imaging* medico. Per sviluppare tutte le sue potenzialità, però, il Metaverso deve integrare su vasta scala tutte le tecnologie ICT, come Internet ad alta velocità, 5G/6G⁹¹, Realtà Virtuale, Realtà Aumentata, Realtà Mista, Realtà Estesa, Gemelli Digitali, Olografia⁹², Calcolo computazionale⁹³ e Intelligenza Artificiale, consentendo altresì ai medici di poter interagire tra di loro e con gli Avatar⁹⁴.

⁹⁰ Cfr. BEKELIS *et al.* 2017.

⁹¹ I termini 5G e 6G (acronimi di 5 *Generazione* e 6 *Generazione*) indicano l'insieme di tecnologie di telefonia mobile e cellulare, i cui standard definiscono la quinta e la sesta generazione della telefonia mobile con una significativa evoluzione rispetto alle tecnologie precedenti.

⁹² L'*olografia* è una tecnologia ottica di memorizzazione di un'informazione visiva sotto forma di un finissimo intreccio di frange di interferenza con impiego di luce laser coerente, opportunamente proiettata e l'immagine creata dalle frange di interferenza è caratterizzata da una illusione di tridimensionalità.

⁹³ La *teoria della complessità computazionale* è una branca della *teoria della computabilità* che studia le risorse minime necessarie (principalmente tempo di calcolo e memoria) per la risoluzione di un problema. Con *complessità di un algoritmo* o *efficienza di un algoritmo* ci si riferisce dunque alle risorse di calcolo richieste. I problemi sono classificati in differenti *classi di complessità*, in base all'efficienza del migliore algoritmo noto in grado di risolvere quello specifico problema.

⁹⁴ Per *Avatar* s'intende, in Informatica, la rappresentazione grafica utilizzata da un utente di servizi Internet (per es. le *chat line* o alcuni videogiochi) per identificarsi quando entra in comunicazione con altri utenti.

Vediamo ora un esempio esplicativo di applicazione del Metaverso con Gemelli Digitali e Intelligenza Artificiale esposto in un disegno grafico⁹⁵.

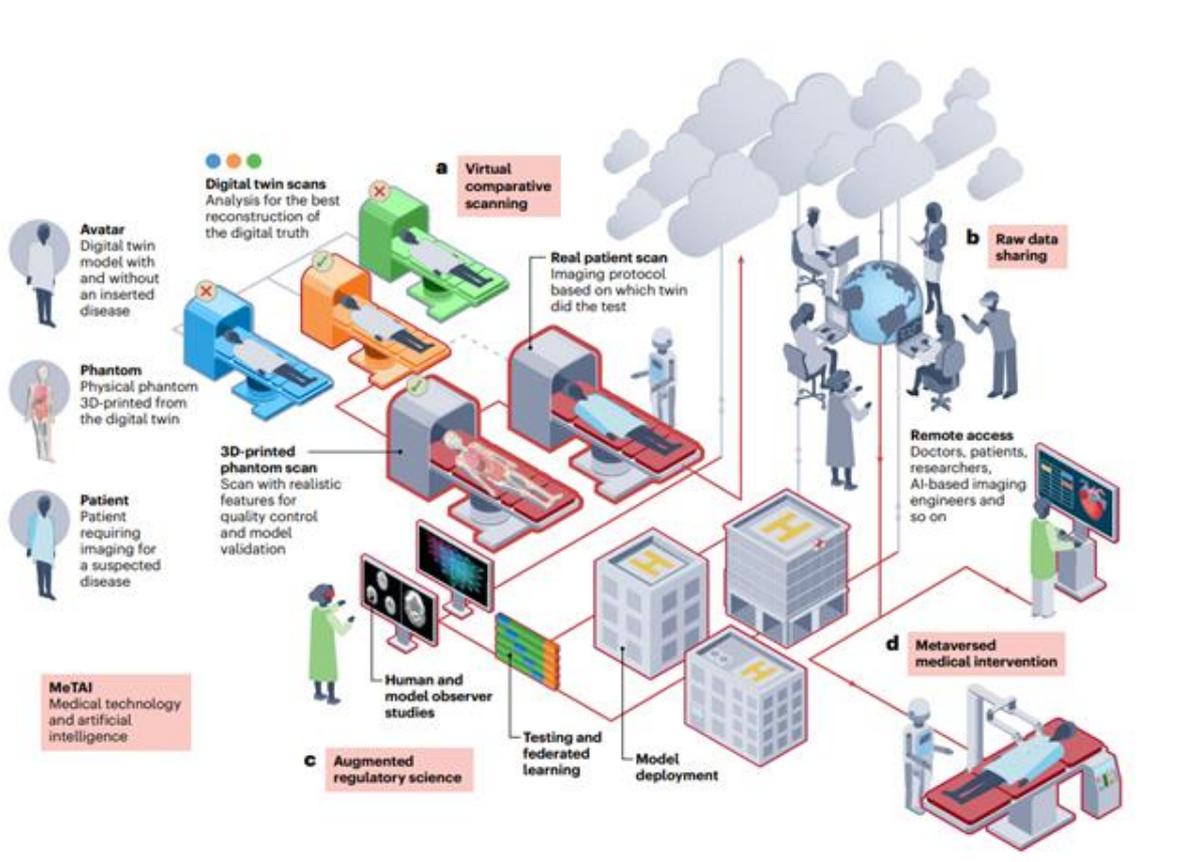


Figura 19. Descrizione del processo: a) scansione comparativa virtuale (per trovare la migliore tecnologia di *imaging* in una situazione specifica); b) condivisione dei dati grezzi (per consentire un accesso aperto e controllato ai dati grezzi tomografici); c) scienza conoscitiva aumentata (per estendere le sperimentazioni cliniche virtuali in termini di portata e durata); d) intervento medico in “metaverso” (per eseguire un intervento medico aiutato dal metaverso). In un’implementazione esemplare di questo ecosistema, prima che un paziente venga sottoposto a una vera Tomografia Computerizzata, le sue scansioni vengono simulate su varie macchine virtuali per trovare il miglior risultato di *imaging*. (a) Sulla base di queste conoscenze viene eseguita una vera e propria scansione. Quindi, le immagini del metaverso vengono trasferite al team di assistenza medica del paziente e, previo consenso del paziente e in base a protocolli di calcolo sicuri, le immagini e i dati grezzi tomografici possono essere resi disponibili ai ricercatori (b). Tutte queste immagini e i dati reali e simulati, nonché altre informazioni rilevanti dal punto di vista medico, possono essere integrati nel metaverso e utilizzati in studi clinici aumentati (c). Infine, se clinicamente indicato, il paziente verrà sottoposto a un intervento chirurgico robotico remoto assistito dal metaverso e seguito nel metaverso per la riabilitazione (d)⁹⁶.

⁹⁵ Cfr. WANG *et al.* 2022.

⁹⁶ Fonte: WANG *et al.* 2022.

9. CONCLUSIONI

Abbiamo visto che l'utilizzo dell'intelligenza artificiale, della realtà virtuale e del Metaverso in medicina porterà a traguardi sempre più entusiasmanti, aumentando le potenzialità del medico, in quanto queste tecnologie fungono già oggi da estensioni e amplificazioni dei suoi sensi. Inoltre lo coadiuveranno sia per la loro velocità di elaborazione e di trasferimento delle informazioni, sia nella precisione delle diagnosi e nella formulazione delle adeguate terapie, sino a effettuare predizioni prognostiche sul decorso e l'evoluzione di determinati quadri morbosi o l'insorgere di eventuali patologie. Nel bel mezzo di tutte queste tecnologie innovative e sempre più performanti, la figura professionale del medico rifulgerà in maniera sempre più importante, essenziale e centrale del processo ergonomico, cioè antropocentrico, di studio, prevenzione e cura riguardo alle malattie delle persone. E la sua arma più importante consisterà nella forma di comunicazione empatica riguardo ai suoi assistiti, costantemente memore di aver prestato il proprio giuramento di Ippocrate.

RINGRAZIAMENTI

Si desidera ringraziare per l'attenzione concessaci **IBM Italia** e **Red Hat** per quanto riguarda l'Intelligenza Artificiale e i Gemelli Digitali e **Philips Healthcare** per le specifiche delle grandi apparecchiature diagnostiche.

BIBLIOGRAFIA

BALAJI V., SONG T. A., MALEKZADEH M., HEIDARI P., DUTTA J.
2024, «Artificial Intelligence for PET and SPECT Image Enhancement», *J Nucl Med*, 2 gennaio 2024; 65(1), pp. 4-12, DOI: 10.2967/jnumed.122.265000.

BEKELIS K., CALNAN D., SIMMONS N., MACKENZIE T. A., KAKOULIDES G.
2017, «Effect of an immersive preoperative virtual reality experience on patient reported outcomes: A randomized controlled trial», *Ann. Surg.*, 265 (6), pp. 1068-1073, DOI: 10.1097/SLA.0000000000002094.

BIANCO V., D'AGOSTINO M., PIRONE D., GIUGLIANO G., MOSCA N., DI SUMMA M., SCERRA G., MEMMOLO P., MICCIO L., RUSSO T., STELLA E., FERRARO P.
2023, «Label-Free Intracellular Multi-Specificity in Yeast Cells by Phase-Contrast Tomographic Flow Cytometry», *small methods*, Volume 7, Issue 12, 2300447, <<https://doi.org/10.1002/smt.202300447>>.

GROSSI F.
1999, *Tecnologie per l'Informazione e la Comunicazione, conoscere e comprendere i nuovi media*, Milano, Franco Angeli.

2000, *An Ergonomic Model of Communication*, in *Ergonomic Evaluation of Audio Components and Systems*, presentato alla AES Convention: 108 (February 2000) Paper Number: 5086, Publication Date: February 1, 2000.

2014a, «Human Communication», Kazan State University of Architecture and Engineering, inside the Course “Architecture of Communications”, scaricabile dall’indirizzo web:

<https://www.academia.edu/6654760/GROSSI_Franco_Human_Communication>.

2014b, *The AUAF Model of Communication*, Academia.edu, scaricabile dall’indirizzo web:

<https://www.academia.edu/12200615/The_AUAF_Model_of_Communication>.

2023a, «Ergonomia Medica / Medical Ergonomics», *QuaderniCIRD*, 26 (2023), pp. 33-56.

2023b, «Diseño Industrial y Comunicación Visual vs. Calidad Percibida / Industrial Design and Visual Communication vs. Perceived Quality, by A3manos», *Revista académica de la Universidad de La Habana*, Instituto Superior de Diseño (ISDi), vol. 10, No. 20, Julio - Diciembre 2023, pp. 16-22.

2023c, *Safety, Quality and Ergonomics Applied to Medicine*, presented at the XI International Scientific and Practical Forum *Effective Management Systems: Quality. Circular Economy. Technological Sovereignty*, Sessione *Qualità, sicurezza e innovazioni in medicina*, Kazan, November 23, 2023.

KAPLAN N., MARQUES M., SCHARF I., YANG K., ALKUREISHI L., PURNELL C., PATEL P., ZHAO L.

2023, «Virtual Reality and Augmented Reality in Plastic and Craniomaxillofacial Surgery: A Scoping Review», *Bioengineering Review*, Basel CH, 2023 Apr., 10(4): 480. Published online 2023 Apr 17. DOI: 10.3390/bioengineering10040480.

KWO L.

2021, «The power of AI in surgery», *Mobi Health News*, Australia, Global Edition, Digital Health, November 23, scaricabile dal sito web: <<https://www.mobihealthnews.com/news/contributed-power-ai-surgery>>.

LINDGREN BELAL S., FRANTZ S., MINARIK D., ENQVIST O., WIKSTRÖM E., EDENBRANDT L., TRÄGÅRDH E.

2024, «Applications of Artificial Intelligence in PSMA PET/CT for Prostate Cancer Imaging», *ELSEVIER, Seminars in Nuclear Medicine*, 54 (1), pp. 141-149, <<https://doi.org/10.1053/j.semnuclmed.2023.06.001>>.

MARR B.

2023, «The Difference Between Generative AI And Traditional AI: An Easy Explanation For Anyone», *Forbes*, Jersey City, New Jersey, USA, Jul 24, 2023, scaricabile dal sito:

<<https://www.forbes.com/sites/bernardmarr/2023/07/24/the-difference-between-generative-ai-and-traditional-ai-an-easy-explanation-for-anyone/?sh=319f7af8508a>>.

MARTÍNEZ-SELLÉS M., MARINA-BREYSSE M.

2023, «Current and Future Use of Artificial Intelligence in Electrocardiography», *J Cardiovasc Dev Dis.*, National Library of Medicine, USA, 2023 Apr 17, 10(4):175. DOI: 10.3390/jcdd10040175.

MCLUHAN M.

2001 (seconda edizione), *Understanding Media: The Extension of Man*, Milton Park, Abingdon (U.K.), Routledge.

MCLUHAN M., POWERS B. R.

1992 (prima edizione), *The Global Village: Transformations in World Life and Media in the 21st Century*, Oxford University Press (U. K.).

SHANNON C. E.

1948, «A Mathematical Theory of Communication», *Bell System Technical Journal*, Vol. 27, pp. 379-423, 623-656, scaricabile dall'indirizzo web:

<[https:// people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf](https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf)>.

SHIMRON E., PERLMAN O.

2023, «AI in MRI: Computational Frameworks for a Faster, Optimized, and Automated Imaging Workflow», *Bioengineering (Basel)*, 2023 Apr 20, 10(4):492, doi: 10.3390/bioengineering10040492.

SHU H., LIANG R., LI Z., GOODRIDGE A., ZHANG X., DING H., NAGURURU N., SAHU M., CREIGHTON F. X., TAYLOR R. H., MUNAWAR A., UNBERATH M.

2023 *Twin-S: A Digital Twin Paradigm for Skull Base Surgery*, Johns Hopkins Medicine University, Baltimore, MD, United States. arXiv:2211.11863v2 [cs.HC], presentato il 7 Maggio 2023 alla 14^a *International Conference on Information Processing in Computer-Assisted Interventions (IPCAI)*, tenutasi a Monaco di Baviera.

SIONTIS K. C., NOSEWORTHY P. A., ATTIA Z. I., FRIEDMAN P. A.

2021, «Artificial intelligence-enhanced electrocardiography in cardiovascular disease management», *Nat. Rev. Cardiol.*, 18, pp. 465-478, scaricabile dal sito web:

<<https://doi.org/10.1038/s41569-020-00503-2>>.

VAID A., JIANG J., SAWANT A., LERAKIS S., ARGULIAN E., AHUJA Y., LAMPERT J., CHARNEY A., GREENSPAN H., NARULA J., GLICKSBERG B., NADKARNI GIRISH N.

2023, «A foundational vision transformer improves diagnostic performance for electrocardiograms», *npj Digit. Med.*, 6, 108, scaricabile dal sito web <<https://doi.org/10.1038/s41746-023-00840-9>>.

WANG G., BADAL A., JIA X., MALTZ J. S., MUELLER K., MYERS K. J., NIU C., VANNIER M., YAN P., YU Z., ZENG R.

2022, «Development of metaverse for intelligent healthcare», *Nature Machine Intelligence*, 4, pp. 922-929, (2022 November), scaricabile dal sito web: <<https://doi.org/10.1038/s42256-022-00549-6>>.

WATZLAWICK P., BAVELAS J. B., JACKSON D. D.

2014 (riedizione), *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies and Paradoxes*, WW Norton & Co, N.Y. (N.Y., USA).

SITI WEB

CAROBENE A.

Realtà virtuale,

<<https://www.treccani.it/enciclopedia/realta-virtuale/>>, sito consultato il 9.1.2024.

GROSSI F.

Un modello della comunicazione in chiave ergonomica,

<<https://sites.units.it/ergocom/docs/it/1.0.html>>, sito consultato il 9.1.2024.

ISTITUTO EUROPEO DI ONCOLOGIA – IEO

Radiomica: un passo verso la medicina di precisione,

<<https://www.ieo.it/it/CHI-SIAMO/Come-siamo-organizzati/Dipartimenti/Dipartimento-di-Immagini-e-Scienze-Radiologiche/Radiomica/>>, sito consultato il 9.1.2024.

SILEO F.,

Cos'è il data lake, quali vantaggi offre e come si differenzia dal data warehouse,

<<https://www.bigdata4innovation.it/big-data/cosae-il-data-lake-quali-vantaggi-offre-e-come-si-differenzia-dal-data-warehouse/>>, sito consultato il 9.1.2024.

TRECCANI

Big Data,

<https://www.treccani.it/vocabolario/big-data_res-007d6462-8995-11e8-a7cb-00271042e8d9_%28Neologismi%29/>, sito consultato il 9.1.2024.

Dossier “Matematica e ...”

Introduzione al Dossier “Matematica e ...”*

Nonostante non manchino le iniziative di divulgazione e cura dell’immagine della matematica e queste si siano intensificate fin da quando l’anno 2000 fu proclamato *Anno Mondiale della Matematica*¹ e il 14 marzo di ogni anno – giornata dapprima dedicata solo a *pi greco* – si festeggia dal 2020 quale *Giornata Internazionale della Matematica*², ci sono tuttora molte persone che ritengono che la matematica sia solo un insieme di rigide e astruse regole e, soprattutto, non ne conoscono le applicazioni oltre al mero calcolo aritmetico, e ancor meno la storia.

Entrambe le importanti iniziative prima citate, promosse dall’*IMU – International Mathematical Union*, hanno avuto il supporto dell’*UNESCO*, che le ha proclamate ufficialmente. Ne possiamo ben comprendere i motivi, se esaminiamo, ad esempio, i principali obiettivi della *Giornata Internazionale della Matematica* (qui riportati in traduzione italiana)³:

1. Migliorare la comprensione tra il grande pubblico, coloro che prendono le decisioni e nelle scuole, dell’importanza della matematica nell’istruzione.
2. Contribuire allo sviluppo delle competenze nell’educazione matematica e scientifica, con particolare attenzione alle ragazze e ai bambini dei paesi in via di sviluppo (SDG4).
3. Raggiungere l’uguaglianza di genere e potenziare le donne e le ragazze in matematica (SDG5).
4. Migliorare la comprensione tra il grande pubblico, coloro che prendono le decisioni e nelle scuole dell’importanza della matematica come strumento per uno sviluppo che porti a situazioni economiche più prospere (SDG9).
5. Sottolineare l’importanza della ricerca di base nelle scienze matematiche come germe di scoperte tecnologiche e di gestione della società (SDG8).
6. Evidenziare il ruolo della matematica nell’organizzazione della società moderna, compresi i sistemi economici, finanziari, sanitari e di trasporto, le telecomunicazioni, nella ricerca del benessere umano, ecc. (SDG3).

* Title: Introduction to the Dossier “Mathematics and ...”.

¹ Le indicazioni originali si possono ancora trovare a questo indirizzo:

<[² Per il sito ufficiale dell’*International Day of Mathematics*, dove è possibile registrare gli eventi organizzati e scaricare materiali utili anche per la didattica, vedi: <<https://www.idm314.org/>>.](https://www.emis.de/misc/cdrom/WMY2000/Jussieu/wmy2000.html#:~:text=On%20May%2C%206th%2C%201992%2C,be%20the%20World%20mathematical%20Year.&text=Many%20Mathematical%20Institutions%20are%20already,has%20to%20be%20really%20international!>.</p>
</div>
<div data-bbox=)

³ Per la versione inglese, vedi: <<https://www.idm314.org/about>>.

7. Sensibilizzare al ruolo della matematica nella lotta ai disastri, alle epidemie, alle malattie emergenti, alle specie invasive (SDG11).
8. Evidenziare il ruolo della matematica nel passaggio a un'economia circolare di sostenibilità compatibile con la conservazione della biodiversità (SDG 14 e 15).
9. Fornire al grande pubblico e ai giovani gli strumenti per comprendere le sfide planetarie e la capacità di rispondere come cittadini consapevoli.
10. Aumentare il networking [condivisione di informazioni o servizi tra persone o aziende] e le collaborazioni internazionali nella sensibilizzazione del pubblico alla matematica.
11. Aumentare l'accesso alle informazioni, fornendo un modo semplice per dare ai cittadini una scelta in tutti gli aspetti della loro vita quotidiana.

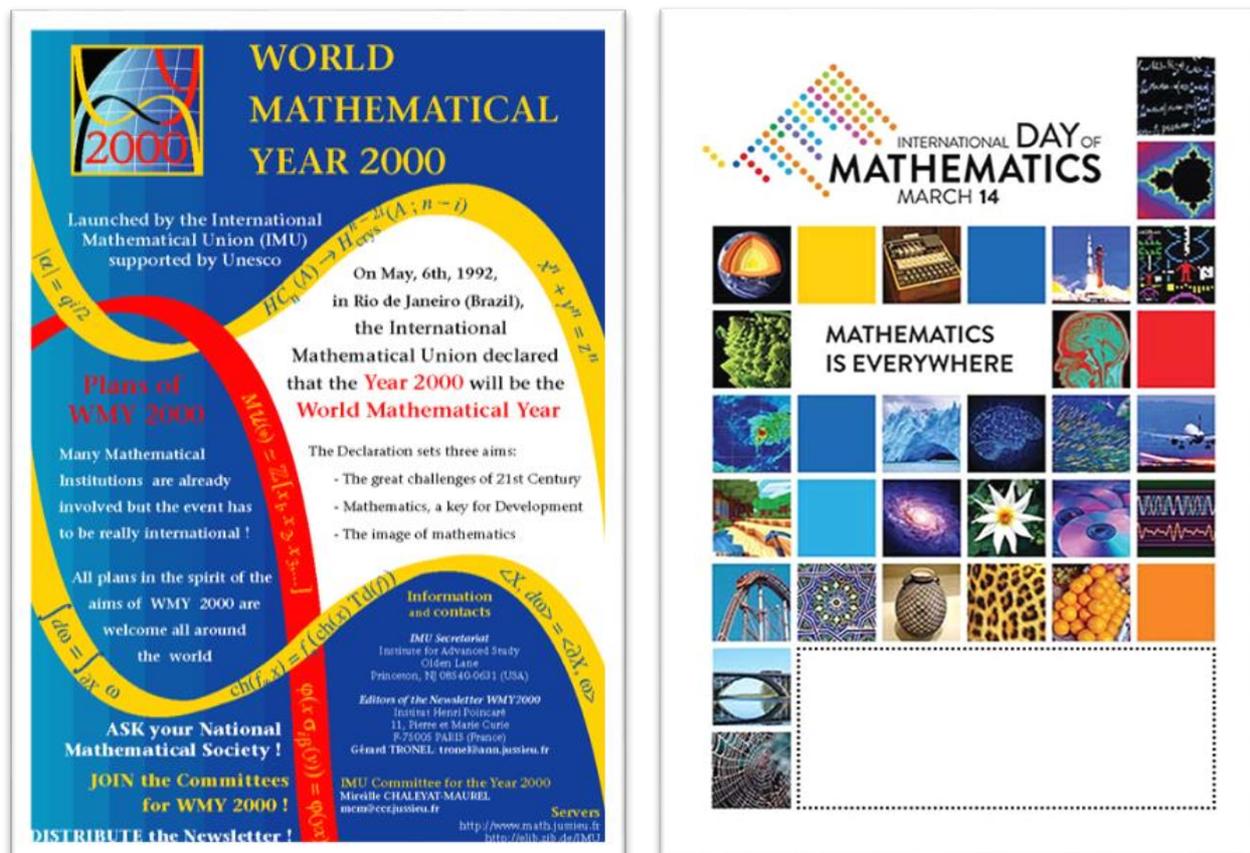


Figura 1. Il poster originale della proclamazione del 2000 Anno Internazionale della Matematica⁴ e quello della prima Giornata Internazionale della Matematica (2020)⁵.

In accordo con le indicazioni ministeriali per la Scuola primaria e secondaria, anche la Scuola deve farsi carico di porre in atto iniziative finalizzate a migliorare l'interesse degli allievi e delle allieve per la matematica e a sviluppare in essi un atteggiamento positivo nei suoi confronti.

⁴ Vedi <<https://www.emis.de/misc/cdrom/WMY2000/Jussieu/affiche.html>>.

⁵ Vedi <<https://www.idm314.org/decorate>>.

Con l'intento di fornire spunti ai docenti, non solo di matematica, ma anche di altre discipline, atti ad ampliare l'orizzonte culturale dei propri studenti, in questo numero di *QuaderniCIRD* si è deciso di raccogliere in un *dossier* un certo numero di contributi di carattere matematico che, nonostante la diversità dei temi trattati, hanno in comune l'attenzione alla divulgazione al grande pubblico di argomenti “non standard”, ma che, soprattutto, possono sollecitare la discussione in classe sul ruolo della matematica nelle altre discipline e nella vita sociale, collegandosi anche a temi di attualità. Naturalmente possono essere utilizzati, in tutto o in parte, direttamente nella didattica della matematica stessa.

Il primo in ordine di presentazione è *Distanze, lingue e parole*, di **Andrea Sgarro** e **Laura Franzoi**: in questo articolo si può vedere come la nozione matematica di *distanza*, opportunamente “addomesticata”, possa essere utile nella linguistica e in altri campi, come anche nell'informatica.

Segue *Sophie Germain e l'Ultimo Teorema di Fermat*, di **Elisabetta Podda** e **Alessandro Logar**. Qui gli autori – oltre a descriverne diffusamente l'opera in relazione ai tentativi di dimostrazione del cosiddetto *Ultimo Teorema di Fermat* – presentano l'interessante figura di Sophie Germain, matematica francese vissuta a cavallo tra il Settecento e l'Ottocento, e pongono in evidenza come la condizione femminile dell'epoca non le abbia permesso di frequentare gli ambienti accademici né di avere i giusti riconoscimenti scientifici che, ora, finalmente le sono stati attribuiti.

Il terzo contributo del dossier trae origine da una tesi di laurea in matematica. Si tratta di: *Come performa ChatGPT in matematica? Prime sfide su alcuni problemi elementari di Analisi 1*, di **Agnese Valentini**. In esso l'autrice “sfida” ChatGPT, il ben noto sistema che consente agli utenti di “comunicare” con un'intelligenza artificiale, ponendogli quesiti di matematica di tipo computazionale e dimostrativo e ne osserva le performance, con interessanti commenti che conducono, alla fine, a suggerimenti e avvertenze per un suo eventuale utilizzo nella didattica.

L'ultimo contributo qui presentato, *La crittografia classica come occasione di ragionamento*

matematico, di **Fabio Pasticci**, è particolarmente dedicato ai docenti della Scuola primaria. Qui si forniscono esempi di utilizzo in classe della crittografia, considerandola uno strumento utile per il recupero di abilità operative e il consolidamento di conoscenze matematiche, come già avvenuto in esperienze didattiche effettuate in precedenza da vari autori.

LUCIANA ZUCCHERI
Dipartimento MIGe
Università di Trieste
zuccheri@units.it

Distanze, lingue e parole*

ANDREA SGARRO

Dipartimento di Matematica, Informatica e Geoscienze
Università di Trieste
sgarro@units.it

LAURA FRANZOI

Dipartimento di Scienze Economiche, Aziendali, Matematiche e Statistiche
Università di Trieste
laura.franzoi@deams.units.it

ABSTRACT

Even if surprising for many mathematicians, quite a large number of the distances described in the Encyclopedia of distances, are not metric distances, i.e., they do not comply with some or other of the metric axioms which appear to be so natural or even unexpendable to those who tackle this multifaceted geometric and topological notion. Using examples taken from linguistics and word strings, we argue that the notion of distance is so rich and fruitful that the metric axioms in some cases risk to be an unreasonably narrow cage.

PAROLE CHIAVE

DISTANZA / DISTANCE; DISTANZA METRICA / METRIC DISTANCE; DISTANZA DI EDIT / EDIT DISTANCE;
LINGUISTICA COMPUTAZIONALE / COMPUTATIONAL LINGUISTIC.

1. CHE COS'È UNA DISTANZA?

Alla domanda appena formulata molti matematici risponderebbero senza esitazione enunciando la definizione di una *distanza metrica* o più concisamente di una *metrica*, definizione che ci sarà preziosa. Si parte da un insieme o se volete uno *spazio* S , i cui oggetti (i cui elementi), in numero finito o infinito, denoteremo con lettere come x , y o z (per evitare banalità converrà pensare che gli oggetti distinti siano almeno tre, anche se i matematici più diligenti si accontenterebbero di imporre che S non sia vuoto).

* Title: Distances, languages and words.

Ad ogni coppia (x,y) di oggetti distinti o anche coincidenti (può darsi che $x = y$) viene associato un numero non negativo $d(x,y)$ che è appunto la distanza fra x e y . Inutile dire che per poter usare un termine impegnativo come “distanza” conviene imporre qualche condizione di regolarità in modo da escludere “comportamenti inammissibili”. La distanza d si dice metrica se, qualunque siano gli oggetti x, y e z , distinti o coincidenti, valgono i seguenti tre assiomi:

- i) $d(x,y) = 0$ se e solo se $x = y$ (solo le auto-distanze sono nulle);
- ii) $d(x,y) = d(y,x)$ (simmetria della distanza);
- iii) $d(x,z) + d(z,y) \geq d(x,y)$ (disuguaglianza triangolare, cfr. Figura 1).

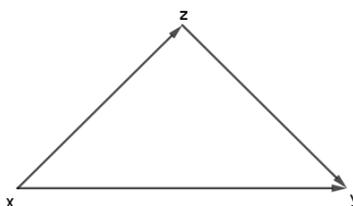


Figura 1. La disuguaglianza triangolare.

(i simboli \geq e \leq significano “maggiore o uguale a” e rispettivamente “minore o uguale a”); inutile osservare che la normalissima distanza euclidea nel piano o nello spazio è appunto metrica. Non cambia nulla di essenziale, ma precisiamo subito che talvolta vengono ammesse anche distanze infinite $d(x,y) = +\infty$.

Potremmo iniziare una discussione filosofica sull’opportunità o sulla naturalezza dei tre assiomi, ma il discorso si incaglia appena partiti. In quello che è il “libro sacro” delle distanze, vale a dire nell’*Encyclopedia of distances* di Deza e Deza¹, buona parte delle distanze elencate non sono affatto metriche, e ce ne sono a centinaia provenienti dai più diversi campi teorici e applicati. A giudizio degli autori, va da sé sindacabile, la richiesta minima al di sotto della quale non si dovrebbe scendere se si vuole evitare che l’uso del termine “distanza” diventi fuorviante è l’unico debolissimo assioma per ogni coppia di oggetti distinti x e y :

¹ Cfr. DEZA, DEZA 2016.

$$iv) \quad 0 \leq d(x,x) \leq \min[d(x,y), d(y,x)]$$

dove $\min[,]$ indica il minimo dei due argomenti nelle parentesi quadre. Per far funzionare meglio la nostra intuizione è auspicabile che valga la disuguaglianza triangolare ordinata *iii*) dove stavolta l'ordine in cui vengono specificati gli oggetti è essenziale, non essendo più garantita la simmetria. A questo punto la discussione filosofica è inevitabile, ma prima di impegnarsi nelle sezioni seguenti, anzi parallelamente alla discussione, vediamo di introdurre tre distanze estremamente notevoli che abbiamo scelto perché mettono in crisi gli assiomi metrici, pur essendo quelle dei §§ 2 e 3 delle distanze registrate a pieno titolo nell'*Encyclopedia of Distances*.

2. LA DISTANZA DI MULJAČIĆ

Nell'*Encyclopedia* è chiamata forse troppo generosamente *Sgarro distance*; altro nome usato in letteratura è quello di *fuzzy Hamming distance* visto che ha a che fare appunto con la logica *fuzzy* o sfocata o sfumata². Per non appesantire il discorso per il momento ci accontenteremo della distanza di Muljačić “monodimensionale”, limitata a singoli valori logici x e y che sono numeri reali dell'intervallo $[0,1]$: $0 \leq x \leq 1, 0 \leq y \leq 1$.

Nella *logica binaria* (o se volete nella logica matematica di tipo aristotelico, crisippino) esistono solo i valori logici 0=falso e 1=vero, mentre nella *logica fuzzy* sono ammessi tutti i valori intermedi: oltre al bianco e al nero ci sono tutte le sfumature di grigio. Potremmo ad esempio decidere, sulla base delle nostre opinioni, del nostro stato di conoscenze, che il valore logico della proposizione $P=\{\text{Alice parla correntemente il francese}\}$ sia 0,9, mentre il valore logico della proposizione $Q=\{\text{Bruno è un abile chitarrista}\}$ sia solo 0,3 (stiamo dando dei “voti” come a scuola, solo sulla scala da 0 a 1, ma si veda SGARRO 2019³ per maggiori dettagli).

Se x e y sono due valori logici, x il valore logico della proposizione P e y quello della proposizione Q , come possiamo definire la loro distanza $d(x,y)$ nello spirito della logica fuzzy? Indichiamo con AND, OR e NOT i tre *operatori logici* fondamentali di

² Circa la logica *fuzzy* o sfocata o sfumata si veda SGARRO 2019, pp. 37-45.

³ Cfr. SGARRO 2019.

congiunzione (P e anche Q, sia P sia Q), *disgiunzione* (P oppure indifferentemente Q, almeno uno dei due) e *negazione*. Una risposta, argomentata in SGARRO 2019⁴, potrebbe essere, usando i valori logici che sono numeri fra 0 e 1:

$$d(x,y) = (P \text{ è vera AND } Q \text{ è falsa}) \text{ OR } (P \text{ è falsa AND } Q \text{ è vera})$$

Nella logica fuzzy⁵ si definisce, argomentando opportunamente le definizioni:

$x \text{ AND } y = \min[x,y]$, $x \text{ OR } y = \max[x,y]$, $\text{NOT}x = 1-x$, per cui è adeguato porre, come appunto si fa con la distanza di Muljačić:

$$d(x,y) = \max[\min[x, 1 - y], [\min[1 - x, y]]]$$

Con un po' di pazienza si potrebbe dimostrare che valgono la *ii*) e la *iii*)⁶, ma ovviamente non vale la *i*), perché le auto-distanze possono essere strettamente positive. Un problema? Al contrario, un grande vantaggio nell'ambito della logica sfocata, visto che l'*autodistanza* $d(x,x)$ diventa una preziosa misura della "sfocatezza" del valore logico x , nulla solo se x è nitido o *crisp* (verità totale 1 o falsità totale 0) e massima in corrispondenza al valore totalmente "ambiguo" $x = \frac{1}{2}$, a metà strada esatta fra il vero e il falso.

Qualche chiarimento sul nome di Žarko Muljačić, un linguista spalatino che è stato uno dei massimi esperti delle lingue romanze o neo-latine. Scelse 40 parametri lessicali o sintattici da associare a 12 lingue neolatine⁷; i parametri potevano essere presenti o assenti nella lingua rispettiva, ma in alcuni casi la loro presenza/assenza era poco chiara per cui era saggio ricorrere a valori logici sfumati, di auto-distanza strettamente positiva. Ad esempio, al parametro 15 (comparativo formato con *plus* come nell'italiano *più buono* o nel francese *plus bon*, e non con *magis* come nel catalano *més bo* o nel rumeno *mai bun*) viene dato il valore di verità $\frac{1}{2}$ nel caso del provenzale, che oscilla fra i due.

Lo scopo di Muljačić era di dimostrare che il dalmatico, ormai estinto, è una "lingua-ponte" fra la latinità occidentale (italiano, francese, spagnolo, friulano, ecc.) e quella

⁴ Cfr. SGARRO 2019.

⁵ Cfr. SGARRO 2019.

⁶ Si veda ad esempio FRANZOI, SGARRO 2017.

⁷ Cfr. MULJAČIĆ 1967.

orientale, rappresentata essenzialmente dal rumeno. In dalmatico sono scritti ad esempio i *Praecepta Rectoris* della Repubblica di Ragusa-Dubrovnik (siamo nel 1280); l'ultimo parlante di dalmatico fu Antonio Udina (Tuane Udaina), morto a Veglia-Krk nel 1898, dopo essere stato, per fortuna della linguistica, ampiamente intervistato dal linguista albanese Matteo Bartoli.

Un'osservazione tecnica: la distanza fra stringhe di valori logici, nel nostro caso di 40 valori logici, è additiva e si trova sommando le distanze fra i corrispettivi valori (nel nostro caso la somma ha 40 addendi); le proprietà metriche si conservano e una lingua ha auto-distanza strettamente positiva quando almeno uno dei 40 parametri che la descrivono è mal definito, ossia né 0 né 1.

3. LA DISTANZA DI LEVENŠTEJN

D'accordo, il nome yiddish di Vladimir Levenštejn è un po' impronunciabile, ma potete chiamarla anche *distanza di edit* oppure, se siete degli incorreggibili puristi, *distanza redazionale*; la sua importanza per gli utenti del web è fin troppo ovvia: se scrivete "distamzza" il sistema vi chiederà «Volevi dire "distanza"?» o magari vi correggerà "distamzza" in "distanza" senza neppure chiedervi niente. Ciò implica che "lui" si sia accorto che la stringa non registrata "distamzza" è vicina alla stringa pienamente legittima "distanza", la loro distanza è piccola (e scusateci il bisticcio di parole).

Occupiamoci proprio di parole o se preferite di stringhe costruite, per dire, sull'alfabeto delle 21 lettere italiane maiuscole, come TRIESTE o MONFALCONE di lunghezza rispettiva 7 e 10 (numero delle lettere che compongono la stringa).

L'idea delle distanze di edit è di servirsi di trasformazioni ognuna delle quali ha il proprio costo, e di usare la sequenza di trasformazioni a costo minimo che consente di passare dalla prima alla seconda stringa (il costo di più trasformazioni usate una dopo l'altra è la somma dei singoli costi): è per l'appunto questo costo additivo minimo che è la distanza di edit fra le due stringhe. Qualche esempio chiarirà le cose. Supponiamo di avere a disposizione due trasformazioni, il *twiddle* (scambio di due lettere

adiacenti) e la sostituzione di una lettera con un'altra, la prima trasformazione a costo unitario, la seconda a costo 2.

Trasformare CERTA in CETRA costa 1, mentre passare da CARO a CASO costa 2 - non ci sono possibilità più economiche. Passare da CARRO a CARO non si può, la distanza, se si vuole, è infinita. Se ci fossero anche gli inserimenti e le cancellazioni, diciamo a costo 1, la distanza fra CARRO e CARO sarebbe solo 1 (una singola cancellazione), come quella fra CARO e CARRO (un inserimento), ma se gli inserimenti costassero 2 e le cancellazioni soltanto 1 la simmetria si perderebbe. Se poi i *twiddle* fossero gratuiti la distanza fra le due stringhe, pur distinte, CERTA e CETRA scenderebbe a zero.

Abbiamo smontato tutti gli assiomi delle distanze metriche, tranne uno, la triangolarità ovviamente *ordinata* visto che non è più garantita la simmetria. Possiamo passare a costo minimo da x a z per poi arrivare sempre a costo minimo da z a y; questo è un modo legittimo di passare da x a y, ma non è affatto detto che sia il più economico, visto che stiamo ignorando i percorsi legittimi che non rispettano il vincolo di dover "toccare" per forza z. In altre parole: le trasformazioni e i loro costi possono essere capricciosi come vi pare, ma la disuguaglianza triangolare ordinata è sempre verificata dalle distanze di edit.

Ci corre l'obbligo di precisare che la distanza di edit di gran lunga più popolare consente tre trasformazioni, *sostituzione*, *cancellazione* e *inserimento*, tutte e tre a costo unitario, ed è perfettamente metrica. A differenza di quanto succedeva con i nostri esempietti, scovare il percorso minimo non è sempre facile "a occhio" ma per fortuna esiste uno splendido algoritmo basato sui principi della programmazione dinamica che risolve il problema a bassa complessità di calcolo: è per questo che "lui" è così bravo a calcolare le distanze in tempo reale, come vi siete ben accorti usando la rete o il vostro telefonino.

4. LE DUE DISTANZE DI LONGOBARDI

Finora non siamo riusciti a smentire la triangolarità, ma neanche questa resiste come mostrano numerosi casi trattati nell'*Encyclopedia* e come mostra la situazione che

segue, situazione che nell'*Encyclopedia* non appare ancora.

La scuola di linguisti tradizionali e computazionali diretta da Giuseppe Longobardi (York, UK) ha portato a risultati straordinari nella filogenesi delle lingue⁸. I metodi sono basati sulla *sintassi* piuttosto che sul *lessico*, come si fa di norma, in base all'osservazione che il "segnale" sintattico evolve nel tempo più lentamente del segnale lessicale: è questo che spiega l'uso del termine "preistoria" nel titolo del lavoro di CEOLIN *et al.*⁹.

Nel caso in questione¹⁰ le lingue trattate, tutte del Vecchio mondo – Europa, Asia e Africa – sono 58, classificate ognuna mediante 94 parametri sintattici. A differenza di quanto accadeva con Muljačić, qui il problema non è più la *fuzziness*, ma piuttosto il fatto che certi parametri sono totalmente non informativi, è come se non esistessero.

Per intenderci, anche se siamo costretti a semplificare e a chiedere scusa ai linguisti: se al parametro «In questa lingua esiste la declinazione dei sostantivi» abbiamo risposto no è inutile chiederci, come previsto da un parametro successivo, se genitivo e dativo possano avere desinenze distinte, come in tedesco, cui spetterebbe l'1, mentre in rumeno, cui spetterebbe lo 0, sono sempre coincidenti (i sostantivi tedeschi e rumeni si declinano)¹¹. Le stringhe associate a ciascuna lingua sono dunque ternarie, 0 per i parametri no, 1 per i parametri sì e una stella * per i parametri non informativi che di fatto in quella lingua sono inesistenti; nel caso appena accennato, del genitivo e del dativo, l'italiano, che non ha la declinazione dei sostantivi, avrebbe una stella.

Per giungere alle classificazioni delle lingue e cercare di capire se il giapponese e il coreano, per dire, abbiano qualche remoto antenato che i metodi tradizionali, lessicali, non riescono a intravedere, abbiamo bisogno di distanze fra stringhe ternarie.

Se non ci fossero le stelle e le stringhe fossero solo binarie potremmo usare la *distanza di Hamming* o la *distanza di Jaccard*. Nel caso di Hamming si contano le posizioni in cui le stringhe differiscono e si normalizza dividendo per la lunghezza comune delle stringhe

⁸ Si veda ad esempio CEOLIN *et al.* 2021.

⁹ Cfr. CEOLIN *et al.* 2021.

¹⁰ Cfr. CEOLIN *et al.* 2021.

¹¹ Ad esempio in romeno *fata, fetei* (la ragazza, della/alla ragazza), *fetele, fetelor* (le ragazze, delle/alle ragazze).

(la normalizzazione non si fa sempre, come invece faremo noi), nel caso di Jaccard si accorciano le stringhe eliminando le posizioni in cui figura lo zero sia in una stringa sia nell'altra e si calcola la *distanza normalizzata di Hamming* fra le stringhe così accorciate.

Ad esempio $d_H(001, 011) = \frac{1}{3}$, $d_J(001, 011) = \frac{1}{2}$: nel caso di Jaccard la prima posizione è come se non esistesse. Va subito detto che queste due distanze sono rigorosamente metriche, anche se nel caso di Jaccard non è così semplice dimostrarlo.

Purtroppo le nostre stringhe sono ternarie, per cui dobbiamo saper gestire anche le stelle modificando quanto abbiamo detto in modo da ottenere distanze di tipo ternario; chiameremo le due nuove distanze *pseudo-Hamming* e *pseudo-Jaccard*.

Quello che si fa, date le due stringhe, è eliminare le posizioni in cui compaiono le stelle, comprese le posizioni in cui una sola delle due lingue ha una stella, accorciare di conseguenza la lunghezza e finalmente calcolare la distanza normalizzata di Hamming o di Jaccard fra le stringhe così accorciate (che nel caso di Jaccard possono essere soggette a un ulteriore accorciamento).

Ad esempio con $x = 1^*0^*01$ e $y = 100^*10$ si ha $d_{PH}(x, y) = \frac{2}{4} < \frac{2}{3} = d_{PJ}(x, y)$.

Purtroppo entrambe le pseudo-distanze possono violare la disuguaglianza triangolare; si pensi alle tre stringhe “astratte” di lunghezza 21:

$x = 100000000001111111111,$

$y = 111111111110000000000,$

$z = 1^*****.$

Le stringhe x e z sono diverse, ma $d(x,z) = 0$; come se non bastasse

$$d(x,z) + d(z,y) = 0 + 0 < d(x,y) = \frac{20}{21},$$

e dunque quasi 1, che è il valore più grande che una distanza normalizzata possa assumere. Non abbiamo specificato il tipo di distanza perché in questo caso, in mancanza di posizioni in cui entrambe le stringhe abbiano uno 0, pseudo-Hamming e pseudo-Jaccard coincidono.

Il nostro antipatico esempio riguarda stringhe “astratte”: che cosa succede con le 58 lingue “concrete” e con le loro stringhe di lunghezza 94?

Ci sono sì lingue a distanza 0, come l’irlandese e il gallese, ma questo solo perché le loro stringhe ternarie sono identiche (si rammenti che stiamo usando distanze fra stringhe che solo indirettamente sono distanze fra lingue). In quanto alla disuguaglianza triangolare le triple di lingue x , y e z che vanno controllate sono 92568, come mostra il calcolo combinatorio (basta contare il numero delle coppie non ordinate di elementi x, y distinti e moltiplicare per il numero degli elementi z che rimangono).

Con la distanza pseudo-Hamming le triple difettose sono 993, vale a dire lo $0.0107 \approx 1\%$ del numero totale. Le triple “peggiori”, quelle con la differenza $d_{pH}(x,y) - d_{pH}(x,z) - d_{pH}(z,y)$ massima e pari a 0.1, sono (SerboCroato, Hindi, Mandarino), (SerboCroato, Hindi, Cantonese), (Sloveno, Hindi, Mandarino), (Sloveno, Hindi, Cantonese), (Polacco, Hindi, Mandarino), (Polacco, Hindi, Cantonese), (Russo, Hindi, Mandarino), (Russo, Hindi, Cantonese).

Con la distanza pseudo-Jaccard la situazione è migliore. Ci sono 143 triple difettose pari allo $0.015... = 0.15\%$ del numero totale. Lo scarto massimo di 0.089 corrisponde a (Hindi, Tamil, Mandarino), (Hindi, Tamil, Cantonese), (Hindi, Telugu, Mandarino), (Hindi, Telugu, Cantonese)¹².

Ai linguisti poco interessa la mancata metricità: ecco piuttosto, nella Figura 2, l’albero filogenetico cui sono giunti. Fra le parentele che suggerisce c’è quella fra il giapponese e il coreano, con un progenitore comune che affonda nel tempo e che è “invisibile” ai tradizionali strumenti della filogenesi lessicale.

5. CONCLUSIONI

Dati gli interessi e le preferenze dei due autori, non sorprende che gli esempi trattati rivestano tutti carattere linguistico, ma le distanze dell’*Encyclopedia*, metriche o non metriche che siano, riguardano situazioni prese dalla geometria, dall’algebra, dalla teoria dei numeri, dall’analisi funzionale, dalla statistica e dal calcolo delle probabilità,

¹² Per maggiori dettagli si veda DINU *et. al.* 2023.

dalla teoria dei grafi e dalla combinatorica, dall'ingegneria matematica e dall'informatica (immagini, audio, internet), dalla biologia e dalla fisica, dalla chimica e dalla medicina, dalla geografia, dalla geofisica e dall'astronomia fino alla cosmologia, alla teoria della relatività e perfino alla criminologia: se c'è un concetto pervasivo, dunque, è quello di *distanza*, e neppure la gabbia della metricità, pur ampia e generosa, riesce a contenerlo.

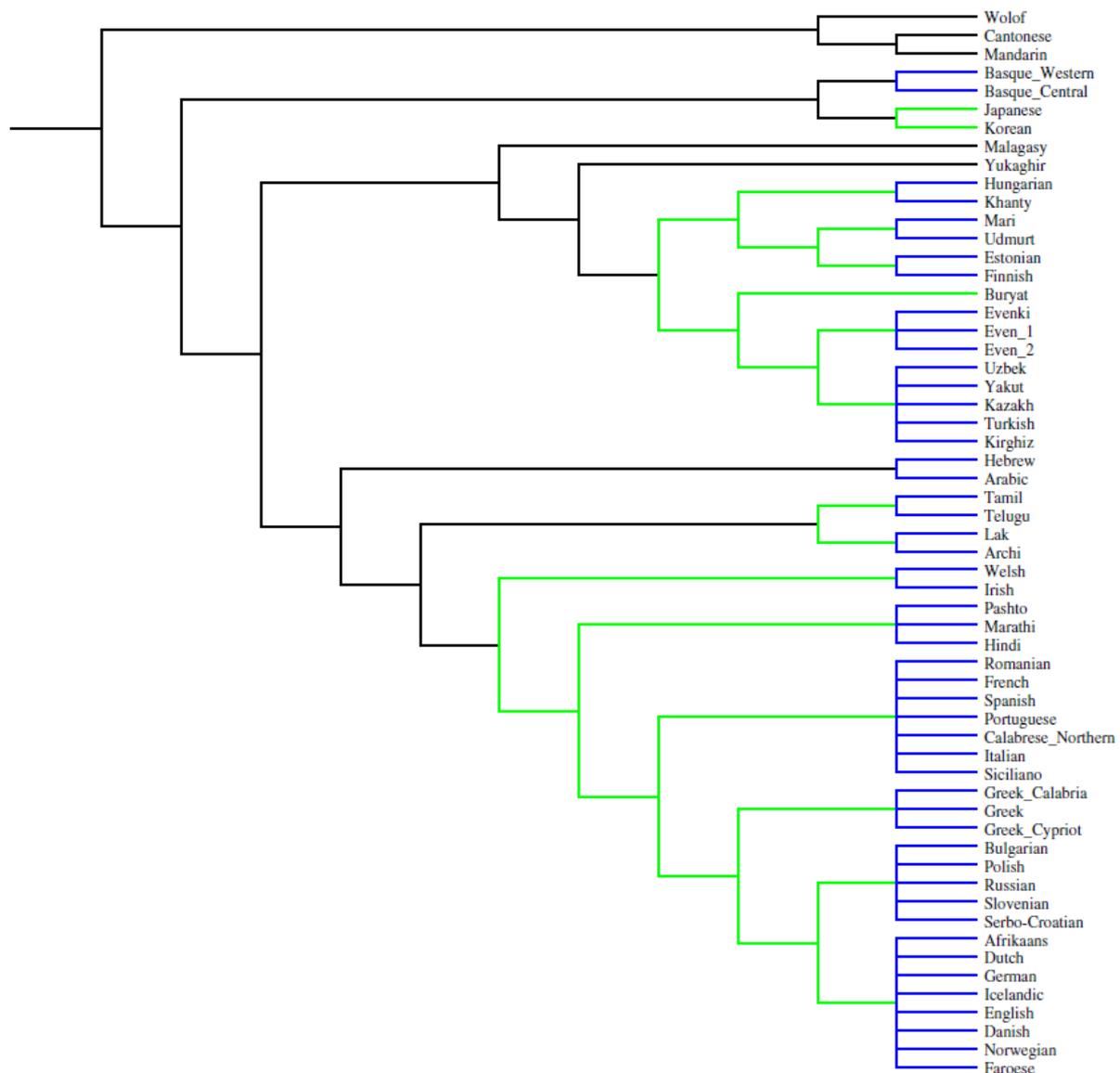


Figura 2. L'albero filogenetico basato sulla sintassi (Fonte: CEOLIN *et al.* 2021).

BIBLIOGRAFIA

CEOLIN A., GUARDIANO C., LONGOBARDI G., IRIMIA M. A., BORTOLUSSI L., SGARRO A. 2021, «At the boundaries of syntactic prehistory», *Philosophical Transactions of the Royal Society B*, 376, pp. 1-10.

DEZA M. M., DEZA E.

2016, *Encyclopedia of distances*, Springer.

DINU A., DINU L., FRANZOI L., SGARRO A.

2023, *Distances in syntax-based linguistic phylogeny*, to be submitted to a computational linguistic forum.

FRANZOI L., SGARRO A.

2017, *Fuzzy Hamming distinguishability*, Napoli, FUZZ-IEEE, pp. 1-6.

MULJAČIĆ Ž.

1967, «Die Klassifikation der romanischen Sprachen», *Romanistisches Jahrbuch*, 18, pp. 23-37.

SGARRO A.

2019, «Come contraddirsi rimanendo coerenti: il caso della logica fuzzy», *QuaderniCIRD*, 19, pp. 37-45.

SGARRO A., FRANZOI L.

2021, «Sillogismi sfocati: il modus ponens», *QuaderniCIRD*, 22, pp. 50-62.

Sophie Germain e l'Ultimo Teorema di Fermat*

ELISABETTA PODDA
Trieste
elisabetta.podda99@gmail.com

ALESSANDRO LOGAR
Dipartimento di Matematica, Informatica e Geoscienze
Università di Trieste
logar@units.it

ABSTRACT

The purpose of this paper is to present the figure of Sophie Germain, setting her in the historical period in which she lived, and to try to explain some of the mathematical tools she used in her attempt to solve one of the most famous theorems in the history of mathematics: Fermat's Last Theorem. Though the French mathematician gave some important results concerning Fermat's problem, historically she has not been given credit for what she proved, until recent studies have re-evaluated her works. In the first and second section we present Sophie Germain's biography and introduce Fermat's Last Theorem from the algebraic point of view. The following sections are dedicated to the presentation of the most important results by Sophie Germain and the techniques she uses to prove them. Finally, we briefly discuss the failure of her plan, the so called "Grand Plan", to prove Fermat's conjecture.

PAROLE CHIAVE

MATEMATICA / MATHEMATICS; STORIA DELLA MATEMATICA / HISTORY OF MATHEMATICS; SOPHIE GERMAIN; ULTIMO TEOREMA DI FERMAT / FERMAT'S LAST THEOREM; TEORIA DEI NUMERI / NUMBER THEORY; PRIMI DI GERMAIN / GERMAIN'S PRIMES.

1. INTRODUZIONE

Sophie Germain è stata una matematica vissuta in Francia a cavallo tra il diciottesimo e il diciannovesimo secolo. La sua figura è rimasta poco nota fino agli anni recenti, quando alcuni storici della matematica le hanno dedicato varie ricerche e approfondimenti. È così risultato che la Germain, per l'originalità dei temi studiati e

* Title: Sophie Germain and Fermat's Last Theorem.

delle tecniche introdotte, è meritevole di occupare un posto di rilievo tra i matematici del XIX secolo.

Nella prima parte di questo articolo viene presentata la biografia di Sophie Germain, in quanto conoscere il contesto storico in cui è vissuta e i matematici dell'epoca con cui è stata in contatto, aiuta a comprendere come questi fattori abbiano influenzato la sua carriera. Si può subito affermare che molto di quanto si conosce della sua vita è arrivato a noi grazie al suo amico e collega Guglielmo Libri che, dopo la morte della Germain nel 1831, pubblicò nel 1832 un necrologio dal titolo *Notices sur M.lle Sophie Germain*¹. Inoltre, il lavoro *Sophie Germain: Un articolo nella storia della teoria dell'elasticità* di L. Bucciarelli e N. Dworsky, fornisce una visione completa della sua vita personale e professionale, analizzando poi in dettaglio i suoi contributi alla fisica matematica. Successivamente, si passa ad analizzare l'intenso lavoro di ricerca svolto dalla Germain, nel corso di molti anni, nel tentativo di dimostrare uno dei teoremi più famosi della storia della matematica, il cosiddetto *Ultimo Teorema di Fermat*.² Sebbene i suoi sforzi si siano dimostrati vani, l'originalità dei suoi risultati parziali e le tecniche introdotte sono senz'altro di grande rilievo.

Si presentano quindi i principali teoremi trovati da Sophie Germain e si cerca di esemplificare alcuni aspetti dimostrativi per mostrare come, pur avendo a disposizione pochi strumenti matematici, con raffinati ragionamenti, la Germain sia riuscita a ottenere delle conclusioni che sono ancora del tutto degne di nota.

Il contributo persegue quindi il duplice scopo di introdurre una figura importante della storia della matematica, ancora poco conosciuta, e spiegare brevemente alcuni degli strumenti da lei usati.

Per quanto riguarda la trattazione degli aspetti più prettamente matematici, si è cercato di non addentrarsi troppo in dettagli tecnici, ma di presentare l'argomento in modo che

¹ Cfr. LIBRI 1832.

² Di questo teorema si parlerà più diffusamente nella Sezione 3; si tratta di una congettura che il matematico francese Pierre de Fermat formulò nel 1637, senza darne una dimostrazione. La dimostrazione fu ottenuta dopo tre secoli e mezzo, nel 1994, dal matematico britannico Andrew Wiles.

anche chi non è esperto del settore sia comunque in grado di afferrare i punti salienti dei ragionamenti sviluppati dalla studiosa.

2. LA BIOGRAFIA DI SOPHIE GERMAIN



Figura 1. Sophie Germain.
(Fonte: Sophie Germain | Torino Scienza)

Sophie Germain nacque a Parigi il primo aprile del 1776 in una famiglia benestante. Il padre era un commerciante di seta e aveva un incarico politico come deputato dell'Assemblea Nazionale. Gli anni della sua giovinezza furono segnati dalla Rivoluzione francese e dal Regime del Terrore per cui, essendo costretta a rimanere in casa, si appassionò alla lettura come attività per trascorrere il tempo e distrarsi.

All'età di tredici anni lesse e studiò per conto proprio la monografia *Storia della matematica* del matematico francese Jean-Étienne Montucla. In particolare fu colpita dal racconto dell'episodio della morte di Archimede: secondo la leggenda, il giorno della caduta di Siracusa, attorno al 200 a. C., un soldato romano si sarebbe presentato per uccidere Archimede a colpi di spada e quest'ultimo lo avrebbe pregato invano di lasciargli terminare la dimostrazione nella quale era impegnato³.

La lettura del volume le fece scoprire quanto la matematica fosse una disciplina affascinante e, da quel momento in poi, la giovane Germain fu completamente immersa nell'approfondimento della materia, incominciando a studiare alcuni testi di

³ Cfr. O'CONNOR, ROBERTSON in Siti web.

Bézout che aveva trovato nella biblioteca del padre. Imparò anche il latino per poter comprendere le opere di Newton e di Eulero. Essendo una giovane donna, per lei all'epoca non era prevista un'educazione di tipo matematico; perciò, si ritrovò a studiare da autodidatta, senza l'aiuto di un insegnante.

Inizialmente la famiglia tentò in tutti i modi di dissuaderla, in quanto il suo interesse per la matematica era considerato inusuale e poco consono a una ragazzina della sua età: per impedirle di leggere i libri durante la notte, le toglievano perfino le candele e il fuoco dalla biblioteca. Successivamente però, vedendo la sua grande passione e dedizione, i genitori saggiamente non la ostacolarono più e la lasciarono libera di seguire la sua vocazione.

All'età di diciotto anni, la Germain voleva ardentemente ricevere un'adeguata istruzione matematica, senza dover più studiare da sola. Il suo desiderio era quello di iscriversi all'*École Polytechnique* di Parigi, inaugurata proprio in quell'anno, che però permetteva soltanto agli studenti maschi di partecipare alle lezioni.

Per ovviare a questo problema Sophie Germain si fece venire in mente uno stratagemma: assunse l'identità di uno studente del Politecnico di nome Antoine-August LeBlanc che aveva da poco abbandonato gli studi e così riuscì a procurarsi le lezioni dei vari corsi, non potendo però mai parteciparvi di persona.

Uno dei corsi che seguiva con più interesse era il corso di analisi matematica tenuto dal ben noto Joseph-Louis Lagrange, il quale di tanto in tanto chiedeva ai suoi studenti di consegnare le loro osservazioni e commenti sugli argomenti del corso e anche alcuni esercizi lasciati per casa, per poi discuterne insieme⁴. Quando gli capitò di leggere il foglio consegnato dalla Germain sotto il nome di LeBlanc, fu positivamente sorpreso dal fatto che questo studente, che non era fino ad allora sembrato particolarmente incline alla matematica, scrivesse delle osservazioni così acute e interessanti, mostrando un grande talento.

Lagrange fu quindi spinto dalla curiosità di conoscerlo e chiese di poterlo incontrare: appena scoprì la vera identità di LeBlanc, mostrò grande stupore e le fece molti

⁴ Cfr. PETROSINO in Siti web.

commenti lusinghieri per le sue doti. Per Lagrange l'aspetto più sorprendente fu il venire a conoscenza del fatto che la Germain si fosse auto-istruita nel corso degli anni, fino a raggiungere un livello di conoscenze matematiche pari a quello degli insegnamenti dei primi anni di università.

Col passare degli anni, la Germain divenne sempre più interessata alla *teoria dei numeri*, studiò la monografia *Teoria dei numeri* di Adrien-Marie Legendre del 1798 nonché le *Disquisitiones Arithmeticae*⁵ che Carl Friedrich Gauss aveva scritto all'età di ventun anni e che erano state pubblicate nel 1801.

Lagrange e Legendre divennero i suoi due mentori principali, la guidarono e incoraggiarono durante gli anni di studio all'Università, percorso non facile e non privo di ostacoli per una giovane ragazza dell'epoca. La comparsa di una giovane donna nei circoli intellettuali parigini destò molta curiosità: da una parte alcuni, come Lagrange e Legendre, erano ammirati delle sue capacità e della sua bravura, altri invece, come Poisson, la sminuivano e la facevano sentire intellettualmente inferiore.

La Germain, per un paio di anni, si dedicò allo studio dettagliato del volume *Disquisitiones Arithmeticae*, cercando di risolvere gli esercizi che erano proposti e dando dimostrazioni personali ai teoremi contenuti al suo interno.

Nel 1804 prese l'audace decisione di scrivere una lettera a Gauss, firmandosi con lo pseudonimo maschile di LeBlanc, ancora una volta per timore di essere rifiutata e non considerata. Nella lettera la Germain condivise con Gauss alcuni risultati di *teoria dei numeri* che aveva trovato e alcune idee e considerazioni sull'*Ultimo Teorema di Fermat*. Incominciò così una lunga corrispondenza tra la Germain e Gauss, basata sulla stima e sul supporto reciproco: si contano in tutto 14 lettere, 10 lettere inviate dalla Germain a Gauss tra il 1804 e il 1808, a cui seguiranno altre due datate 1819 e 1829. Gauss rispose con interesse ed entusiasmo alla prima lettera, tant'è che scrivendo al suo amico H. W. Olbers (un medico tedesco) disse:

⁵ Cfr. GAUSS 1801.

*Ho ricevuto con piacere una lettera da un giovane geometra di Parigi, LeBlanc, che ha studiato l'algebra superiore con grande entusiasmo e mi ha fornito prova di esser penetrato nei meandri più profondi delle mie Disquisitiones Arithmeticae.*⁶

Questa parte significativa della lettera che Gauss mandò a Olbers è riportata anche nel volume *Wilhelm Olbers, sein Leben und seine Werke* di Schilling⁷.

Dopo tre anni di corrispondenza sotto falso nome, Gauss scoprì la vera identità di LeBlanc a seguito di alcuni eventi della guerra franco-prussiana allora in corso. Nell'ottobre del 1806 la vittoria dell'esercito di Napoleone a Jena aprì la strada all'invasione di gran parte dei territori prussiani. Appena la Germain venne a sapere che le truppe di Napoleone erano in procinto di invadere Brunswick, la città natale di Gauss, ricordandosi della vicenda dalla morte di Archimede, fu presa dalla preoccupazione per l'incolumità del suo mentore.

A tal proposito allora Sophie Germain chiamò il generale J. M. Pernety⁸, che era un amico di famiglia e comandante dell'artiglieria francese, e gli chiese di andare a trovare Gauss e garantirgli la dovuta sicurezza. Per svolgere la missione che la Germain gli aveva affidato, il generale Pernety incaricò uno dei suoi comandanti di battaglione. L'incontro tra Gauss e il comandante è raccontato nella lettera che quest'ultimo scrisse a Pernety una volta conclusa la sua missione. In una parte della lettera, riportata nella biografia di Sophie Germain di Bucciarelli e Dworsky, è scritto:

*[...] Monsieur Gauss disse di non aver l'onore di conoscere né voi né Mademoiselle Germain... Dopo aver parlato dei diversi punti contenuti nel vostro ordine, è sembrato un po' confuso e mi ha chiesto di trasmettergli i suoi ringraziamenti per la vostra considerazione a suo nome.*⁹

In questa circostanza la Germain si sentì quindi in obbligo di rivelare a Gauss la sua vera identità.

*Non sono così completamente sconosciuto a voi come potreste credere, ma temendo il ridicolo collegamento a una scienziata, ho precedentemente preso il nome di M. LeBlanc nel comunicare con voi. [...] Spero che le informazioni che vi ho confidato oggi non mi privino dell'onore che mi avete accordato sotto un nome preso in prestito.*¹⁰

⁶ Cfr. LAUBENBACHER, PEGELLEY 2010.

⁷ Cfr. SCHILLING 1900.

⁸ J. M. Pernety fu un generale di artiglieria che prese parte a molte campagne militari importanti come le guerre rivoluzionarie francesi e le successive guerre napoleoniche, oltre a prestare servizio per alcuni periodi come membro dell'Armata d'Italia e dell'esercito inglese. Si distinse per il suo coraggio e talento, tant'è che il suo nome compare tra quelli scritti in uno dei pilastri sotto l'Arco di Trionfo a Parigi.

⁹ Cfr. BUCCIARELLI, DWORSKY 1998.

¹⁰ Vedi GIORDANO in Siti Web.

Scrivendo queste parole, Germain temeva l'ira di Gauss per averlo ingannato, ma sorprendentemente, a seguito di questa rivelazione, Gauss rimase positivamente meravigliato e al contempo la elogiò per la sua determinazione e il suo talento, e la incoraggiò a continuare l'attività di ricerca matematica. In proposito, nel 1807, sempre in una lettera a Olbers, Gauss scrisse:

*Di recente il mio testo Disquisitiones Arithmeticae mi ha recato molte sorprese. Come già sapete, la corrispondenza con LeBlanc da Parigi ha evidenziato quanto LeBlanc padroneggi completamente tutti gli argomenti trattati nell'opera. LeBlanc si è da poco rivelato essere un nome fittizio di una giovane donna, Sophie Germain: sicuramente questo fatto vi meraviglierà tanto quanto ha meravigliato me.*¹¹

Inoltre, H. M. Edwards nel suo libro *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, riporta quello che è un vero e proprio elogio di Gauss a Sophie Germain, scritto subito dopo aver scoperto la sua vera identità. In questa lettera Gauss si dice ammirato dalla facilità e semplicità con cui la Germain comprendeva e analizzava i vari aspetti dell'aritmetica, dalla saggezza con la quale generalizzava e perfezionava i risultati ottenuti e dal coraggio e dalla perseveranza dimostrati nell'occuparsi di matematica. Gauss esprime la sua ammirazione con queste parole:

*Ma come posso descrivere il mio stupore e la mia meraviglia nel vedere il mio apprezzato corrispondente M. LeBlanc trasformarsi in questo illustre personaggio Sophie Germain che dà un brillante esempio di quello che difficilmente riuscirei a credere? Avere una passione per le scienze astratte in generale e soprattutto per i numeri è estremamente raro: non è una materia che colpisce tutti; l'incantevole fascino di questa scienza sublime rivela se stessa solo a coloro che hanno il coraggio di penetrarla. Una donna, a causa del suo sesso e dei nostri pregiudizi, incontra molti più ostacoli di un uomo nel familiarizzarsi con problemi complessi. Tuttavia, quando supera queste barriere e penetra nelle profondità più recondite, rivela di possedere il coraggio più nobile, un talento straordinario e un genio superiore. [...] Gli appunti di cui le vostre lettere sono piene mi hanno dato moltissime gioie. Li ho studiati con attenzione e ammiro la facilità con cui penetrate tutti i rami dell'aritmetica e la saggezza con cui generalizzate e perfezionate i risultati.*¹²

Questo atteggiamento lodevole, era alquanto insolito per l'epoca, vista la poca considerazione che gli uomini avevano nei confronti delle capacità femminili in matematica e nella scienza in generale. Va detto però che, nonostante queste buone conoscenze e le referenze di cui disponeva, la Germain lavorò in quasi totale isolamento, non potendo partecipare agli incontri all'Accademia e non avendo accesso

¹¹ Cfr. LAUBENBACHER, PEGELLEY 2010.

¹² Cfr. EDWARDS 1977.

al materiale scientifico che veniva pubblicato in quegli anni. Questo fece sì che la sua figura di matematica rimanesse per molto tempo nascosta e poco considerata.

Nella stessa lettera, Gauss ringrazia la Germain per il pensiero e l'aiuto datogli durante l'invasione di Brunswick, rassicurandola del fatto che fino a quel momento la guerra non lo aveva coinvolto direttamente.

Dal 1808 fino al 1819 il carteggio tra Germain e Gauss si interruppe a seguito della nomina di Gauss a direttore dell'Osservatorio astronomico di Gottinga. Questo nuovo incarico lo costrinse a mettere da parte tutti i suoi impegni legati alla matematica ed egli non riuscì più a trovare il tempo per rispondere alle lettere della Germain. Nel corso del 1808 Sophie Germain gli scrisse altre tre lettere, ma, non ricevendo risposte, smise di scrivergli nel 1809.

A questo punto, priva del supporto di Gauss, Germain mise temporaneamente da parte il suo interesse per la *teoria dei numeri* e decise di dedicarsi ad altri campi della matematica. Proprio nel 1808 il fisico tedesco E. F. Chladni aveva organizzato a Parigi delle esibizioni di esperimenti riguardo al complesso fenomeno delle *vibrazioni* di sottili lastre. L'Accademia delle Scienze di Parigi ne fu incuriosita e mise in palio un premio per chi fosse riuscito a dare delle spiegazioni di tipo matematico al fenomeno. La Germain, nel corso degli anni successivi, si mise a studiare questo problema con molta dedizione ed elaborò tutta una sua *teoria sulla vibrazione* di membrane elastiche, grazie anche all'aiuto e al supporto di Legendre. Nel 1811 consegnò all'Accademia delle Scienze di Parigi, come unica partecipante, il lavoro da lei svolto. Sfortunatamente il lavoro consegnato, basato su una generalizzazione della *teoria di Eulero* sulla vibrazione di travi, era contrassegnato da un errore. L'Accademia decise quindi di prolungare la competizione per altri due anni.

Approfittando di questa opportunità, nel 1813 Sophie Germain presentò una versione modificata e migliorata del suo primo lavoro. Anche questa volta la strada seguita dalla Germain per ottenere l'equazione fondamentale fu giudicata parzialmente scorretta. La competizione fu ulteriormente prolungata per due anni, alla fine dei

quali Germain consegnò, di nuovo come unica partecipante, il suo terzo lavoro in cui studiava anche le *vibrazioni di superfici curve*. La commissione composta da Legendre, Poisson e Laplace, valutando il lavoro della Germain come interessante ma allo stesso tempo incompleto e lacunoso, decise di conferirle il premio con riserva nel 1816.

L'Accademia prese la decisione di non pubblicare il suo elaborato, perciò, alcuni anni dopo, la Germain lo pubblicò a sue spese. Ad oggi questo testo è considerato il suo contributo più importante alla fisica matematica, un lavoro ricco di brillanti intuizioni che getta le fondamenta della moderna *teoria dell'elasticità*. Il premio conferitole dimostra che Sophie Germain non è soltanto un personaggio importante nella storia della matematica per i contributi dati alla *teoria dei numeri*, ma che i suoi studi sono stati di ben più ampio spettro, fino ad occuparsi di fisica matematica e di scienza in generale. Bucciarelli e Dworsky¹³ riferiscono che nel 1826 la Germain presentò all'Accademia un nuovo studio sull'elasticità, che essa considerava una versione più chiara rispetto all'ultimo lavoro consegnato. Cauchy, che aveva l'incarico di leggere e valutare il lavoro, dopo averlo analizzato, incoraggiò fortemente la Germain a pubblicarlo: lo studio quindi apparve nel 1828 sulla rivista *Annales de Chimie*.

Nel 1816, l'Accademia delle Scienze mise in palio un altro premio, questa volta per chi fosse riuscito a dimostrare quello che è noto come *Ultimo Teorema di Fermat*, anche se all'epoca l'affermazione di Fermat era solo una congettura tutta ancora da verificare.

La competizione fu estesa al 1818 e poi ritirata nel 1820. Sophie Germain, da parte sua, non presentò mai una soluzione all'Accademia per concorrere al premio, ma grazie all'autobiografia di Legendre¹⁴ pubblicata nel 1823 e alla documentazione trovata dopo la sua morte, sappiamo che la Germain lavorò a lungo a questo problema. Di certo il fatto che i suoi lavori e manoscritti non fossero stati all'epoca pubblicati è uno dei motivi principali per cui tutti i suoi contributi all'*Ultimo Teorema di Fermat* sono rimasti a lungo sconosciuti e solo di recente sono stati riscoperti e rivalutati.

¹³ Cfr. BUCCIARELLI, DWORSKY 1998.

¹⁴ Cfr. LEGENDRE 1832.

Probabilmente la Germain riprese a studiare l'Ultimo Teorema di Fermat solo nel 1818, molto tempo dopo aver scritto di questo argomento nel 1804 nella sua prima lettera a Gauss, in cui si diceva convinta del fatto che la *teoria delle congruenze modulari e dei residui* era lo strumento giusto per affrontare il problema. Sappiamo inoltre che mise a punto un piano completo, da lei denominato “*Grand Plan*”, con il fine di dimostrare l'Ultimo Teorema di Fermat nella sua totalità.

Legendre, nella sua autobiografia attribuisce, in maniera riduttiva, alla Germain soltanto il teorema principale per cui è maggiormente ricordata, che oggi prende il nome di *Teorema di Sophie Germain*, ma dall'analisi dei suoi lavori si viene a conoscenza di molti altri risultati a lei dovuti. Da quando la Germain riprese le ricerche nell'ambito della *teoria dei numeri*, mise a punto vari risultati riguardanti la dimostrazione dell'Ultimo Teorema di Fermat in alcuni casi particolari.

Dopo undici anni di pausa, nel 1819, la visita di un amico di Gauss, un certo H. C. Schumacher, diede alla Germain l'occasione di mandare a Gauss una nuova lettera, mettendolo al corrente in particolare dei suoi studi circa l'Ultimo Teorema di Fermat¹⁵.

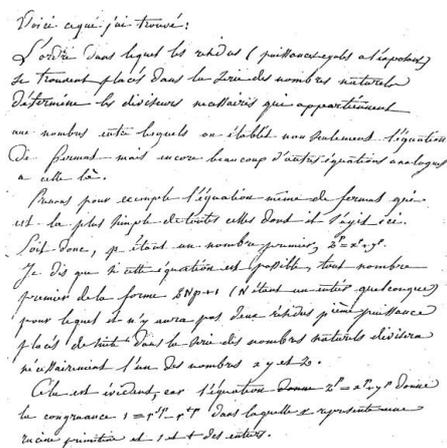


Figura 2. Lettera del 12 maggio 1819 a Gauss.
(Fonte: <https://amslaurea.unibo.it>)

Nella lettera datata 12 maggio 1819, dove esordisce dicendo «Ecco cosa ho trovato», la Germain espone a Gauss tutte le ricerche che aveva condotto negli ultimi anni sia

¹⁵ Cfr. DEL CENTINA 2007.

in ambito matematico sia fisico e spiega dettagliatamente il suo programma per arrivare alla dimostrazione completa dell'*Ultimo Teorema di Fermat*.

Nella lettera scrive:

Sebbene abbia lavorato per qualche tempo alla teoria dell'elasticità, non ho mai cessato di pensare alla teoria dei numeri. Vi darò un'idea del mio coinvolgimento nelle ricerche in questo campo della matematica ammettendo che, anche senza speranze di successo, la preferisco ad altri lavori che potrebbero comunque interessarmi e darmi dei risultati certi. Molto prima che la nostra Accademia proponesse un premio per la dimostrazione dell'impossibilità dell'equazione di Fermat, questo tipo di sfida, che fu portata alle moderne teorie da un geometra che non possedeva le risorse di cui disponiamo noi oggi, mi ha spesso tormentato. Ho intravisto vagamente una connessione tra la teoria dei residui e la famosa equazione. Penso di aver già parlato con voi di questa mia idea molto tempo fa, perché mi ha colpito appena ho letto il vostro libro.¹⁶

Nel 1829 Gauss ricevette quella che sarebbe stata l'ultima lettera da parte della Germain, la quale a causa della malattia che la colpì riuscì a lavorare sempre meno. Le parole con cui la Germain chiuse questa lettera esprimono tutto il suo rimpianto per non aver collaborato a sufficienza con Gauss: per sviluppare tutto il suo potenziale nella teoria dei numeri avrebbe avuto sicuramente bisogno nel corso degli anni di molto più supporto e interazione da parte sua. Del Centina nel suo articolo *The correspondence between Sophie Germain and Carl Friedrich Gauss* riporta il seguente frammento della lettera:

Rimpiango di essere stata privata del vantaggio che avrei trovato nel partecipare a delle conversazioni con voi, come fa Mr Bader. Quello che mi racconta non è sorprendente per me, ma è oggetto di invidia. Oltre a quello che potrei imparare da voi, mi rammarico di non aver potuto sottoporre al vostro giudizio molte idee che non ho pubblicato e che sarebbero troppo lunghe da spiegare in una lettera.¹⁷

Nel giugno del 1831, proprio quando l'Università di Gottinga aveva deciso di conferirle il titolo di dottore onorario per il lavoro svolto nel corso di molti anni, Sophie Germain morì a causa di un tumore al seno.

Per evidenziare ancora una volta il poco riconoscimento ottenuto per i suoi studi, è interessante vedere che sul suo certificato di morte figurò come “redditiera” e non come “matematica”. Quello che possiamo dire complessivamente della sua vita è che, nonostante le difficoltà incontrate nel suo percorso formativo e lavorativo in quanto

¹⁶ Cfr. LAUBENBACHER, PENGELLEY 2010.

¹⁷ Cfr. DEL CENTINA, FIOCCA 2012.

donna, il suo contributo ha permesso di sistemare un tassello importante nella storia della dimostrazione della congettura di Fermat.

Inoltre, dall'analisi dei suoi scritti concludiamo che il suo nome dovrebbe essere incluso tra quelli dei più importanti studiosi di *teoria dei numeri* della sua epoca, come Gauss stesso affermò nel 1837, quando disse che Sophie Germain avrebbe meritato una laurea onoraria se fosse stata ancora viva, in quanto aveva realizzato qualcosa di utile nella più rigorosa e astratta delle scienze.

3. L'ULTIMO TEOREMA DI FERMAT

Quello che è passato alla storia come *Ultimo Teorema di Fermat* è in realtà una congettura che Fermat aveva enunciato a seguito dei suoi studi ma di cui non fornì mai una dimostrazione scritta. L'enunciato del teorema è apparentemente molto semplice, ma dimostrarlo è stato una sfida, durata secoli, per moltissimi matematici di altissimo livello. Pierre de Fermat è stato un magistrato francese del Diciassettesimo secolo che si occupava anche di matematica. Il campo della matematica in cui fu più attivo è indubbiamente la *teoria dei numeri*, sebbene abbia dato importanti contributi al *calcolo differenziale*, alla *geometria analitica* e al *calcolo delle probabilità*. Quando morì, nel 1665, era considerato uno dei più famosi matematici del suo tempo a livello europeo.

Nel 1637, leggendo l'*Arithmetica* del matematico greco Diofanto del terzo secolo d. C., che tratta di *equazioni diofantee*, scrisse la famosa (o addirittura famosissima) annotazione nel margine della pagina in cui era riportato il teorema di Pitagora:

*È impossibile dividere un cubo in altri due cubi, una quarta potenza o in generale una potenza qualsiasi in due potenze dello stesso valore maggiore del secondo. Dispongo di una meravigliosa dimostrazione di questo teorema che non può essere contenuta nel margine troppo stretto della pagina.*¹⁸

Fermat quindi, non avendo abbastanza spazio nella pagina, non rese mai nota la dimostrazione che diceva di aver trovato. Com'è ben noto, in termini algebrici il teorema si può esprimere in questa forma:

¹⁸ Cfr. HOWARD 1964.

Teorema. Dato n numero naturale non nullo, l'equazione $x^n + y^n = z^n$ non ammette soluzioni intere positive se $n > 2$.

Seguendo il testo di Diofanto, l'obiettivo di Fermat era quello di trovare soluzioni razionali dell'equazione $x^n + y^n = z^n$. Infatti, se abbiamo una soluzione (x, y, z) razionale dell'equazione, da essa si ricava facilmente una soluzione intera moltiplicando x, y e z per il minimo comune multiplo d dei denominatori.

Si escludono le soluzioni irrazionali in quanto si osserva che l'equazione ne ammette infinite, infatti, scegliendo x, y positivi, basta calcolare $z = \sqrt[n]{x^n + y^n}$.

Inoltre, si cercavano soluzioni positive in quanto, nel Diciassettesimo secolo, i numeri negativi e lo zero venivano ancora visti con sospetto e non erano considerati come possibili soluzioni. Perciò le terne del tipo $(0, y, y)$, $(z, 0, z)$ e la terna nulla $(0,0,0)$, pur soddisfacendo l'equazione, erano escluse dall'insieme delle soluzioni.

Da queste osservazioni segue che, come riportato nell'enunciato del teorema, il problema si riconduce alla ricerca di soluzioni intere (e positive).

Notiamo che, quando $n = 1$ o $n = 2$, l'equazione $x^n + y^n = z^n$ ammette invece soluzioni intere positive. Se $n = 1$, l'equazione diventa $x + y = z$ e ha infinite soluzioni intere positive che sono rappresentate nello spazio dai punti a coordinate intere positive del piano $\pi: x + y - z = 0$. Se invece $n = 2$, si ottiene l'equazione $x^2 + y^2 = z^2$, che è proprio l'equazione che nasce dalla relazione tra i cateti e l'ipotenusa espressa dal *Teorema di Pitagora*. Per questo motivo le soluzioni intere dell'equazione sono dette *terne pitagoriche*.

La *terna pitagorica* più conosciuta è senz'altro $(3,4,5)$, sicuramente nota agli Egizi attorno al 2000 a. C., ma alcune tavolette Babilonesi la riportano già in una lista assieme ad altre *terne pitagoriche* anche complesse come $(4961, 6480, 8161)$. È facile da vedere che se (a, b, c) è una *terna pitagorica*, allora lo è anche $(\lambda a, \lambda b, \lambda c)$ con λ numero naturale non nullo. In particolare, le *terne pitagoriche* sono infinite.

I primi risultati relativi alla prova del *Teorema di Fermat* risalgono a Fermat stesso, che fornì una dimostrazione del caso $n = 4$. Eulero, nel 1760, dimostrò (anche se con un

errore risultato poi correggibile) il caso $n = 3$ (dimostrazione poi pubblicata nel 1770 nel suo testo *Algebra*) e Legendre e Dirichlet, entrambi nel 1825, provarono il caso $n = 5$. Dal punto di vista storico, ci sono opinioni contrastanti circa la validità della dimostrazione di Eulero: sebbene ad oggi la maggior parte degli studiosi sostenga che egli ha effettivamente provato il teorema nel caso $n = 3$, è necessario mettere in evidenza l'incompletezza della dimostrazione a causa di un errore che apparentemente non era stato notato.

Per dimostrare il caso $n = 4$, Fermat introdusse una tecnica che si è dimostrata efficace nel provare molti altri problemi in teoria dei numeri: il *metodo della discesa infinita*. La tecnica consiste nell'assumere (per assurdo) di avere una soluzione, data da una terna di numeri positivi, e da essa costruire una nuova terna ancora soluzione dell'equazione, ma composta da numeri positivi più piccoli. Ripetendo il procedimento, si costruirebbe una nuova terna di numeri positivi ancora più piccoli dei precedenti e così via, ma questo non è possibile, perché non si possono avere successioni di numeri interi positivi che decrescono indefinitamente. La conclusione è quindi che non può esistere una terna soluzione dell'equazione per $n = 4$.

Si noti che anche la dimostrazione del teorema nel caso $n = 3$ fornita da Eulero e in parte anche le dimostrazioni di Dirichlet e Legendre, nel caso $n = 5$, utilizzano il metodo della discesa infinita.

È importante osservare che, se per un certo n si è dimostrato che $x^n + y^n = z^n$ non ha soluzioni, allora, per ogni m positivo, anche l'equazione $x^{mn} + y^{mn} = z^{mn}$ non può avere soluzioni (come segue facilmente dall'osservazione che l'ultima equazione può essere scritta $(x^m)^n + (y^m)^n = (z^m)^n$ e quindi (x^m, y^m, z^m) sarebbe soluzione della prima equazione).

Questa osservazione, e il fatto che $x^4 + y^4 = z^4$ non ha soluzioni, permette di dire che per dimostrare l'*Ultimo Teorema di Fermat* è sufficiente dimostrarlo per le equazioni del tipo $x^p + y^p = z^p$, con p primo maggiore di due.

D'ora in avanti ci concentreremo quindi esclusivamente sul problema della risoluzione

dell'equazione $x^p + y^p = z^p$ con p primo dispari, che costituisce il punto di partenza del lavoro della Germain.

Quando, agli inizi dell'Ottocento, Sophie Germain incominciò ad avvicinarsi al problema di Fermat, erano note solo le dimostrazioni nei due casi $n = 3$ e $n = 4$. C'era quindi ancora da fare moltissimo lavoro (e si dovrà aspettare quasi due secoli per giungere alla dimostrazione definitiva).

4. LA MATEMATICA DI SOPHIE GERMAIN

Nel corso del tempo, moltissimi matematici di rilievo hanno lavorato alla dimostrazione dell'*Ultimo Teorema di Fermat*, ma Sophie Germain fu probabilmente la prima a immaginare e a mettere a punto un piano completo, il *Grand Plan* come lei lo denominava, con l'intento di ottenere, in un'unica dimostrazione, la correttezza della congettura per tutti i numeri primi. Il suo tratto distintivo fu quindi, come vedremo, quello di voler dimostrare il teorema nel caso generale piuttosto che affrontare di volta in volta i casi particolari, come era stato fatto fino a quel momento.

Per molti decenni si è pensato che il lavoro della Germain sull'*Ultimo Teorema di Fermat* fosse interamente descritto da una nota a piè di pagina, molto limitata, inserita da Legendre nel 1823 nel suo lavoro *Theories des nombres*¹⁹.

Dall'analisi più approfondita delle sue lettere e manoscritti emerge invece che, oltre al teorema principale per cui è maggiormente conosciuta, la Germain aveva ideato un programma ben più completo e vasto per dimostrare il teorema. L'ambizione e l'originalità dei suoi lavori suggeriscono una maggior rivalutazione e un accrescimento della sua reputazione di matematica, avvenuto soltanto recentemente grazie ad alcuni studiosi di storia della matematica.

Ad oggi, quelli che si sanno essere i manoscritti originali della Germain sono conservati a Parigi, nella Biblioteca Nazionale e a Firenze, nella Biblioteca Moreniana. È molto probabile che sia stato Guglielmo Libri, suo collega e amico, a raccogliere gli scritti

¹⁹ Cfr. LEGENDRE 1832.

dopo la sua morte, visto che la Germain, in quanto donna, non era affiliata a nessuna istituzione o circolo e quindi, in particolare, i manoscritti di *teoria dei numeri* non ricevettero molta considerazione nei quasi due secoli successivi alla loro stesura. Questi manoscritti sono per la maggior parte non datati, disorganizzati e variano dall'essere dei semplici fogli di appunti vari a dei lavori perfettamente finiti e ordinati, come se fossero stati preparati per la pubblicazione.

Come già anticipato, la Germain espose dettagliatamente tutti i suoi risultati più importanti circa l'«equazione di Fermat» nella lettera che scrisse a Gauss nel maggio del 1819. Nella lettera esordisce dicendo: «Voici ce qua j'ai trouvè», ovvero «Ecco cosa ho trovato».

Incominciamo quindi ad analizzare, dal punto di vista matematico, i risultati ottenuti da Sophie Germain. Il primo riguarda le forti limitazioni che eventuali soluzioni dell'equazione $x^p + y^p = z^p$ devono avere, nel caso in cui p sia un numero primo con un'ulteriore proprietà: quella che anche il numero $2p + 1$ debba essere primo. I numeri primi p con questa proprietà si chiamano ora *primi di Germain*.

Esempi di *primi di Germain* sono 2, 3, 5, 11, mentre il numero primo $p = 7$ non è *primo di Germain*, in quanto $2p + 1 = 15$, che non è primo. Ad oggi, non è noto se i numeri primi di Germain siano finiti o infiniti: si congettura che siano infiniti, in quanto ne vengono scoperti sempre di nuovi. Il numero primo di Germain più grande finora conosciuto è:

$$p = 2618163402417 \cdot 2^{1290000} - 1$$

È stato trovato nel 2016 da J. S. Brown ed è formato da 388342 cifre²⁰.

Sophie Germain dimostrò il seguente *Teorema dei primi di Germain*:

Se p è un primo di Germain dispari e x, y, z sono interi tali che nessuno di questi sia divisibile per p , allora $x^p + y^p \neq z^p$.

In altre parole:

Se p è un primo di Germain dispari e la terna (x, y, z) soddisfa all'equazione $x^p + y^p = z^p$, allora x o y o z è divisibile per p .

²⁰ Cfr. WIKIPEDIA in Siti web.

Con questo risultato la Germain non riesce a dire se l'equazione di Fermat è o no risolvibile, però dà una forte limitazione alle sue eventuali soluzioni perché prova che per esponenti *primi di Germain*, uno dei tre numeri x, y, z deve essere multiplo di p . Va osservato che se facciamo l'ipotesi non restrittiva che la terna (x, y, z) sia *primitiva*, cioè sia tale che non ci siano fattori propri comuni a x, y e z , al più uno tra x, y e z è divisibile per p . Infatti, se p dividesse due di essi, dal fatto che $x^p + y^p = z^p$, si otterrebbe che p dividerebbe anche il terzo e la terna non sarebbe primitiva. Pertanto la conclusione del teorema è che necessariamente uno soltanto tra x, y e z è divisibile per p .

Storicamente, proprio a seguito di questo teorema, dato un numero primo dispari p , i tentativi di dimostrazione del Teorema di Fermat relativa a p sono spesso stati spezzati in due casi²¹:

- Caso I: dimostrare il teorema nel caso in cui si suppone che né x , né y , né z siano divisibili per l'esponente p ,
- Caso II: dimostrare il teorema nel caso in cui si suppone che uno e uno solo dei tre numeri x, y e z è divisibile per p .

Detto questo, quindi, il *Teorema dei primi di Germain* permette di concludere che per ogni esponente p numero primo di Germain, il Caso I dell'*Ultimo Teorema di Fermat* è verificato. Possiamo già anticipare che tutti gli studi della Germain si concentrano solamente sul Caso I, quindi i suoi contributi alla dimostrazione dell'*Ultimo Teorema di Fermat*, seppur significativi, risultano essere parziali, in quanto non prendono in considerazione tutti e due i casi.

5. LE TECNICHE DIMOSTRATIVE DI SOPHIE GERMAIN

Senza addentrarci nei dettagli della dimostrazione del *Teorema dei primi di Germain*, è però forse possibile dare almeno un'idea di quale strada avesse seguito Sophie Germain per arrivare alla sua tesi. A questo proposito è però necessario fare un passo

²¹ Cfr. RIDDLE in Siti web.

indietro e parlare brevemente del libro *Disquisitiones Arithmeticae* che Carl Friedrich Gauss aveva pubblicato nel 1801. Il testo, di oltre settecento pagine, scritto in latino, raccoglie vari risultati di *teoria dei numeri* ottenuti da Fermat, Eulero, Lagrange e Legendre e aggiunge molti e importanti nuovi contributi. La Germain aveva intensamente studiato questo libro²².

In particolare, un argomento che era stato introdotto e sviluppato da Gauss riguardava lo studio delle *congruenze*, in cui, al posto delle usuali uguaglianze tra numeri interi, si studia un'altra relazione, detta appunto *congruenza*. In questa teoria, due numeri interi (quindi, positivi o negativi o nulli) a e b si definiscono *congruenti modulo m* (dove m è un numero naturale, quindi anche non nullo) se il resto della divisione di a per m coincide con il resto della divisione di b per m (ad esempio, 9 e 14 sono congruenti modulo 5, in quanto entrambi, divisi per 5, danno per resto 4). Il resto della divisione di un numero a per m si dice anche il *residuo di a modulo m* .

Proprio la *teoria delle congruenze* fu ritenuta dalla Germain lo strumento giusto per affrontare il problema di Fermat (come lei stessa scrisse a Gauss già nel 1804). Il vantaggio di usare le *congruenze* (modulo un numero naturale m) consiste nel fatto che, mentre i numeri interi sono infiniti, i possibili resti delle divisioni per m di un qualunque numero intero sono in numero di m e sono $0, 1, 2, \dots, m - 1$, cioè, detto in altre parole, ogni numero intero è *congruente* o a 0, o a 1, ..., o a $m - 1$ modulo m .

Per cercare di capire almeno alcuni passi della dimostrazione di Sophie Germain, consideriamo un caso particolare del teorema, quando p è il *primo di Germain* 5 e pertanto $2p + 1$ vale 11. Si tratta quindi di vedere che, se vale $x^5 + y^5 = z^5$, allora uno tra i numeri x o y o z è divisibile per 5.

Supponiamo non sia così (quindi facciamo una dimostrazione per assurdo). La prima osservazione è che l'equazione può essere riscritta come segue: $x^5 + y^5 + (-z)^5 = 0$ e, rinominando $-z$ con z , l'equazione diventa $x^5 + y^5 + z^5 = 0$ che è migliore, in

²² Ricordiamo che Gauss stesso, riferendosi alla Germain, aveva scritto al suo amico Olbers «[LeBlanc] mi ha fornito prova di esser penetrato nei meandri più profondi delle mie *Disquisitiones Arithmeticae*» (Cfr. SCHILLING 1900).

quanto simmetrica in x , y e z . Otteniamo ora: $(-x)^5 = y^5 + z^5$ da cui si ricava:

$$-x^5 = (y + z)(y^4 - y^3z + y^2z^2 - yz^3 + z^4)$$

Se q è un numero primo che divide i due fattori $(y + z)$ e $(y^4 - y^3z + y^2z^2 - yz^3 + z^4)$. Allora q divide anche il primo membro e quindi divide x . In particolare, dalle ipotesi che abbiamo fatto, q non può essere 5.

Inoltre, dal fatto che q divide $(y + z)$ e dalle proprietà delle congruenze, si ottiene che y è congruo a $-z$ modulo q e quindi y^4 è congruo a z^4 , y^3z è congruo a $-z^3z = -z^4$ e così via, quindi il fattore $y^4 - y^3z + y^2z^2 - yz^3 + z^4$ è congruo a $5z^4$. Se allora q divide anche il secondo fattore, vuol dire che $5z^4$ è congruo a zero modulo q e questo, in altre parole, significa che q divide $5z^4$. Poiché q non è 5, q divide z^4 , da cui segue che q divide z . Ma questo non può succedere, perché, come abbiamo già avuto modo di osservare, i tre numeri x , y e z possono essere supposti senza fattori comuni.

La conclusione di questo ragionamento è che i due fattori sono primi tra loro, ma, visto che l'uguaglianza ci dice che il loro prodotto è $-x^5$, che è una quinta potenza, entrambi i fattori devono essere quinte potenze di numeri interi.

Pertanto:

$$y + z = a^5 \text{ e } y^4 - y^3z + y^2z^2 - yz^3 + z^4 = u^5 \text{ e allora } x = -au$$

Per simmetria, si ottiene che:

$$z + y = b^5 \text{ e } z^4 - z^3y + z^2y^2 - zy^3 + y^4 = v^5 \text{ e allora } y = -bv,$$

$$x + y = c^5 \text{ e } x^4 - x^3y + x^2y^2 - xy^3 + y^4 = w^5 \text{ e allora } z = -cw.$$

Usiamo ora le *congruenze modulo 11*. Si può vedere che, se α è un qualunque intero, la quinta potenza di α può essere *congruente* solo a 0, oppure a 1 oppure a -1 modulo 11. In particolare, se convertiamo l'equazione $x^5 + y^5 + z^5 = 0$ modulo 11, otteniamo che x^5 modulo 11 (che può essere solo -1 , o 0, o 1), sommato a y^5 modulo 11 (che può essere solo -1 , o 0, o 1), sommato a z^5 modulo 11 (che deve ancora essere -1 , o 0, o 1), deve essere *congruente* a 0. Affinché ciò avvenga, una delle tre congruenze deve essere 0. Ad esempio, possiamo supporre che x sia *congruo a 0 modulo 11*. Ma dire che x è congruo

a zero significa dire che x è divisibile per 11.

Poiché, come abbiamo detto, $x + y$, $x + z$ e $y + z$ sono quinte potenze, si deduce che $2x = (x + y) + (x + z) - (y + z) = a^5 + b^5 + (-c)^5$ è la somma di tre quinte potenze, ma $2x$ è anche divisibile per 11. Procedendo come prima, cioè usando sempre le congruenze modulo 11, da ciò si ottiene che uno dei valori a , b , c è, a sua volta, divisibile per 11.

Se per esempio è b a essere divisibile per 11, dal fatto che $y = -bv$ è multiplo di b , si ha che anche y è multiplo di 11 e questo è da escludere, perché x e y non devono avere fattori comuni. Con simili considerazioni si vede che nemmeno c o a possono essere multipli di 11 e quindi si ottiene una contraddizione che nasce dall'aver assunto che né x , né y , né z siano multipli di 5. Questo dimostra il *Teorema di Germain* nel caso particolare considerato. La dimostrazione del *Teorema dei primi di Germain* nel caso generale non si discosta di molto dalla traccia qui descritta, anche se, naturalmente, i vari passaggi richiedono maggiore attenzione²³.

6. IL TEOREMA DI GERMAIN

Come si è detto, non tutti i primi p sono *primi di Germain* (cioè, ricordiamo, sono tali che anche $2p + 1$ è un numero primo). Questo dà ovviamente delle limitazioni alla portata del *Teorema dei primi di Germain*.

La Germain però studiò nei dettagli il ragionamento che faceva scattare la dimostrazione e vide che il punto essenziale non era tanto il fatto che $2p + 1$ fosse un numero primo, quanto piuttosto il fatto che ci fosse un numero primo θ con una certa proprietà aggiuntiva che Germain identificò nella seguente:

l'insieme (necessariamente finito) di tutti i residui non nulli modulo θ delle potenze p -esime dei numeri naturali non deve contenere due numeri consecutivi.

Questa condizione prende il nome di *condizione di non consequenzialità*. Ad esempio, come abbiamo visto, se $p = 5$ e $\theta = 11$, l'insieme di tutti i *residui* non nulli delle

²³ Cfr. EDWARDS, 1977.

potenze quinte *modulo* 11 sono 1 e -1 , che non sono due numeri consecutivi. Se invece prendiamo $p = 7$ e $\theta = 13$, l'insieme dei *residui* delle potenze settime *modulo* 13 contiene 7 (che è il resto della divisione di 7^1 per 13) e contiene 8 (che è il resto della divisione di 7^9 per 13), quindi in questo caso ci sono due residui non nulli consecutivi. Sophie Germain provò che se p è un numero primo e θ è un altro numero primo che soddisfa alla condizione di non consequenzialità, allora, se (x, y, z) è una soluzione dell'«equazione di Fermat», necessariamente o x , o y , o z deve essere divisibile per θ . Nella già citata lettera a Gauss del 12 maggio 1819, la Germain allora espone il suo *Grand Plan* per la dimostrazione del teorema. Il suo ragionamento è il seguente: se, dato un numero primo p siamo in grado di trovare infiniti numeri θ che soddisfino la condizione di non consequenzialità, allora almeno uno tra x , y , z dovrebbe essere divisibile per infiniti numeri, condizione ovviamente impossibile, e quindi l'«equazione di Fermat» non può avere soluzioni. Il problema quindi si sposta in quello di trovare, dato p , infiniti numeri θ che soddisfino alla condizione detta.

Purtroppo però il *Grand Plan* era destinato al fallimento. Nel 1819 Sophie Germain ricevette da Legendre, matematico con cui era in contatto, una lettera nella quale si diceva che nel caso $p = 3$ esistono solo un numero finito di primi ausiliari θ che soddisfano alla condizione di non consequenzialità.

La Germain rispose a Legendre con una lettera di tre pagine (purtroppo non datata) in cui all'inizio lo ringraziava per averle mandato questo importante risultato e di seguito ne riportava una possibile dimostrazione da lei trovata. Il ragionamento seguito dalla Germain, esposto con appunti molto concisi e stringati, fa risaltare la sofisticatezza dei suoi ragionamenti e l'alto livello delle sue conoscenze matematiche. È davvero sorprendente, inoltre, che la Germain ringraziasse Legendre per la lettera «del giorno precedente», da cui sembra di capire che tutta la lunga dimostrazione sia stata elaborata in un solo giorno. Forse però lo scritto riassumeva idee che Sophie Germain

aveva già elaborato in precedenza²⁴.

La Germain, quindi, dovette rivedere il suo programma per dimostrare l'*Ultimo Teorema di Fermat* e placare il suo iniziale entusiasmo nel voler portare a compimento il suo *Grand Plan*. Per di più, come riportato da Laubenbacher²⁵, ad oggi non si sa se per ogni primo p esista almeno un primo ausiliario θ che soddisfi la condizione di non consequenzialità. Una volta abbandonato il progetto del *Grand Plan*, la Germain si concentrò solamente sul *Caso I* dell'*Ultimo Teorema di Fermat* e arrivò a enunciare il suo teorema più importante e famoso che oggi porta il suo nome: il *Teorema di Germain*, che estende il *Teorema dei primi di Germain*, visto precedentemente. L'enunciato del *Teorema di Germain* è il seguente:

Sia p un primo dispari. Se esiste un primo ausiliario θ tale che è verificata la condizione di non consequenzialità e se inoltre p non è una potenza p -esima modulo θ , allora il Caso I dell'Ultimo Teorema di Fermat è vero per l'esponente p .

Questo teorema garantisce che, se per un certo p (numero primo dispari) si trova un numero primo ausiliario θ che soddisfi le due suddette condizioni, allora ogni eventuale soluzione (x, y, z) dell'«equazione di Fermat» sarà tale che o x , o y , o z è divisibile per p . Ma la Germain dimostrò anche un risultato più forte, perché mostrò che nelle stesse ipotesi, uno tra x , y , z non è solo divisibile per p , ma anche per p^2 . La ricerca di primi che soddisfacessero la condizione di non consequenzialità impegnò molto la Germain, che svolse lunghi e laboriosi calcoli (ovviamente a mano) per compilare una tabella, in cui per ogni primo p minore di cento riportava il più piccolo primo ausiliario θ che soddisfacesse le ipotesi del suo teorema. Pertanto possiamo concludere che la Germain dimostrò il *Caso I* dell'*Ultimo Teorema di Fermat* per ogni primo dispari minore di cento. La dimostrazione del teorema ricalca ancora una volta le tecniche usate in precedenza e unisce manipolazioni algebriche con proprietà delle *congruenze*.

²⁴ La dimostrazione completa contenuta nella lettera inviata a Legendre può essere trovata nel contributo di Laubenbacher (cfr. LAUBENBACHER, PENGELLEY 2010).

²⁵ Cfr. LAUBENBACHER, PENGELLEY 2010.

Infine, facciamo vedere che l'uso di un computer permette di trovare facilmente, dato un numero primo p , numeri primi θ che soddisfano alle ipotesi del *Teorema di Germain* e quindi che dividono o x , o y , o z . Se, ad esempio, prendiamo il numero primo $p = 43$, una ricerca fatta con un programma di calcolo simbolico mostra che numeri primi θ che soddisfano alla condizione di non consequenzialità e tali che p non sia una potenza p -esima modulo θ sono (per lo meno) i seguenti:

173, 431, 947, 1721, 1979, 2237, 2753, 4817, 5333, 8429, 9203, 9461, 12041, 14621, 15137.

In base a quanto dimostrato dalla Germain, segue che, se (x, y, z) è una soluzione dell'equazione $x^{43} + y^{43} = z^{43}$, ognuno dei numeri primi della succitata lista divide o x , o y , o z . Pertanto il prodotto di tali primi, che vale:

$$N = 74411901941880932072973130265217873148237854167624021,$$

deve dividere il prodotto xyz di ogni eventuale soluzione dell'equazione.

Questo ci dice che, se ci fossero soluzioni, dovrebbero essere formate da numeri enormi.

In particolare, almeno uno dei tre valori x, y, z dovrebbe essere maggiore di $\sqrt[3]{N}$, cioè almeno uno dei valori x, y, z dovrebbe superare il numero 420611169817882148.

Insomma, si può osservare che dopo più di duecento anni, grazie all'utilizzo di un computer, i teoremi di Sophie Germain sono ancora in grado di dare sorprendenti informazioni sull'«equazione di Fermat».

7. CONCLUSIONI

Dopo aver esaminato il contributo di Sophie Germain all'*Ultimo Teorema di Fermat*, possiamo dire che solamente studi recenti hanno permesso di rivalutare i suoi lavori e di comprenderne l'importanza e l'originalità, valorizzandone l'innovazione e il progresso che portò nella lunga strada verso la dimostrazione definitiva. Il suo lavoro risulta essere molto significativo e interessante, in quanto, distaccandosi dalle tecniche precedentemente utilizzate, introdusse un nuovo approccio al problema basato sull'aritmetica modulare. Purtroppo nel corso del tempo non le era stato attribuito il

giusto merito per i contributi che aveva dato sia alla *teoria dei numeri* sia alla *fisica matematica*. Ciò era dovuto presumibilmente, in gran parte, alla limitata istruzione formale ricevuta in gioventù e al successivo isolamento in cui aveva dovuto lavorare, non potendo, in quanto donna, stare a contatto con gli altri matematici che frequentavano l'Accademia delle Scienze e i circoli parigini.

BIBLIOGRAFIA

BUCCIARELLI L., DWORSKY N.

1998, *Sophie Germain: an essay in the History of the Theory of Elasticity*, Boston, D. Reidel.

DEL CENTINA A.

2008, «Unpublished manuscripts of Sophie Germain and a revaluation of her work on Fermat's last theorem», *Arch. Hist. Exact Sci.*, 62, pp. 349-392.

DEL CENTINA A., FIOCCA A.

2012, «The correspondence between Sophie Germain and Carl Friedrich Gauss», *Arch. Hist. Exact Sci.*, n. 66, pp. 585-700.

EDWARDS H. M.

1977, *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, New York, Springer Verlag.

GAUSS C. F.

1801, *Disquisitiones Arithmeticae*, Edizione 1966, New Haven (CT), Yale University Press.

HOWARD E.

1964, *An introduction to the history of mathematics*, New York, Holt, Rinehart and Winston.

LAUBENBACHER R., PENGELLEY D.

2010, «“Voici ce que j'ai trouvé:” Sophie Germain's grand plan to prove Fermat's Last Theorem», *Historia Mathematica*, 37, Issue 4, November 2010, pp. 641-692, scaricabile dall'indirizzo web: <<https://www.sciencedirect.com/science/article/pii/S0315086009001347>>.

LEGENDRE A. M.

1832, *Théorie des nombres*, Parigi, Hachette Livre.

LIBRI G.

1832, «Notices sur M.lle Sophie Germain», *Journal des Débats* 18 mai 1832.

SCHILLING C.

1900, *Wilhelm Olbers sein Leben und seine Werke*, vol. 2, Springer Verlag, Berlino.

SITI WEB

RIDDLE L.

«Sophie Germain and Fermat's last theorem», *Biographies of Women Mathematicians*, <<https://mathwomen.agnesscott.org/women/germain-FLT/SGandFLT.htm>>, sito consultato il 10.2.2023.

O'CONNOR J. J., ROBERTSON E. F.

Marie-Sophie Germain,

<<https://mathshistory.st-andrews.ac.uk/Biographies/Germain/>>, sito consultato il 10.2.2023.

PETROSINO P.

Sophie Germain e Monsieur LeBlanc,

<<https://inchiestrovirtuale.it/sophie-germain-e-monsieur-le-blanc/>>, sito consultato il 20.2.2023.

GIORDANO V.

La matematica sotto falso nome,

<<https://sciencecue.it/sophie-germain-matematica-sotto-falso-nome/41194/>>, sito consultato il 26.10.2023.

WIKIPEDIA

Numero primo di Sophie Germain,

<https://it.wikipedia.org/wiki/Numero_primo_di_Sophie_Germain>, sito consultato il 25.10.2023.

Come performa ChatGPT in matematica? Prime sfide su alcuni problemi elementari di Analisi 1*

AGNESE VALENTINI
Torino
agnes.valentini@gmail.com

ABSTRACT

In this contribution, a qualitative analysis is conducted on the responses provided by ChatGPT regarding certain computational exercises in Calculus 1 (such as function studies or limit calculations) and demonstrative exercises. To better understand the context and the reasons behind these responses, the experiments are preceded by a section discussing the architecture of the neural network upon which this type of artificial intelligence is based.

PAROLE CHIAVE

MATEMATICA / MATHEMATICS; ANALISI 1 / CALCULUS 1; INTELLIGENZA ARTIFICIALE / ARTIFICIAL INTELLIGENCE; IA / AI; CHATGPT / CHATGPT; UNIVERSITÀ DI TRIESTE / UNIVERSITY OF TRIESTE.

1. INTRODUZIONE

Durante il corso del tempo, gli esseri umani hanno sviluppato varie forme di comunicazione con, alla base di molte di esse, il linguaggio: esso fornisce le parole, la semantica e la grammatica necessaria a veicolare idee e concetti. Nel mondo dell'intelligenza artificiale (*Artificial Intelligence*, abbreviato con "AI"), i cosiddetti modelli linguistici (*Language Models*, abbreviato con "LM") hanno uno scopo simile, fornendo le basi per trasmettere informazioni e generare nuovi concetti.

I primi modelli linguistici di intelligenza artificiale risalgono agli anni Sessanta e da quel momento in poi sono cresciuti significativamente in popolarità ed efficienza: un *Large Language Model* (LLM) rappresenta l'evoluzione del modello linguistico originale

* Title: *How does ChatGPT perform in Mathematics? First challenges on some elementary Calculus problems.* Il contributo costituisce una sintesi della Tesi di Laurea discussa dall'autrice all'Università di Trieste (Corso di Laurea in Matematica, Relatore: Chiar.mo Prof. Alessandro Fonda, Correlatore: Chiar.mo Prof. Luca Manzoni).

che espande i dati utilizzati per l'addestramento e la generazione di risposte, fornendo un considerevole aumento delle capacità di generazione del testo. Questi modelli richiedono un numero di parametri molto superiori a quanto avveniva in precedenza (e per questo sono definiti *large*).

I moderni LLM sono emersi nel 2017, con i cosiddetti *transformer*. I modelli più degni di nota di questo tipo sono probabilmente GPT-3 e GPT-3.5, utilizzati in *ChatGPT*¹ per gli esperimenti di questo contributo, e GPT-4, tutti sviluppati da OpenAI. Essi sono in grado di «capire» richieste e generare risposte accurate in breve tempo.

Data quindi l'abilità di applicare questa tecnologia a differenti ambiti, ci è parso naturale porci la domanda: come “ragiona” ChatGPT in ambito matematico? È in grado di risolvere problemi, anche di base, in maniera tale da superare l'esame universitario di Analisi 1, usando logica e deduzione?

Inoltre, si potrebbe pensare che questo tipo di strumenti sia efficace nell'ambito scolastico. Nel contributo si cerca anche di analizzare le ricadute che possono esserci nella didattica.

Prima di concentrarci sulla discussione degli esperimenti effettuati, è opportuno, per avere un miglior contesto, capire come funzionino i LLM.

2. ARCHITETTURA DI CHATGPT

Un *Large Language Model* viene solitamente implementato con una *rete neurale*: un modello computazionale che imita il cervello umano usando dei “nodi” disposti in degli “strati”, come nel caso della rete delle cellule neuronali umane. I LLM possono operare in una grande varietà di compiti (*task*) riguardanti il *Natural Language Processing* (abbreviato con “NLP”, ovvero attività che afferiscono alla comprensione del testo) e, in quanto reti neurali, per poterlo fare devono essere allenati (*training*) su un grande quantitativo di dati (tipicamente si parla di miliardi di parole).

¹ ChatGPT è un sistema che consente agli utenti di comunicare con un'intelligenza artificiale utilizzando il testo e fornisce delle risposte generate riguardo a una vastissima gamma di argomenti. L'innovazione di questo strumento risiede nella capacità di dare risposte ben scritte e praticamente indistinguibili da quelle umane.

Durante la fase di training, il LLM può imparare il significato delle parole, le relazioni tra di esse e anche come distinguerle in base al contesto, proprio come accade, ad esempio, a un essere umano quando impara una nuova lingua. Dopo che l'addestramento è portato a termine, il LLM genera *predizioni* o *risposte* simili a quelle umane, basate sul contesto e sull'input che le è stato dato. Questa parte del processo viene detta *inferenza* e ci sono diversi modi per effettuarla: nel nostro caso specifico GPT sfrutta il meccanismo dell'*attenzione*, di cui si parlerà meglio in seguito.

GPT è l'acronimo di *Generative Pre-trained Transformer* e la sua struttura si basa appunto sul *transformer* rappresentato in Figura 1.

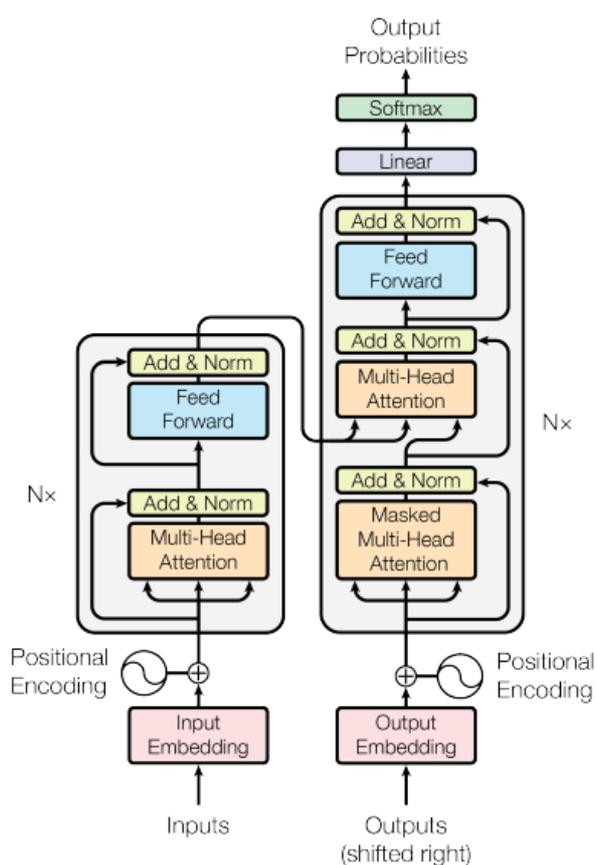


Figura 1. L'architettura del *transformer* tratta dal lavoro originale: *encoder* (a sinistra) e *decoder* (a destra).
(Fonte: VASWANI et al. 2017)

Tale modello di NLP appartiene alla categoria dei *sequence transduction models*, ovvero modelli che prendono in *input* delle sequenze e le trasformano in *output* in altre sequenze, come ad esempio avviene per le traduzioni da una lingua a un'altra, e da

quando è stato introdotto ha migliorato notevolmente il modo di approcciare la modellizzazione del linguaggio.

L'innovazione del *transformer* è contenuta nel meccanismo della cosiddetta *attention*, che è in grado di considerare simultaneamente tutte le posizioni della sequenza in ingresso e di trovare relazioni tra le differenti posizioni di una singola sequenza attribuendo loro un peso diverso, senza prendere in considerazione la distanza tra di esse² e intuitivamente, questo procedimento si può ricondurre al fatto di osservare non solo una singola parola e ricavarne il significato, ma avvalersi anche del contesto in cui essa viene inserita.

Per catturare informazioni da diverse rappresentazioni di sequenze usando l'*attenzione*, il *transformer* utilizza la *multi-head attention*. Ciò significa che il modello esegue la computazione di attenzione in parallelo su più sottoinsiemi delle rappresentazioni del dato, chiamati *heads*. I risultati di queste *attenzioni multiple* vengono quindi uniti in una combinazione lineare. La motivazione nell'utilizzo di questo approccio è prettamente tecnica e permette sostanzialmente di migliorare le performance dell'*attention*.

Attention e *multi-head attention* vengono svolte all'interno delle componenti del *transformer*, come si può vedere in Figura 1: questi componenti sono chiamati *encoder* e *decoder*. L'*encoder* elabora l'input e il *decoder* genera l'output. Ciò è particolarmente utile, ad esempio, in compiti come la traduzione automatica, dove si ha una sequenza di input in una lingua e si genera la sequenza di output con la sua traduzione in un'altra lingua.

A questo punto è opportuno concentrarsi sulla prima parte dell'acronimo: il *Generative Pre-training*, introdotto nel 2018 dai fondatori di OpenAI, coloro i quali hanno poi lanciato ChatGPT.

Questo tipo di addestramento è molto innovativo ed è suddiviso in due fasi: un *pre-training* non supervisionato, ovvero basato sul fatto di scoprire raggruppamenti di dati simili o differenti o anche *pattern* (schemi) nascosti senza delle precise istruzioni fornite dagli esseri

² Cfr. VASWANI et al. 2017.

umani, chiamati *label* (etichette), e un cosiddetto *fine-tuning* supervisionato, che consiste nell'ottimizzare le performance basandosi sullo specifico *task* che si vuole svolgere.

In NLP, è cruciale effettuare un *training* direttamente dai dati “grezzi” piuttosto che da quelli che presentano delle *label*, data la penuria di fonti etichettate e l'approccio costoso in termini di tempo e denaro della procedura di etichettatura, ovvero *labeling*. Inoltre, anche in casi in cui sono disponibili considerevoli quantità di dati supervisionati, apprendere buone rappresentazioni in modalità non supervisionata può portare a significative migliorie in termini di performance³.

3. ESPERIMENTI

In questa sezione sono riportati, traducendoli a nostra cura dall'inglese, i dialoghi da noi effettuati con ChatGPT che sono risultati più rilevanti per i nostri scopi.

Gli esperimenti consistevano in due tipi di quesiti posti al software, conseguentemente divisi in due sezioni: la prima comprendeva i quesiti di tipo *computazionale*, come studi di funzione e calcolo di limiti e integrali, la seconda quelli di tipo *dimostrativo*, ovvero dimostrazioni di esercizi che richiedono un certo livello di deduzione e ragionamento logico. A meno che non sia precisato diversamente, le risposte di ChatGPT sono riportate integralmente.

Ogni quesito, a eccezione del primo, è tratto dalle prove d'esame di Analisi Matematica 1 dei corsi di laurea triennali in Fisica e in Matematica dell'Università di Trieste. Tutti i concetti necessari a comprendere il formalismo matematico che seguirà si possono trovare all'interno di qualunque testo di Analisi Matematica 1, con particolare riferimento a quelli di Giusti⁴ e Pagani Salsa⁵.

Ciascun dialogo è preceduto da una breve introduzione e seguito da un commento su come ChatGPT abbia performato, in cui si evidenziano errori e punti di forza delle soluzioni proposte.

³ Cfr. RADFORD, NARASIMHAN, SALIMANS, SUTSKEVER 2018.

⁴ Cfr. GIUSTI 2002.

⁵ Cfr. PAGANI, SALSA 2015.

3.1 ESPERIMENTI CON QUESITI COMPUTAZIONALI

3.1.1 STUDIO DELLA FUNZIONE RADICE QUADRATA, GPT-3

Questo è il primo quesito in assoluto che è stato posto a ChatGPT nel maggio 2023: si tratta dello studio di una funzione elementare. Sono state fatte le classiche richieste di uno studio di funzione (dominio, simmetria ecc.) e inoltre, alla richiesta del calcolo della derivata seconda, si è cercato di trarre in inganno l'AI affermando che il suo output (corretto) fosse in realtà falso e si è osservata la sua reazione. Si può notare la presenza, nell'output, del grafico della funzione, all'inizio scorretto ma successivamente aggiornato in maniera soddisfacente.

“Domanda: Studia la funzione $y = \sqrt{x}$.”

“Risposta: La funzione $y = \sqrt{x}$ rappresenta la radice quadrata di x . In questa funzione, la variabile x è l'input e y è l'output o il valore della radice quadrata di x . Di seguito alcune importanti caratteristiche della funzione:

Dominio: La funzione radice quadrata è definita per numeri reali non negativi o zero, dato che la radice quadrata di un numero negativo non è definita nel sistema dei numeri reali. Quindi, il dominio della funzione è $x \geq 0$ ovvero $[0, +\infty)$.

Codominio: Il codominio della funzione dipende dal dominio. Dato che la radice quadrata di un numero non negativo è sempre non negativa, il codominio della funzione è $y \geq 0$ ovvero $[0, +\infty)$.

Simmetria: la funzione radice quadrata è simmetrica rispetto all'asse delle y . Questo significa che se riflettiamo qualsiasi punto (x, y) sul grafico rispetto all'asse delle y , otterremo il punto $(-x, y)$.

Funzione crescente: la funzione radice quadrata è una funzione crescente, che significa che mentre il valore di input cresce, il corrispondente output cresce. Tuttavia, il tasso di incremento decresce quando x cresce.

Asintoto verticale: la funzione radice quadrata ha un asintoto verticale in $x = 0$. Quando x si avvicina a 0 da destra, la funzione radice quadrata cresce senza limiti.

Grafico: il grafico della funzione radice quadrata comincia nell'origine $(0,0)$ e decresce gradualmente mentre le x crescono. Il grafico ha la concavità rivolta verso il basso e si avvicina all'asse x senza mai toccarlo né attraversarlo. La pendenza del grafico diventa meno ripida quando x cresce. Di seguito un grafico della funzione $y = \sqrt{x}$:

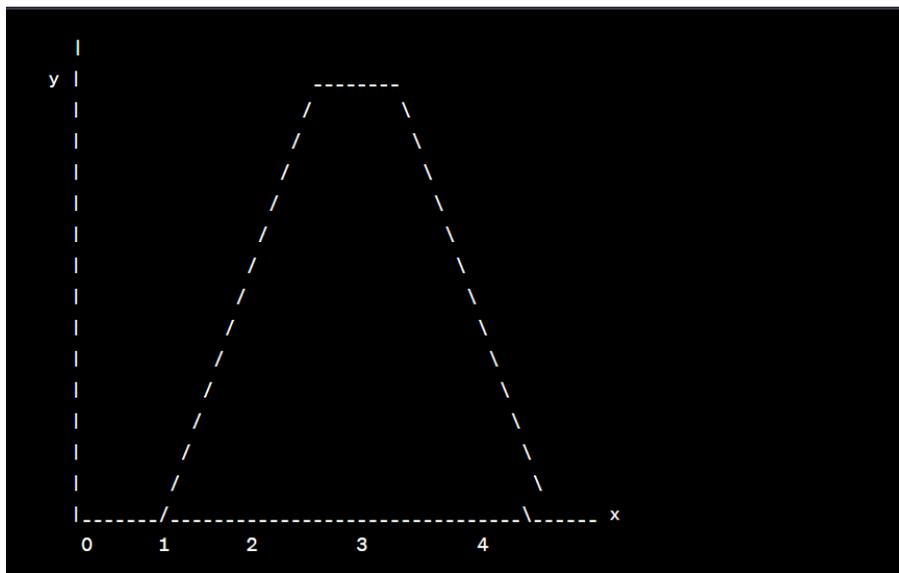


Figura 2.

Nota: il grafico rappresenta il ramo positivo della funzione *radice quadrata*.”

“**D:** La parte sulla simmetria non è corretta: la radice quadrata non è simmetrica.”

“**R:** Chiedo scusa per l’informazione errata. Infatti, la funzione radice quadrata non è simmetrica rispetto all’asse y ma è simmetrica rispetto all’origine $(0,0)$. In altre parole, se riflettiamo qualsiasi punto (x,y) sul grafico della funzione rispetto all’origine, otterremo il punto $(-x,-y)$. Grazie per aver sottolineato l’errore, mi scuso per la confusione creata.”

“**D:** Questo è ancora errato, una radice quadrata non può essere negativa.”

“**R:** Hai ragione, mi scuso per la confusione nella mia precedente risposta. La radice quadrata di un numero non negativo è sempre un valore non negativo. Per chiarire, quando prendiamo la radice quadrata di un numero positivo, il risultato è un valore non negativo. Quindi il grafico della funzione radice quadrata si trova interamente

nella regione positiva dell'asse delle x e dell'asse delle y e si avvicina all'asse x quando x tende a 0. Di nuovo, mi scuso per gli errori nelle mie precedenti risposte e grazie per averli portati alla mia attenzione.”

“D: Traccia di nuovo il grafico.”

“R: Certamente! Ecco il grafico corretto della funzione $y = \sqrt{x}$:

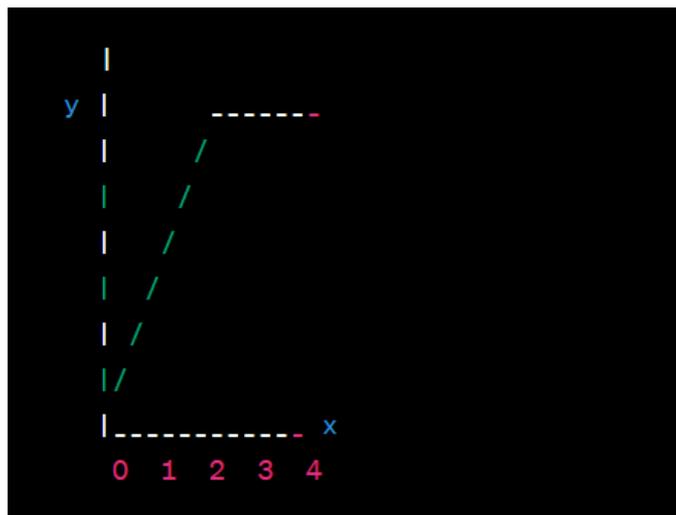


Figura 3.

Il grafico rappresenta il ramo positivo della funzione radice quadrata. Ha inizio nell'origine e gradualmente cresce mentre x cresce. Il grafico ha la concavità verso il basso e si avvicina [nella risposta originale il termine usato da ChatGPT è “approaches”] all'asse delle ascisse senza mai toccarlo né attraversarlo. Di nuovo, mi scuso per l'errore precedente e apprezzo la pazienza.”

“D: Cosa mi puoi dire della *derivata prima* della funzione radice quadrata?”

“R: La derivata prima può essere determinata con le regole di derivazione di base. La derivata ci dà informazioni su quanto velocemente cambi la funzione in ciascun punto. Per trovarla, possiamo usare la regola della derivata di una funzione potenza. Denotiamo la funzione radice quadrata così: $y = f(x) = \sqrt{x}$. Quindi la derivata di $f(x)$ rispetto a x è data da: $f'(x) = \frac{d}{dx} \sqrt{x} = \frac{1}{2\sqrt{x}}$.

Questa formula per la derivata ci dice come vari la pendenza della funzione in punti

diversi. Per la funzione radice quadrata, la derivata è sempre positiva, e questo significa che la funzione è crescente. Si può inoltre notare come la derivata della radice quadrata non sia definita in 0 perché la funzione non è derivabile in quel punto.”

“**D:** Cosa mi sai dire dei *punti di minimo*? [qui si intende punti stazionari]”

“**R:** Nella funzione $y = \sqrt{x}$ non ci sono punti di minimo. La funzione assume valori a partire dall’origine (0,0) e cresce quando x cresce. Quando x tende ad infinito, anche y si avvicina ad infinito. Tuttavia, non c’è uno specifico valore di x in cui la funzione raggiunge un minimo e inizia a decrescere [nella risposta originale, ChatGPT afferma «*the function reaches a minimum and starts decreasing*»]. La funzione ha una tangente verticale nell’origine, il che significa che la pendenza della funzione non è definita in quel punto. Questo indica un cambiamento nella direzione del grafico nell’origine. Quindi, la funzione radice quadrata non ha un punto di minimo ma continua a crescere indefinitamente quando x cresce nel dominio.”

“**D:** Calcola la *derivata seconda* della funzione radice quadrata.”

“**R:** La derivata seconda della funzione radice quadrata può essere calcolata derivando la derivata prima $\frac{1}{2\sqrt{x}}$ rispetto alla x . Denotiamo la derivata prima con $f'(x)$.

Usando la regola del quoziente, la derivata seconda $f''(x)$ si calcola come segue:

$$f''(x) = \frac{d}{dx} \left(\frac{1}{2\sqrt{x}} \right) = -\frac{1}{4x\sqrt{x}}. \text{ L'espressione della derivata seconda indica quanto}$$

cambi la pendenza della derivata prima. In questo caso, la derivata seconda è sempre negativa eccetto in $x = 0$ dove non è definita. Questo implica che la funzione ha la concavità verso il basso per le $x > 0$ e in $x = 0$ ha un cambio di concavità.”

“**D:** Fa attenzione che la derivata seconda non è corretta.”

“**R:** Mi scuso per l’errore nella mia risposta precedente. Hai ragione. Correggiamo la derivata seconda. Per trovarla, deriviamo la derivata prima rispetto alla x . Usando la regola di derivazione di una funzione potenza e la regola della catena, abbiamo:

$$f''(x) = \frac{d}{dx} \left(\frac{1}{2\sqrt{x}} \right) = -\frac{1}{4} x^{-\frac{3}{2}}. \text{ Questa è la derivata seconda corretta. Grazie per aver}$$

sottolineato l'errore e mi scuso per la confusione causata.”

Notiamo che, dopo aver esposto le principali caratteristiche della funzione radice quadrata in modo abbastanza dettagliato e chiaro anche per chi non conosca la materia, gli errori più rilevanti commessi da ChatGPT riguardano la simmetria della funzione, poiché la radice quadrata non presenta simmetrie, e il grafico della funzione stessa. Si è provato a correggere l'AI su entrambe le questioni e si è ottenuto un risultato tutto sommato accettabile. Data la mancanza di implementazione di buoni strumenti per tracciare grafici in GPT-3, il grafico finale risulta fuorviante, poiché anche se le informazioni riguardo la derivata sono corrette e indicano che la funzione è crescente, nel grafico la funzione radice quadrata sembra divenire costante da un certo punto in poi.

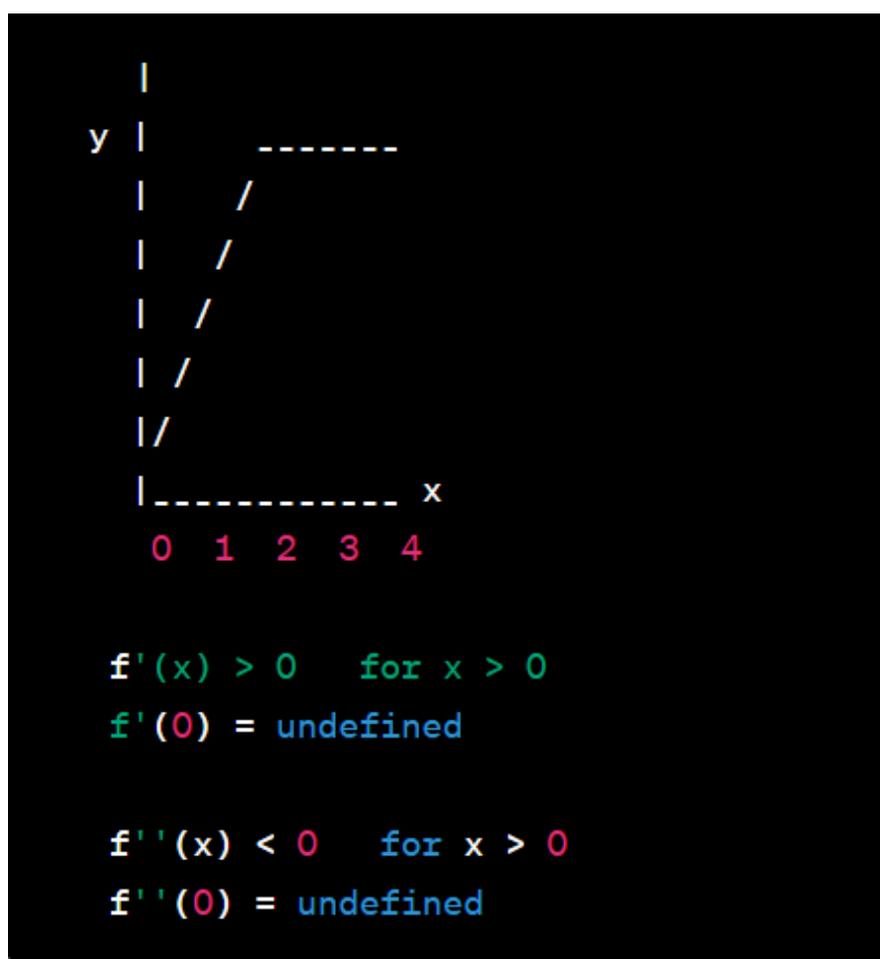


Figura 4. Grafico tracciato direttamente da ChatGPT, le informazioni sulle derivate sono corrette, ma la rappresentazione grafica è fuorviante.

Si possono anche notare degli errori di coerenza dell'esposizione: l'"avvicinarsi" del grafico della funzione all'asse delle ascisse che non coincide con il grafico tracciato né con la correttezza matematica, in quanto in realtà la funzione radice quadrata continua a crescere all'infinito allontanandosi da entrambi gli assi; la "decrescenza" di una funzione intorno a un punto di minimo, infatti, una funzione non può decrescere in un punto di minimo, in quanto non può raggiungere valori inferiori al suo minimo.

Un'altra osservazione va fatta sulla riformulazione della derivata seconda: le risposte che GPT ha dato sono entrambe corrette; tuttavia, l'AI non si è minimamente resa conto del fatto che le due espressioni erano matematicamente equivalenti, anche se scritte in modo diverso, e quindi le ha interpretate come se fossero distinte.

3.1.2 CALCOLO DI UN INTEGRALE, GPT-3.5

In questo esperimento si è provato a far risolvere a ChatGPT un *integrale indefinito*. Il test è composto di due *chat* distinte: la motivazione dietro a questa scelta risiede nel fatto che la prima risposta era completamente errata e quindi si è voluto provare a riproporre la stessa domanda dopo qualche minuto in un'altra *chat*. Di seguito si riporta la prima *chat*.

“Domanda: Calcola il seguente integrale indefinito: $\int \frac{1}{1+x^2} dx$.”

“Risposta: Questo integrale può essere valutato usando una sostituzione trigonometrica. Procediamo con il calcolo. Useremo la sostituzione $x = \tan(\theta)$.

Questo ci dà $dx = \sec^2(\theta)d\theta$ e $\frac{1}{1+x^2} = (\cos(\theta))^2$. Quindi l'integrale diventa: $\int (\cos(\theta))^2 (\sec(\theta))^2 d\theta$. Adesso possiamo semplificare l'integranda usando l'identità trigonometrica $(\sec(\theta))^2 = 1 + (\tan(\theta))^2$ nel modo seguente: $\int 1 + (\tan(\theta))^2 d\theta$. Integrando termine a termine otteniamo: $\int 1 d\theta + \int (\tan(\theta))^2 d\theta$. Il primo integrale è semplicemente θ mentre per il secondo possiamo usare l'integrale della funzione tangente al quadrato: $\int (\tan(\theta))^2 d\theta = \tan(\theta) - \theta$. Quindi globalmente l'integrale

diventa $\theta + (\tan(\theta) - \theta) + C$, dove C è la costante di integrazione. Adesso, ricordando che $x = \tan(\theta)$, quindi $\theta = \arctan(x)$ e sostituendo nel passaggio precedente si ottiene: $\arctan(x) + (\tan(\arctan(x)) - \arctan(x)) + C$. Dato che $\tan(\arctan(x)) = x$, otteniamo $\arctan(x) + (x - \arctan(x)) + C$. Semplificando ulteriormente: $x + C$. Quindi, l'integrale indefinito di $\frac{1}{1+x^2}$ rispetto alla x è $x + C$, dove C è la costante di integrazione.”

“**D:** Manca il $(\cos(\theta))^2$ nella semplificazione dell'integranda.”

“**R:** Hai assolutamente ragione, e mi scuso per la svista. Correggiamo quell'errore e continuiamo con il conto. Eravamo arrivati al punto: $\int (\cos(\theta))^2 (\sec(\theta))^2 d\theta$. Per semplificare l'integranda usiamo l'identità trigonometrica $(\sec(\theta))^2 = 1 + (\tan(\theta))^2$ ma abbiamo anche bisogno di riscrivere $(\cos(\theta))^2$ in termini del $(\sin(\theta))^2$ usando l'identità $(\cos(\theta))^2 = 1 - (\sin(\theta))^2$: $\int (1 - (\sin(\theta))^2)(1 + (\tan(\theta))^2) d\theta$. Adesso distribuiamo e semplifichiamo: $\int 1 - (\sin(\theta))^2 + (\tan(\theta))^2 - (\sin(\theta))^2 (\tan(\theta))^2 d\theta$ [...] Il risultato finale sarà una combinazione di θ , $\tan(\theta)$ e funzioni logaritmiche di $\cos(\theta)$. Di nuovo, mi scuso per l'errore precedente e apprezzo la tua pazienza.”

Osserviamo che per il calcolo di un integrale indefinito bisogna ricercare una funzione la cui derivata è la funzione integranda e ci sono diversi metodi per farlo. In questo caso è utile effettuare un cambio di variabile e quello suggerito dall'AI è corretto; successivamente, avvalendosi di un'identità trigonometrica corretta, vengono svolte delle sostituzioni nell'integranda e la proprietà di linearità dell'integrale permette di procedere nella computazione di integrali più semplici, che poi attraverso sostituzioni danno la soluzione finale a meno di una costante. In queste computazioni, viene fatto notare all'AI un errore di conto nella prima semplificazione che viene corretto prontamente. Nonostante ciò, la soluzione a cui ChatGPT giunge alla fine è completamente errata (alcuni passaggi di mero conto sono stati rimossi per evitare di dilungarsi troppo). Di seguito si riporta il secondo tentativo.

“**Domanda:** Calcola $\int \frac{1}{1+x^2} dx$.”

“**Risposta:** Per calcolare l’integrale possiamo usare la sostituzione dell’arcotangente. Sia $u = x$, allora $du = dx$. L’integrale diventa $\int \frac{1}{1+x^2} dx = \int \frac{1}{1+u^2} du$. Adesso possiamo integrare rispetto a u : $\int \frac{1}{1+u^2} du = \arctan(u) + C$. Sostituendo la u con la x , otteniamo: $\int \frac{1}{1+x^2} dx = \arctan(x) + C$. Quindi l’integrale indefinito di $\frac{1}{1+x^2}$ rispetto alla x è $\arctan(x) + C$, dove C è la costante di integrazione.”

“**D:** Nota che questa sostituzione non cambia nulla.”

“**R:** Chiedo scusa per la confusione. Hai ragione, la sostituzione che ho suggerito inizialmente non cambia l’integrale. Usiamo un approccio differente, come una sostituzione trigonometrica.

Procediamo con $x = \tan(\theta)$, da cui si ottiene $dx = (\sec(\theta))^2 d\theta$. Adesso, l’integrale diventa $\int \left(\frac{1}{1+(\tan(\theta))^2} \right) \cdot (\sec(\theta))^2 d\theta$. Semplificando il denominatore usando l’identità trigonometrica $(\tan(\theta))^2 + 1 = (\sec(\theta))^2$, abbiamo

$$\int \left(\frac{1}{(\sec(\theta))^2} \right) \cdot (\sec(\theta))^2 d\theta = \int d\theta = \theta + C$$

Sostituendo x con $x = \tan(\theta)$, otteniamo $\theta + C = \arctan(x) + C$. Quindi, l’integrale indefinito ha come risultato $\arctan(x) + C$, dove C è la costante di integrazione.”

Osserviamo che in questo secondo caso, nonostante il cambio di variabile proposto in principio sia sostanzialmente inutile (si tratta semplicemente di cambiare la lettera che dà il nome alla variabile), si è riusciti a sottolineare l’errore e ChatGPT, dopo essersi corretta, giunge alla stessa soluzione proponendo un approccio diverso. Si può notare come lo svolgimento preveda dei passaggi simili a quelli visti nella prima chat ma la correzione ha permesso di partire da un’assunzione iniziale corretta e quindi conseguentemente di effettuare dei passaggi logici giusti.

Più nel dettaglio, nella prima chat la sostituzione iniziale era corretta ma l’errore di conto dato dalla dimenticanza del termine $(\cos(\theta))^2$ ha portato, nello svolgimento dei conti, a una risposta sbagliata. Cosa che invece non è accaduta nel secondo tentativo, dove la correzione proposta è stata effettuata e i conti sono invece stati svolti correttamente.

3.1.3 CALCOLO DI UN LIMITE, GPT-3.5

In questo esperimento vengono testate le capacità ChatGPT di valutare un *limite*. Un possibile svolgimento corretto richiede un cambio di variabile e un limite notevole.

“**Domanda:** Calcola $\lim_{x \rightarrow +\infty} x^2(\arcsin(\frac{1}{x}) - \arcsin(\frac{1}{x^2}))$.”

“**Risposta:** Per calcolare il limite dell’espressione data possiamo usare alcune proprietà trigonometriche e manipolazioni algebriche. Vediamo il procedimento passo a passo. Prima di tutto ci concentriamo sulla differenza tra arcoseni nelle parentesi. Usando l’identità trigonometrica

$$\arcsin(a) - \arcsin(b) = \arcsin(a\sqrt{1-b^2})$$

possiamo semplificare l’espressione:

$$\arcsin\left(\frac{1}{x}\right) - \arcsin\left(\frac{1}{x^2}\right) = \arcsin\left(\frac{1}{x}\sqrt{1-\left(\frac{1}{x^2}\right)^2}\right) = \arcsin\left(\frac{1}{x}\sqrt{1-\frac{1}{x^4}}\right).$$

Quando x tende a $+\infty$, $\frac{1}{x}$ tende a 0 e $\frac{1}{x^4}$ tende a 0 più velocemente di $\frac{1}{x^2}$. Quindi,

possiamo approssimare $\sqrt{1-\frac{1}{x^4}}$ con 1 per x grande. Questo significa:

$$\arcsin\left(\frac{1}{x}\sqrt{1-\frac{1}{x^4}}\right) \approx \arcsin\left(\frac{1}{x} \cdot 1\right) = \arcsin\left(\frac{1}{x}\right).$$

Adesso, sostituendo questo nell’espressione originale si ottiene

$$\lim_{x \rightarrow +\infty} x^2(\arcsin\left(\frac{1}{x}\right) - \arcsin\left(\frac{1}{x^2}\right)) \approx \lim_{x \rightarrow +\infty} x^2 \cdot \arcsin\left(\frac{1}{x}\right).$$

Quando x tende a $+\infty$, $\frac{1}{x}$ tende a 0 e $\arcsin(0) = 0$. Dunque, il limite diventa

$\lim_{x \rightarrow +\infty} x^2 \cdot 0 = 0$ e quindi il limite dell’espressione data quando x tende a $+\infty$ è 0.”

“**D:** Forse una strategia migliore potrebbe essere sostituire $\frac{1}{x}$ con t , in modo tale che quando x tende a $+\infty$, t tende a 0.”

“**R:** Assolutamente, questo è un ottimo approccio per semplificare l’espressione e valutare il limite. Seguiamo il tuo suggerimento e usiamo la sostituzione $t = \frac{1}{x}$.

Quando x va a infinito, t tende a 0. Quindi sostituendo in

$$\lim_{x \rightarrow +\infty} x^2 \left(\arcsin\left(\frac{1}{x}\right) - \arcsin\left(\frac{1}{x^2}\right) \right)$$

otteniamo $\lim_{t \rightarrow 0} \frac{1}{t^2} \cdot (\arcsin(t) - \arcsin(t^2))$. Quando t tende a 0, t^2 tende a 0 più velocemente; quindi, possiamo effettuare l'approssimazione $\arcsin(t^2) \approx \arcsin(0) = 0$.

[Si noti che il programma ha sostituito erroneamente $\arcsin(0) = 0$ nell'espressione del limite]. Allora l'espressione diventa $\lim_{t \rightarrow 0} \frac{1}{t^2} (\arcsin(t) - \arcsin(0)) = \lim_{t \rightarrow 0} \frac{\arcsin(t)}{t^2}$.

Questa è una forma indeterminata $\frac{0}{0}$. Possiamo usare la regola di de l'Hôpital per

valutare questo limite: $\lim_{t \rightarrow 0} \frac{\arcsin(t)}{t^2} = \lim_{t \rightarrow 0} \frac{\frac{d}{dt} \arcsin(t)}{\frac{d}{dt} (t^2)}$. Derivando si ottiene $\lim_{t \rightarrow 0} \frac{1}{2t \sqrt{1-t^2}}$.

Mentre t tende a 0, l'espressione diventa ∞ . Quindi il risultato finale del limite è

$$\lim_{x \rightarrow +\infty} x^2 \left(\arcsin\left(\frac{1}{x}\right) - \arcsin\left(\frac{1}{x^2}\right) \right) = \infty."$$

Il modo in cui l'AI approccia inizialmente l'esercizio è corretto e anche l'attenzione alla spiegazione dei passaggi è soddisfacente. Tuttavia, possiamo notare immediatamente che la prima soluzione proposta è sbagliata: questo perché si avvale di un'identità non vera (utilizzandola inoltre senza averne provato la validità), giungendo così a una conclusione errata.

Nonostante le varie sostituzioni siano giuste, si può notare inoltre un'altra criticità nei passaggi intermedi della prima soluzione proposta: l'approssimazione dell'arcoseno è formalmente non corretta, difatti non si può sostituire il valore della variabile all'interno del limite usando la giustificazione "tende più velocemente".

Ciò nonostante, dopo aver accolto il nostro suggerimento, il risultato a cui giunge alla fine ChatGPT è corretto, in quanto anche qui svolge i conti senza errori, dilungandosi fin troppo. Si noti inoltre come l'esercizio avrebbe potuto essere svolto in meno passaggi e senza l'utilizzo della regola di de l'Hôpital.

3.2. ESPERIMENTI CON QUESITI DIMOSTRATIVI

3.2.1 LA CONTROIMMAGINE DI UN INSIEME CHIUSO È CHIUSA, GPT-3

Per iniziare a discutere gli esperimenti dimostrativi si è scelta una dimostrazione base di Analisi 1 che verte sulla conoscenza dei concetti di *continuità*, *successioni* e *chiusura* di un insieme e che richiede l'abilità di manipolare tali concetti per svolgere l'esercizio.

“Domanda: Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ una funzione continua. Prova che se C è un sottoinsieme chiuso di \mathbb{R} allora la sua *controimmagine* è chiusa.”

“Risposta: Certamente! Per provare che la controimmagine di un insieme chiuso attraverso una funzione continua è chiusa possiamo usare il criterio sequenziale per insiemi chiusi. Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ una funzione continua e sia C un sottoinsieme chiuso di \mathbb{R} . Dobbiamo provare che la controimmagine di C , denotata con $f^{-1}(C)$, è anche chiusa. Supponiamo che x_n sia una successione in $f^{-1}(C)$ che converge a qualche punto limite x in \mathbb{R} . Dobbiamo provare che anche x è in $f^{-1}(C)$. Dato che x_n è in $f^{-1}(C)$, sappiamo che $f(x_n)$ è in C per ogni n . Dal momento che C è chiuso, ogni successione convergente in C converge a un punto di C . Quindi la successione $(f(x_n))$ converge a un punto y di C . [nonostante ChatGPT non lo sottolinei, $(f(x_n))$ converge perché x_n è convergente per ipotesi e f è continua]. Dato che f è continua, sappiamo che $f(x_n)$ converge a $f(x)$ quando n tende a infinito. Dato che i limiti delle successioni sono unici, abbiamo $y = f(x)$. Dato che y appartiene a C e C è chiuso, sappiamo che y è un punto di limite di C , il che significa che esiste una successione (y_n) in C che converge a y . Allora, per la continuità di f , la successione $(f^{-1}(y_n))$ è una successione in $f^{-1}(C)$ che converge a x .

Dato che i limiti delle successioni sono unici, abbiamo $x = \lim_{n \rightarrow \infty} f^{-1}(y_n)$, il che significa che x è un punto di limite di $f^{-1}(C)$. Quindi, $f^{-1}(C)$ è chiuso. In conclusione, abbiamo provato che la controimmagine di un insieme chiuso attraverso una funzione continua è chiusa.”

Osserviamo che, a un primo impatto, questa risoluzione potrebbe sembrare corretta: usare il criterio sequenziale per insiemi chiusi è una buona idea e inoltre la tesi da

dimostrare è esposta in maniera chiara, cosa fondamentale per capire cosa si debba dimostrare. Anche i passaggi logici effettuati sono ben motivati e seguono sostanzialmente dall'applicazione di definizioni. Tuttavia, l'esercizio si sarebbe potuto concludere alla riga «Dato che i limiti delle successioni sono unici, abbiamo $y = f(x)$ », soltanto aggiungendo che se $f(x) \in C$ allora $x \in f^{-1}(C)$ (per definizione di immagine di una funzione). La questione è che, dopo quella affermazione, ChatGPT non comprende che la dimostrazione è completata e si perde nel tentativo di arrivare a una conclusione aggiungendo informazioni che non sono rilevanti.

3.2.2 INTEGRALE DEFINITO DI UNA FUNZIONE CONTINUA PARI, GPT-3

L'esercizio che segue è un'altra richiesta di dimostrazione “da esame”, che diversamente dalla precedente tocca gli argomenti di *continuità* e *parità di una funzione*, oltre che al concetto di *integrale*. Si può notare come questa sia la soluzione più breve tra quelle proposte dall'AI.

“**Domanda:** Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ una funzione continua pari. Prova che se $\int_{-\pi}^{\pi} f(x)dx = 0$ allora $\int_0^{\pi} f(x)dx = 0$.”

“**Risposta:** Dato che f è una funzione pari, abbiamo che $f(-x) = f(x)$ per ogni x in \mathbb{R} . Quindi possiamo riscrivere l'integrale di f nell'intervallo $[-\pi, \pi]$ come segue:

$$\int_{-\pi}^{\pi} f(x)dx = \int_0^{\pi} f(x)dx + \int_{-\pi}^0 f(x)dx.$$

Usando la sostituzione $u = -x$, abbiamo $\int_{-\pi}^0 f(x)dx = -\int_0^{\pi} f(-u)du$. Dato che f è continua e pari, abbiamo che $f(-u) = f(u)$ per ogni u in \mathbb{R} . Allora, possiamo riscrivere l'integrale precedente così $\int_{-\pi}^0 f(x)dx = -\int_0^{\pi} f(u)du$. Sommando le ultime due equazioni, otteniamo $\int_{-\pi}^{\pi} f(x)dx = 2 \int_0^{\pi} f(x)dx$. Ora, dato che l'integrale definito di f in $[-\pi, \pi]$ è zero, abbiamo $\int_{-\pi}^{\pi} f(x)dx = 0$. Sostituendo questo nella equazione precedente abbiamo $2 \int_0^{\pi} f(x)dx = 0$. Dividendo entrambi i membri per 2, otteniamo $\int_0^{\pi} f(x)dx = 0$. Quindi, abbiamo provato che se f è una funzione continua

pari e l'integrale definito di f in $[-\pi, \pi]$ è zero, allora l'integrale definito di f nell'intervallo $[0, \pi]$ è anche zero.”

Come nel caso precedente, la soluzione sembra accettabile, dal momento che l'ipotesi di parità di f è applicata correttamente, scrivendo anche esplicitamente la definizione, i passaggi sono tutto sommato ben spiegati; infatti, si tratta sostanzialmente di sostituzioni e applicazione di definizioni e delle ipotesi di parità della funzione e del valore dell'integrale, e la tesi viene dimostrata alla fine. Nonostante ciò, è presente un errore di segno nel primo passaggio di sostituzione, $u = -x$, che viene completamente ignorato dall'AI e inoltre non ci sono ulteriori spiegazioni su cosa accada nei passaggi contenuti in «Sommando le ultime due equazioni», in cui ChatGPT riporta direttamente il risultato.

3.2.3 FUNZIONE DIFFERENZIABILE n -VOLTE, GPT-3.5

Per concludere, abbiamo chiesto a ChatGPT un esercizio che per essere risolto sfrutti il *principio di induzione* e un teorema fondamentale di Analisi 1: il *Teorema di Rolle*.

“**Domanda:** Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ una funzione n -volte differenziabile ($n \geq 1$) con esattamente $n + 1$ zeri. Prova che esiste almeno un punto in cui la n -esima derivata è zero.”

“**Risposta:** Per provare che esiste almeno un punto dove l' n -esima derivata è zero possiamo usare il Teorema di Rolle e l'induzione.

Teorema di Rolle: Se una funzione f è continua in un intervallo chiuso $[a, b]$ e differenziabile in (a, b) e se $f(a) = f(b)$, allora esiste almeno un punto c in (a, b) dove la derivata $f'(c) = 0$.

Dimostrazione per induzione:

Caso base $n = 1$: Dato che la funzione f ha esattamente $n + 1 = 2$ zeri, sappiamo che ha almeno un punto x_1 dove $f(x_1) = 0$. Adesso applichiamo il Teorema di Rolle agli intervalli $[a, x_1]$ e $[x_1, b]$, dove a e b sono gli estremi dell'intervallo dove f è definita. In ciascun intervallo, la funzione soddisfa le condizioni del teorema di Rolle, quindi esistono i punti c_1 e c_2 dove $f'(c_1) = 0$ e $f'(c_2) = 0$.

Passo induttivo: Ora assumiamo che il teorema sia vero per $n = k$, il che significa che per una funzione f con $k + 1$ zeri esiste almeno un punto dove la k -esima derivata è zero.

Consideriamo una funzione f con $k + 2$ zeri. Sappiamo che se f ha almeno un punto x_1 dove $f(x_1) = 0$, possiamo applicare il Teorema di Rolle all'intervallo $[a, x_1]$ per trovare un punto c_1 dove $f'(c_1) = 0$. Adesso consideriamo la derivata f' . Questa funzione ha $k + 1$ zeri (perché ciascuno zero di f diventa un punto critico di f'). Per l'ipotesi induttiva, sappiamo che esiste almeno un punto dove la derivata seconda $f''(c_2) = 0$ (dove f'' è la derivata seconda di f).

Questo procedimento può essere ripetuto fino alla derivata n -esima e dato che abbiamo un numero finito di derivate, alla fine raggiungiamo il caso base $n = 1$, dove abbiamo mostrato che esiste un punto dove l' n -esima derivata è zero.

Quindi, per induzione, abbiamo provato che per una funzione f con esattamente $n + 1$ zeri esiste almeno un punto dove l' n -esima derivata è zero.”

Si può notare come GPT-3.5, rispetto alla precedente versione GPT-3, sia molto più prolisso e a tratti possa risultare ridondante.

Anche in questo caso, il ragionamento sembra valido e chiaro, ma, a un occhio più attento, manca di coerenza e correttezza. Difatti c'è un errore nel caso base, nel quale si assume che la funzione abbia due zeri, ma il Teorema di Rolle viene applicato a due intervalli, come se ci fosse un terzo punto x_1 in cui la funzione è zero. Inoltre, ChatGPT afferma di applicare il Teorema di Rolle all'intervallo di definizione della f ma per ipotesi f è definita su tutto \mathbb{R} . L'altro errore è più sottile, e sta nel passo induttivo. Difatti nella dimostrazione del passo induttivo, GPT non usa l'ipotesi induttiva, bensì soltanto nuovamente il Teorema di Rolle, rendendo difatti la sua dimostrazione una dimostrazione non per induzione, ma solo spacciata per tale.

4. CONCLUSIONI

Per riassumere, dopo aver spiegato brevemente il funzionamento di ChatGPT abbiamo voluto testare le sue abilità in un campo che è tutto sommato differente

dalla mera generazione del testo, ovvero *computazione e deduzione matematica*.

Come risultato, dalle risposte che abbiamo analizzato, possiamo comprendere come GPT-3 e GPT-3.5 non siano i migliori strumenti per questo tipo specifico di *task*: non possono essere utilizzati in modo incontrollato e superficiale, ma hanno bisogno di suggerimenti e correzioni in input per poter svolgere un esercizio in maniera quantomeno accettabile, come nel caso della maggior parte degli esperimenti computazionali; compiono diversi errori di conto e molto spesso non sono in grado di accorgersene né di correggerli e inoltre usano identità senza provare la loro validità ed effettuando così ragionamenti logici fallaci, come nel caso del calcolo del limite; infine, il modo in cui vengono date le risposte è ben scritto e risulta talmente convincente, da essere persuasi che non ci siano errori. Questo si nota soprattutto negli esperimenti con quesiti di tipo dimostrativo.

Si può dunque sottolineare come in un contesto didattico e di apprendimento della matematica lo strumento dell'intelligenza artificiale debba essere utilizzato con cautela e in nessun modo possa sostituire un libro di testo o una lezione tenuta da un docente, soprattutto data l'assenza di fonti per verificare la validità delle risposte generate. Tuttavia, invece di vietarne l'utilizzo, cosa che si rivelerebbe controproducente, si potrebbe ad esempio chiedere a GPT lo svolgimento di un calcolo e successivamente far trovare agli studenti gli eventuali errori: per questo tipo di esercizio è necessaria una buona conoscenza della teoria ed esso ha il pregio di far avvicinare allo strumento in maniera guidata. Inoltre, data la capacità dialettica di ChatGPT dovuta alla sua natura intrinseca di modello NLP, si potrebbe eventualmente anche utilizzarlo per far capire agli studenti come scrivere in matematica, analizzando le risposte a quesiti del tipo di quelli qui presentati.

Dobbiamo però considerare, infine, che il lavoro qui presentato si colloca in una fase di evoluzione tecnologica che viene aggiornata in lassi di tempo molto brevi: si può ragionevolmente supporre che nel futuro altri modelli più performanti (come, ad esempio, GPT-4) possano offrire migliori prestazioni in questo specifico ambito e

diventare ausili migliori per la didattica della matematica.

BIBLIOGRAFIA

RADFORD A., NARASIMHAN K., SALIMANS T., SUTSKEVER I.

2018, *Improving Language Understanding by Generative Pre-Training*, OpenAI.

VASWANI A., SHAZEER N., PARMAR N., USZKOREIT J., JONES L., GOMEZ A. N., KAISER Ł., POLOSUKHIN I.

2017, *Attention is all you need*, in: I. GUYON, U. VON LUXBURG, S. BENGIO, H. WALLACH, R. FERGUS, S. VISHWANATHAN, R. GARNETT (a cura di), *Advances in Neural Information Processing Systems (NIPS 2017)*, pp. 6000-6010.

PER APPROFONDIRE

ALAMMAR J.

The Illustrated Transformer tutorial, <<https://jalammar.github.io/illustrated-transformer/>>.

FLEURET F.

2023, *The Little Book of Deep Learning*, Genève, Université de Genève, Version beta-2023.05.03.

GIUSTI E.

2002, *Analisi 1*, Torino, Bollati Boringhieri

PAGANI C. D., SALSA S.

2015, *Analisi Matematica 1*, Seconda Edizione, Bologna, Zanichelli.

RUSSELL S. J., NORWIG P.

2010, *Artificial intelligence. A modern approach*. Third edition, Upper Saddle River (New Jersey) (USA), Pearson Education, inc.

La crittografia classica come occasione di ragionamento matematico*

FABIO PASTICCI
 Dipartimento di filosofia,
 scienze sociali, umane e della formazione
 Università di Perugia
 fabio.pasticci@unipg.it

ABSTRACT

Cryptography can be a tool that, enables students to consolidate knowledge and develop mathematical skills if used as a recreational or playful activity. In this paper we introduce Caesar's cryptosystem, which is one of the simplest encryption schemes. It is very easy to crack the encryption of this system: a brute force attack allows easy recovery of the key and the plaintext. We made a simple modification that allows us to have a greater number of keys, in order to make such a brute force attack infeasible. Finally, we show a statistical method that allows us to force this cryptosystem.

PAROLE CHIAVE

CRITTOSISTEMA DI CESARE / CAESAR CRYPTOSYSTEM; CRITTOGRAFIA / CRYPTOGRAPHY; CRITTOANALISI / CRYPTOANALYSIS; ANALISI DELLE FREQUENZE / FREQUENCY ANALYSIS; DIDATTICA DELLA MATEMATICA / MATHEMATICS EDUCATION.

1. INTRODUZIONE

La *crittografia*, ovvero l'arte e la scienza di trasmettere messaggi facendo in modo che solo il legittimo destinatario sia in grado di leggerli, può essere utilizzata in classe come strumento utile per il recupero di abilità operative e il consolidamento di conoscenze matematiche. Al fine di rendere meno "astratto" lo studio della matematica è possibile trovare schemi crittografici alla cui base ci sono elementi matematici che sono collegati con i bisogni formativi delle singole classi: ad esempio già in una classe terza della Scuola primaria è possibile presentare un *cifrario di Cesare* e far sperimentare

* Title: Classical cryptography as an opportunity for mathematical reasoning.

ai ragazzi la tecnica per cifrare e decifrare i messaggi, mentre già in una classe quarta, sempre della Scuola primaria, è possibile mostrare alcuni collegamenti tra il cifrario di Cesare e l'aritmetica modulare.

Nello specifico si può evidenziare il fatto che spostare di alcuni posti le lettere dell'alfabeto significa addizionare numeri naturali con l'accortezza che quando si arriva all'ultima lettera dell'alfabeto è necessario ripartire dalla prima. Questo, in termini numerici, si traduce nel seguente modo: se si ha un alfabeto con 26 lettere l'addizione 26 più 1 avrà come risultato 1.

Inoltre la crittografia

consente di creare collegamenti con altre discipline, quali la storia (dai metodi antichi di cifratura e decifratura di codici segreti fino al suo utilizzo durante la Seconda Guerra Mondiale), la letteratura (utilizzando testi famosi come esempi di cifratura e decifratura) e l'educazione civica (ad esempio ragionando sulle frodi digitali e sulla necessità di limitare la condivisione di dati personali).¹

Ovviamente esistono anche altri metodi per trasmettere messaggi segreti: scrivere il testo con l'inchiostro simpatico, affidarsi a un messaggero fidato, dividere la comunicazione in più parti per poi inviarle utilizzando canali di trasmissione diversi, solo per citarne alcuni. Quindi, perché si preferisce la crittografia?

La risposta è semplice [...] la matematica fornisce, almeno all'ingrosso, la giustificazione teorica per la forza di un particolare algoritmo o protocollo.²

Un'altra motivazione molto rilevante dal punto di vista didattico è che

cryptography is a charming and rewarding way to introduce into the classroom subjects of traditional or less traditional mathematics.³

In questo lavoro si illustra il *crittosistema* (o *cifrario*) di Cesare e l'attacco esaustivo che è facilmente realizzabile e che permette di risalire alla chiave e al testo in chiaro in breve tempo. Si illustra quindi una modifica al suddetto schema di cifratura che rende inutile l'attacco a forza bruta. Poi viene presentata una ulteriore tipologia di attacco che permette di forzare questo crittosistema derivato da quello di Cesare.

¹ Cfr. CAZZOLA, GRAZIAN 2021.

² Cfr. BERARDI, BEUTELSPACHER 1996.

³ Cfr. BORRELLI, FIORETTO, SGARRO, ZUCCHERI 2002.

In questo percorso viene dato spazio alla concretizzazione del concetto di *funzione* e all'introduzione di nozioni di base di *aritmetica modulare* che, opportunamente presentata, come è stato già visto negli obiettivi collaterali dell'attività di cifrazione e di decifrazione descritti da Zuccheri⁴, amplia quella che può essere definita la *matematica dell'orologio*.

Occorre mettere in atto tutte le strategie possibili, in modo da far scaturire il tutto da attività ludiche da presentare ai ragazzi nell'ottica di far

*acquistare confidenza con la matematica, riuscire ad unire un argomento all'altro, in maniera continua, senza interruzioni, in modo da costruire una specie di «filo di Arianna» che permetta di percorrere il labirinto delle proprie conoscenze matematiche con una certa dimestichezza, senza bruschi salti nel buio.*⁵

Per poter illustrare i vari aspetti della crittografia si precisa qui la terminologia con alcune definizioni.

Si chiama *schema crittografico*, *cifrario* o *crittosistema* il raggruppamento dell'insieme dei possibili *testi in chiaro*, dei corrispondenti *testi cifrati*, degli *algoritmi* (di *cifratura* e di *decifratura*) e delle *chiavi*. Il *testo in chiaro* è la sequenza di lettere, numeri o simboli che si vogliono trasmettere. La *cifratura* è l'operazione che trasforma il *testo in chiaro* in *testo cifrato*. Questa azione viene posta in essere per mezzo di una *chiave segreta* cioè di una informazione che deve essere nota solamente al mittente e al destinatario del messaggio. L'operazione inversa della *cifratura* si chiama *decifratura* se eseguita dal legittimo destinatario per mezzo della chiave e si chiama invece *decrittazione* o *crittoanalisi* se è svolta da un intruso che tenta di risalire al testo in chiaro senza conoscere la chiave. Da quanto appena detto appare evidente che la crittografia permette di creare situazioni didattiche in cui il docente «cattura l'attenzione e rende piacevole fare matematica»⁶. Inoltre

*l'osservazione di processi e competenze permette di riconoscere molteplici gradualità, dando all'insegnante la possibilità di apprezzare i progressi e le potenzialità di ogni allievo.*⁷

⁴ Cfr. ZUCCHERI, 1992.

⁵ Cfr. PERES 1986.

⁶ Cfr. D'AMORE 1999.

⁷ Cfr. DI MARTINO, ZAN 2020.

2. IL CIFRARIO DI CESARE

Il *cifrario di Cesare* è uno dei più semplici schemi crittografici: può essere presentato utilizzando pochi strumenti teorici e soprattutto permette di vedere la «matematica come strumento per indagare e descrivere la realtà»⁸.

Si tratta di uno dei crittosistemi che vengono definiti *simmetrici*, o anche *classici*, poiché la chiave utilizzata dal mittente per cifrare il messaggio è la stessa che permette al destinatario di risalire al testo in chiaro partendo dal testo cifrato che ha ricevuto.

Svetonio, in *Vite dei Cesari* scrive:

*si qua occultius perferenda erant, per notas scripsit, id est sic structo litterarum ordine, ut nullum uerbum effici posset: quae si qui inuestigare et persequi uelit, quartam elementorum litteram, id est D pro A et perinde reliquas commutet.*⁹

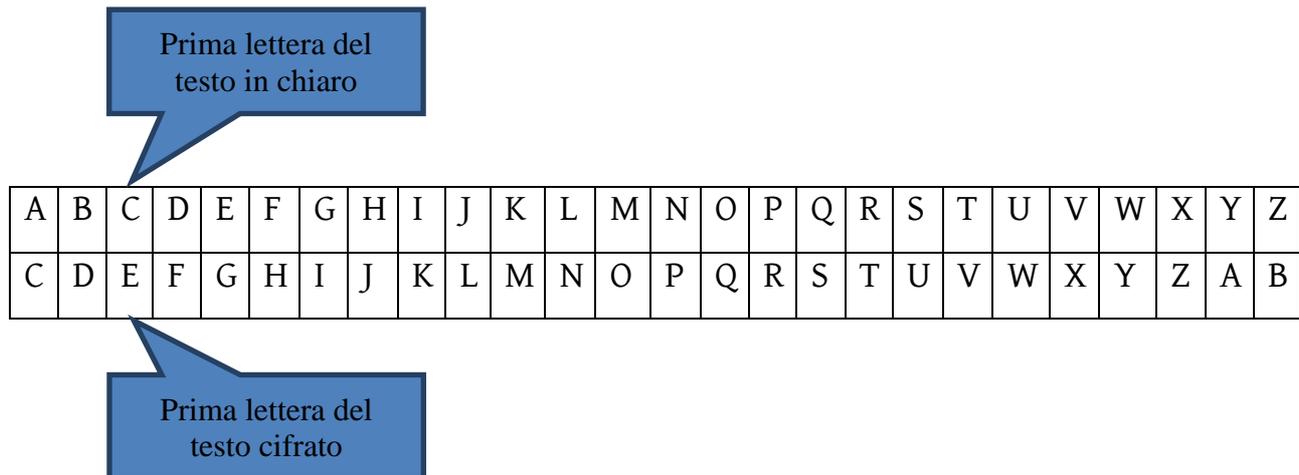
La cifratura di un messaggio, mettendo in pratica quanto appena descritto, avviene nel modo seguente:

- si prende una tabella con due righe e tante colonne quante sono le lettere dell'alfabeto che si considera (l'esempio nella Tabella 1 utilizza l'alfabeto italiano con 26 lettere);
- si scrive sulla prima riga l'intero alfabeto e sulla seconda lo si va a riscrivere iniziando da una lettera diversa dalla lettera A. Quando si arriva alla lettera Z, si continua ripartendo dalla lettera A e proseguendo in ordine fino ad arrivare all'ultima casella disponibile. Questo si può anche definire uno scorrimento circolare dell'alfabeto.

La cifratura avviene utilizzando la seguente procedura. Ogni lettera del testo in chiaro viene cercata sulla prima riga della tabella. La lettera che si trova sulla seconda riga, sotto la lettera trovata, costituisce la corrispondente lettera cifrata. Ad esempio se si deve cifrare la parola “CIAO” si inizia cercando la lettera C sulla prima riga.

⁸ Cfr. SABENA, FERRI, MARTIGNONE, ROBOTTI 2019.

⁹ Cfr. SUETONIUS TRANQUILLUS 2016. «Se doveva comunicare informazioni in modo riservato, le scriveva in linguaggio cifrato, cioè seguendo una sequenza alfabetica disposta in modo tale che non se ne potesse ricavare nessuna parola di senso compiuto: se si desidera comprendere il testo bisogna sostituire la quarta lettera dell'alfabeto alla prima, cioè la D al posto della A e così tutte le altre».



A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B

Tabella 1. Crittosistema di Cesare con chiave A --> C.

Quindi la prima lettera del testo cifrato sarà la lettera E.

Poi si cerca la lettera I sempre sulla prima riga e quindi la lettera K sarà la seconda lettera del testo cifrato. Si procede allo stesso modo per la lettera A e per la lettera O. Quindi “EKCQ” è la parola cifrata che corrisponde alla parola in chiaro “CIAO”.

La chiave di cifratura è il numero di posizioni di cui viene spostata la lettera A dell’alfabeto nella seconda riga. Il destinatario del messaggio, che è a conoscenza della chiave, riceve la parola cifrata “EKCQ”.

Come affermato da Zuccheri¹⁰, questa procedura «fornisce l’occasione di far lavorare i bambini con un’applicazione biunivoca (*cifratura*) e con la sua inversa (*decifrazione*)».

La decifrazione del messaggio avviene nel modo seguente: si prende la stessa Tabella 1 utilizzata per la cifratura del messaggio e ogni lettera del testo cifrato viene cercata sulla seconda riga e sostituita nel corrispondente testo in chiaro dalla lettera che si trova sopra la lettera trovata.

Nel caso in questione si cerca la lettera E sulla seconda riga e si sostituisce con la lettera C, poi si cerca la K sempre sulla seconda riga e si sostituisce con la lettera I poi analogamente per la lettera C e per la lettera Q. Quindi si ottiene la parola “CIAO” che è il testo in chiaro di partenza.

¹⁰ Cfr. ZUCCHERI 1992.

Dopo aver presentato agli alunni – ad esempio, di una classe terza della Scuola primaria – il cifrario di Cesare, si possono organizzare giochi in aula, anche a squadre, in cui vince chi riesce a decifrare un messaggio prima degli altri fornendo la Tabella 1 tutta compilata. Prima di passare alla descrizione delle attività legate ai procedimenti che un intruso può attuare per forzare il crittosistema, possiamo notare che il cifrario di Cesare è di tipo monoalfabetico: «ogni lettera dell’alfabeto in chiaro è cifrata sempre con la stessa lettera dell’alfabeto segreto»¹¹. In altre parole, a lettere uguali tra loro nel testo in chiaro corrispondono lettere uguali tra loro nel testo cifrato.

Per chiarire quanto detto basta osservare che se si cifra la parola “LIBRI” con il cifrario di Cesare in Tabella 1 si ottiene “NKDTK” da cui si evince che le due lettere I del testo in chiaro vengono cifrate con due lettere K. In realtà questa caratteristica appena descritta è uno degli elementi che favorisce il lavoro dei crittoanalisti. È una delle vulnerabilità propria di tutti i cifrari monoalfabetici, compreso lo schema crittografico di Cesare.

3. CRITTOANALISI

Successivamente, dopo aver lavorato su cifratura e decifratura dei messaggi, si propone quindi ai ragazzi un’ulteriore sfida: si mostra loro un testo cifrato, rendendo noto il fatto che è stato criptato utilizzando un cifrario di Cesare, senza comunicare la seconda riga della tabella (cioè si mantiene segreta solamente la *chiave*). A questo punto si invitano gli studenti a decrittare il messaggio.

Partendo dalle soluzioni proposte dai ragazzi si illustra il lavoro del crittoanalista. Occorre una premessa. In generale i crittoanalisti, cioè coloro che vogliono risalire al testo in chiaro senza essere i legittimi destinatari del messaggio, non sono sprovveduti. Spesso possiedono strumenti e capacità di un certo rilievo. Pertanto è opportuno chiedersi quali siano i componenti dello schema crittografico che devono rimanere segreti. Gli elementi candidati a non essere divulgati sono l’*algoritmo* e la *chiave*.

Il crittografo olandese Auguste Kerckhoffs ha evidenziato sei “desiderata” riguardanti

¹¹ Cfr. BERARDI, BEUTELSPACHER 1996.

gli schemi crittografici. In particolare il secondo di questi afferma che «Il faut qu'il [le système] n'exige pas le secret, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi»¹², cioè la sicurezza di uno schema crittografico deve dipendere solamente dalla segretezza della chiave e non dalla segretezza dell'algoritmo di cifratura. Quindi, tornando al gioco proposto ai ragazzi, si lavora sulle strategie proposte da loro per scoprire il testo in chiaro e quindi la chiave del crittosistema, rendendo noto in anticipo l'algoritmo di cifratura utilizzato e mantenendo segreta la chiave.

Dopo un brainstorming si tabulano le risposte e si raggruppano quelle che utilizzano tecniche simili tra loro. Uno dei metodi che generalmente gli studenti propongono è il cosiddetto “attacco a forza bruta”.

3.1 ATTACCO A FORZA BRUTA

Si definisce “attacco a forza bruta” quella procedura, utilizzata dal crittoanalista per risalire al testo in chiaro senza conoscere la chiave, che consiste nel provare tutte le possibili chiavi fino ad arrivare a quella che permette di decifrare il messaggio. Questo tipo di attacco viene chiamato anche *ricerca esaustiva*.

Questa strategia di crittoanalisi, applicata su messaggio che è stato cifrato con il crittosistema di Cesare, si attua mediante la seguente procedura. Si inizia a decifrare il messaggio supponendo che la chiave sia costituita dallo spostamento dell'alfabeto di una sola posizione cioè partendo dalla lettera B nella compilazione della seconda riga della tabella utilizzata per cifrare.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A

Tabella 2. Crittosistema di Cesare con chiave A --> B.

In caso di insuccesso si parte con le lettere successive cioè si prova a iniziare con la

¹² Cfr. KERCKHOFFS 1883, p. 12. «È necessario che [il sistema] non richieda di essere segreto, e che possa cadere senza pericoli nelle mani del nemico».

lettera C, poi D, fino a ottenere la tabella di cifratura giusta che permette di risalire al testo in chiaro. Si prenda ad esempio il testo cifrato dell'esempio precedente, cioè si vuole risalire al testo in chiaro corrispondente alla parola EKCQ.

Partendo dalla Tabella 2, si cerca la lettera E sulla seconda riga e si nota che sopra la E si trova la lettera D. Tale lettera D è la prima lettera del testo in chiaro che si sta ricostruendo. Poi si cerca la lettera K sempre sulla seconda riga e si nota che sopra c'è la lettera J quindi il testo in chiaro inizia con DJ. Già da queste prime lettere ci si accorge che DJ non è una possibile sillaba con cui può iniziare una parola italiana.

Poiché si ipotizza che il testo in chiaro sia una parola italiana, si intuisce che la chiave appena utilizzata è da scartare. Ovviamente, se si volesse proseguire con questa chiave, si otterrebbe un testo in chiaro formato da una parola che non ha significato. Per quanto detto, quindi, si cambia la chiave prendendo in considerazione uno spostamento di due posizioni, cioè si fa partire l'alfabeto della cifratura con la lettera C (cfr. Tabella 1).

Si ripete la stessa procedura e, nel caso in cui non fosse nemmeno questa la chiave, si continua facendo partire l'alfabeto cifrante dello schema crittografico (cfr. Tabella 1) dalla lettera D. In caso di insuccesso si sposta ulteriormente di una lettera l'avvio della procedura fino ad arrivare alla chiave corretta.

Si chiede quindi ai ragazzi di stabilire il numero massimo di tentativi che possono essere necessari per scoprire la chiave. Si coglie l'occasione per ricordare loro che in generale la bontà di un *algoritmo di crittoanalisi* si misura su quello che viene definito "il caso pessimo", cioè si considera il numero massimo di tentativi che un cracker potrebbe essere costretto a compiere per scoprire il messaggio segreto. Ovviamente se nell'esempio appena visto fosse stata utilizzata la Tabella 2 come schema di cifratura, il crittoanalista potrebbe scoprire la chiave con un solo tentativo.

Quindi, qual è questo numero massimo? La risposta corretta è $n-1$ dove n è il numero delle lettere dell'alfabeto utilizzato. Si sottrae 1 perché non si considera la chiave che cifra ogni lettera con se stessa.

4. IL CIFRARIO DI CESARE “RESO PIÙ SICURO”

Come è stato visto il cifrario di Cesare è uno schema di cifratura che può essere forzato molto velocemente, anche senza l’ausilio di strumenti tecnologici, mediante l’attacco a forza bruta. Si chiede quindi agli studenti se e come è possibile incrementare il livello di sicurezza di questo crittosistema.

Una strada che può essere percorsa, per rendere più difficile il lavoro del crittoanalista, è quella di cercare metodi che proteggano dall’attacco a forza bruta che è stato visto in precedenza. Quindi, per vanificare questo tipo di lavoro svolto dall’intruso, si possono “mescolare” le lettere che si scrivono sulla seconda riga della tabella utilizzata per la cifratura.

In questo modo, per conoscere la chiave non è più sufficiente sapere il numero di caselle dopo le quali si inizia a scrivere l’alfabeto partendo dalla lettera A nella seconda riga. Per poter cifrare i messaggi è necessario conoscere tutta la stringa di lettere che forma la seconda riga della tabella.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
E	H	L	N	C	G	P	I	O	S	A	Z	W	U	Y	F	M	B	V	K	R	X	D	J	T	Q

Tabella 3. Crittosistema di Cesare “reso più sicuro”.

Le procedure per cifrare e decifrare i messaggi rimangono le stesse del cifrario di Cesare descritto in precedenza: ad esempio se si deve cifrare la parola “CIAO” si inizia cercando la lettera C sulla prima riga. La prima lettera del testo cifrato sarà quindi la L. Poi si cerca

la lettera I sempre sulla prima riga e quindi la lettera O sarà la seconda lettera del testo cifrato. Si prosegue in questo modo fino ad arrivare alla fine del testo in chiaro. Quindi al testo in chiaro “CIAO” corrisponde il testo cifrato “LOEY”.

Il destinatario del messaggio riceve la parola cifrata e per risalire al testo in chiaro procede, come nel cifrario di Cesare “tradizionale”, cercando ogni singola lettera sulla seconda riga e sostituendola, nel testo in chiaro ricostruito, con la lettera che si trova sopra la lettera appena cercata.

4.1 ATTACCO “A FORZA BRUTA”

L’attacco a forza bruta, come è stato visto, non richiede particolari conoscenze teoriche, basta provare tutte le chiavi possibili. Nel caso del crittosistema di Cesare “tradizionale”, come abbiamo già visto, il numero di tentativi da compiere per risalire al testo in chiaro non preoccupa più di tanto il crittoanalista.

Si chiede quindi agli studenti di formulare ipotesi su quale possa essere il numero massimo di chiavi di un crittosistema di Cesare “reso più sicuro”. In base a quanto detto sulla costruzione della chiave, tale numero equivale al numero di sequenze di 26 lettere che si possono ottenere utilizzando tutte le lettere dell’alfabeto senza ripetere la stessa lettera.

La domanda può essere fatta precedere da un gioco sugli *anagrammi*: ad esempio si può far calcolare agli studenti il numero di tutti i possibili anagrammi di alcune parole.

Per rispondere alla richiesta di calcolare il numero degli anagrammi dell’alfabeto si può partire ponendo agli studenti il seguente quesito: quanti sono gli anagrammi della parola APE? Dopo aver lavorato sulle strategie proposte dagli alunni si può concludere che APE, AEP, PEA, PAE, EAP, EPA sono tutti gli anagrammi richiesti.

Considerando l’ordine in cui sono scritti, peraltro ininfluente sul risultato che si sta cercando, si può notare che occorre scrivere sequenze costituite da tre lettere: per la lettera iniziale ci sono tre possibilità, per la seconda ce ne sono due e, infine, una sola per la terza lettera. Quindi tutti gli anagrammi della parola APE possono essere contati moltiplicando tra loro i numeri delle possibili scelte delle singole lettere quindi

in totale si hanno $3 \cdot 2 \cdot 1 = 6$ anagrammi. Se si volessero contare i possibili anagrammi della parola CIAO allora si dovrebbe calcolare $4 \cdot 3 \cdot 2 \cdot 1 = 24$.

In generale il numero degli anagrammi di una parola formata da n lettere, tutte diverse tra loro, è di $n \cdot (n - 1) \cdot (n - 2) \cdot \dots \cdot 2 \cdot 1$. Questo numero si indica con $n!$ e si legge “ n fattoriale”.

Quindi, tornando al crittosistema “reso più sicuro”, gli studenti arriveranno ad affermare che le possibili chiavi sono $26! = 26 \cdot 25 \cdot 24 \cdot \dots \cdot 2 \cdot 1$ (togliere l’unica chiave che non permuta nessuna lettera dell’alfabeto non influisce sull’ordine di grandezza del numero). Il risultato di questo prodotto, almeno in apparenza, garantisce una certa sicurezza del crittosistema. Cercare di indovinare una chiave, tra tutte queste, “sembra” impossibile.

È interessante notare che è sufficiente un semplice “rimescolamento” delle lettere che si trovano sulla seconda riga della tabella del cifrario di Cesare, come è stato fatto nella Tabella 3, per rinvigorire la sicurezza del crittosistema.

Il numero di tentativi da compiere per forzare il crittosistema rende inapplicabile l’attacco a forza bruta realizzato, come in precedenza, con il solo ausilio di un foglio di carta con cui i ragazzi provano a decifrare il messaggio. Anzi, «anche procedendo al ritmo di una chiave al secondo, per completare il controllo occorrerebbe un tempo di gran lunga superiore all’età dell’universo».¹³

Gli studenti incontreranno nel loro percorso scolastico i possibili modi per stabilire l’ordine di grandezza di questi numeri, ma, anche senza effettuare molti calcoli, si è portati a concludere che il numero che si ottiene rende impossibile la ricerca esaustiva della chiave. Occorre ricordare che in qualsiasi lavoro di crittoanalisi il *tempo* impiegato per ottenere il testo in chiaro è un elemento fondamentale. Basti pensare che le transazioni bancarie, anche di capitali ingenti, avvengono in tempi molto rapidi, quindi riuscire a forzare un crittosistema svolgendo calcoli che durano ore o giorni non ha utilità pratica.

¹³ Cfr. SINGH 2001.

4.2 ATTACCO “STATISTICO”

L’attacco a forza bruta, come è stato visto, pur non richiedendo particolari conoscenze teoriche, spesso non permette di ottenere la chiave in tempi ragionevoli.

Un metodo di attacco più efficiente è quello che si basa sul fatto che in generale in un testo alcune lettere compaiono più frequentemente di altre. Per questo si considera la *frequenza percentuale* (che si ottiene moltiplicando per 100 la *frequenza relativa*¹⁴) che è caratteristica della lingua in cui è scritto il testo in chiaro.

I numeri della Tabella 4 «rappresentano valori medi, che corrispondono solo in modo approssimativo a quelli riscontrabili in un brano specifico»¹⁵.

Lettera	Freq	Lettera	Freq	Lettera	Freq	Lettera	Freq
A	11,74%	G	1,64%	O	9,83%	U	3,01%
B	0,92%	H	1,54%	P	3,05%	V	2,10%
C	4,50%	I	11,28%	Q	0,51%	Z	0,49%
D	3,73%	L	6,51%	R	6,37%		
E	11,79%	M	2,51%	S	4,98%		
F	0,95%	N	6,88%	T	5,62%		

Tabella 4. Frequenza percentuale delle diverse lettere in italiano (Fonte: SINGH 2001).

Per quanto detto, ovviamente, esistono singole frasi in cui la percentuale delle lettere si allontana notevolmente dai valori indicati nella Tabella 1.

Quindi la crittoanalisi, oltre all’applicazione meccanica della sostituzione di lettere in base a frequenze più o meno simili, necessita anche di una buona dose di intuito.

Il lavoro del crittoanalista è, nel caso di cifrari come quello di Cesare (sia quello originale sia quello “reso più sicuro”), agevolato dal fatto che si tratta di cifrari monoalfabetici.

4.3 ESEMPIO

Si chiede ai ragazzi di immaginare di aver intercettato il seguente messaggio segreto (di cui si sa che è scritto in lingua italiana) di cui si vuole comprendere il significato¹⁶:

SGTI P ZQPTBI NPMGUGTP PRQAELPG BEUUP TIAMLP VPDGETDP VEL ZQPTMI MEOVI PTNILP
NIBEAMI MQI SQLILE NG RESSELP P ZQPUE UGOGME AG AVGTHELP UP MQP ASLETPMP PQBPNGP

¹⁴ Si chiama *frequenza relativa* il numero che si ottiene dividendo la frequenza assoluta per il numero di unità statistiche.

¹⁵ Cfr. SINGH 2001.

¹⁶ Per abbreviare il lavoro, senza comunque perdere di generalità, in questo esempio utilizziamo l’alfabeto italiano con 21 lettere.

Per fare questo serve risalire alla *chiave*. In base al *principio di Kerckhoffs*, secondo cui è segreta solamente la chiave e non la procedura utilizzata, si sa che il messaggio è stato cifrato con il crittosistema di Cesare “reso più sicuro”.

Si può notare che, con un attacco a forza bruta, ci si può accorgere che non è stato utilizzato un cifrario di Cesare “convenzionale”, cioè ottenuto facendo scorrere di alcune posizioni l’inizio dell’alfabeto e poi procedendo secondo l’ordine alfabetico. Quindi, sapendo che nella seconda riga della tabella cifrante le lettere sono scritte in ordine casuale, si prova con l’attacco “statistico”.

Il primo passaggio da compiere è quello di contare il numero totale delle lettere che compaiono nel messaggio (nel caso dell’esempio sono 133) e il numero di occorrenze di ogni lettera per costruire una tabella delle frequenze, indicando accanto a ogni lettera la frequenza percentuale con cui tale lettera compare nel testo cifrato.

Lettera	Freq	Lettera	Freq	Lettera	Freq	Lettera	Freq
A	4,51%	G	8,27%	O	1,5%	U	4,51%
B	3%	H	0,75%	P	18,05%	V	3%
C	0%	I	7,52%	Q	6,02%	Z	2,26%
D	1,5%	L	6,77%	R	1,5%		
E	9,78%	M	6,77%	S	3,76%		
F	0%	N	3,76%	T	6,77%		

Tabella 5. Frequenze percentuali delle singole lettere nel testo intercettato.

Confrontando i valori ottenuti con quelli della Tabella 4, si può notare che la lettera del testo cifrato che ha la frequenza maggiore è la P e quindi a tale lettera dovrebbe corrispondere nel testo in chiaro una delle lettere dell’alfabeto che ha frequenza maggiore nella Tabella 1.

Non è automatico però che alla lettera P corrisponda la lettera ‘e’, potrebbe succedere, come vedremo in questo caso, che alla lettera P corrisponda la lettera ‘a’ (utilizziamo lettere maiuscole per il testo cifrato e lettere minuscole per il testo in chiaro). Nell’esempio in esame, come sarà descritto in seguito, ci si accorgerà di questo osservando le prime parole che saranno completamente in chiaro.

In generale, si procede per tentativi ed errori. Se il messaggio cifrato, come in questo caso, è formato da parole divise dagli spazi (cioè non è formato da una stringa di caratteri senza interruzione) il compito potrebbe risultare più immediato perché si possono prendere in considerazione le lettere finali di ciascuna parola cifrata sostituendole con le vocali, procedendo in ordine con le frequenze (anche se in italiano vi sono molte parole che non finiscono per vocale e il testo potrebbe contenere parole tronche). In questo caso la P è la più frequente tra le lettere finali delle parole cifrate e questo supporta l'ipotesi che alla P corrisponda una vocale, che in prima istanza dobbiamo supporre sia la 'e'. Tuttavia ciò non è necessario.

Vediamo come procedere in generale. Innanzitutto, ordiniamo le tabelle 4 e 5 in modo decrescente rispetto alla frequenza e mettiamole a confronto.

Tabella 4 ordinata

E	11,79
A	11,74
I	11,28
O	9,83
N	6,88
L	6,51
R	6,37
T	5,62
S	4,98
C	4,5
D	3,73
P	3,05
U	3,01
M	2,51
V	2,1
G	1,64
H	1,54
F	0,95
B	0,92
Q	0,51
Z	0,49

Tabella 5 ordinata

P	18,05
E	9,78
G	8,27
I	7,52
L	6,77
M	6,77
T	6,77
Q	6,02
A	4,51
U	4,51
N	3,76
S	3,76
B	3
V	3
Z	2,26
D	1,5
O	1,5
R	1,5
H	0,75
C	0
F	0

Proviamo a eseguire, l'una di seguito all'altra, le seguenti sostituzioni nel testo cifrato (nella prima si sostituisce la lettera P con la lettera 'e'), ottenendo testi parzialmente modificati (ricordiamo che nel testo da decifrare la parte modificata è scritta con lettere minuscole mentre la parte restante è scritta con lettere maiuscole).

- P con e:

SGTI e ZQeTBI NeMGUGTe eRQAELeG BEUUE TIAMLe VeDGETDe VEL ZQeTMI MEOVI eTNILe NIBEAMI MQI SQLILE NG RESSELe e ZQeUE UGOGME AG AVGTHELe Ue MQe ASLETeMe eQBeNGe

- E con a:

SGTI e ZQeTBI NeMGUGTe eRQAaLeG BaUUe TIAMLe VeDGaTDe VaL ZQeTMI MaOVI eTNILe NIBaAMI MQI SQLILa NG RaSSaLe e ZQeUa UGOGMa AG AVGTHaLe Ue MQe ASLaTeMe eQBeNGe

- G con i:

SiTI e ZQeTBI NeMiUiTe eRQAaLei BaUUe TIAMLe VeDiaTDe VaL ZQeTMI MaOVI eTNILe NIBaAMI MQI SQLILa Ni RaSSaLe e ZQeUa UiOiMa Ai AViTHaLe Ue MQe ASLaTeMe eQBeNie

- I con o:

SiTo e ZQeTBo NeMiUiTe eRQAaLei BaUUe ToAMLe VeDiaTDe VaL ZQeTMO MaOVO eTNOLe NoBaAMo MQo SQLoLa Ni RaSSaLe e ZQeUa UiOiMa Ai AViTHaLe Ue MQe ASLaTeMe eQBeNie

- L con n

SiTo e ZQeTBo NeMiUiTe eRQAanei BaUUe ToAMne VeDiaTDe Van ZQeTMO MaOVO eTNOne NoBaAMo MQo SQnona Ni RaSSane e ZQeUa UiOiMa Ai AViTHane Ue MQe ASnaTeMe eQBeNie

- M con l:

SiTo e ZQeTBo NeliUiTe eRQAanei BaUUe ToAlne VeDiaTDe Van ZQeTlo laOVO eTNOne NoBaAlo lQo SQnona Ni RaSSane e ZQeUa UiOila Ai AViTHane Ue lQe ASnaTele eQBeNie

- T con r:

Siro e ZQerBo NeliUire eRQAanei BaUUe roAlne VeDiarDe Van ZQerlo laOVO erNone NoBaAlo lQo SQnona Ni RaSSane e ZQeUa UiOila Ai AVirHane Ue lQe ASnarele eQBeNie

- Q con t:

Siro e ZterBo NeliUire eRtAanei BaUUe roAlne VeDiarDe Van Zterlo laOVO erNone NoBaAlo lto Stnona Ni RaSSane e ZteUa UiOila Ai AVirHane Ue lte ASnarele etBeNie

- A con s:

Siro e ZterBo NeliUire eRtsanei BaUUe roslne VeDiarDe Van Zterlo laOVO erNone NoBaslo lto Stnona Ni RaSSane e ZteUa UiOila si sVirHane Ue lte sSnarele etBeNie

- U con c:

Siro e ZterBo Nelicire eRtsanei Bacce roslne VeDiarDe Van Zterlo laOVO erNone NoBaslo lto Stnona Ni RaSSane e Zteca ciOila si sVirHane ce lte sSnarele etBeNie

- N con d:

Siro e ZterBo delicire eRtsanei Bacce roslne VeDiarDe Van Zterlo laOVO erdone doBaslo lto Stnona di RaSSane e Zteca ciOila si sVirHane ce lte sSnarele etBedie

A questo punto, però, si può notare che alcune parole sono formate da tutte lettere modificate, quali ad esempio “delicire”. Purtroppo quest’ultima non è una parola appartenente al vocabolario italiano, quindi occorre rivedere le sostituzioni effettuate.

Dalla Tabella 4 si evince che la differenza tra le frequenze delle lettere ‘a’ ed ‘e’ è 0,05% quindi nel prossimo tentativo di decrittazione si prova a sostituire la P del testo cifrato con la ‘a’ e la E del testo cifrato con la ‘e’.

Si procede quindi con nuove sostituzioni.

- P con a:

SGTI a ZQaTBI NaMGUGTa aRQAELaG BEUUa TIAMLa VaDGETDa VEL ZQaTMI MEOVI aTNILa NIBEAMI MQI SQLILE NG RESSELa a ZQaUE UGOGME AG AVGTHeLa Ua MQa ASLeTaMa aQBaNga

- E con e:

SGTI a ZQaTBI NaMGUGTa aRQAeLaG BeUUa TIAMLa VaDGeTDa VeL ZQaTMI MeOVI aTNILa NIBeAMI MQI SQLILe NG ReSSeLa a ZQaUe UGOGMe AG AVGTHeLa Ua MQa ASLeTaMa aQBaNga

Poi come già fatto in precedenza, si procede con le seguenti sostituzioni.

- G con i:

SiTI a ZQaTBI NaMiUiTa aRQAeLai BeUUa TIAMLa VaDieTDa VeL ZQaTMI MeOVI aTNILa NIBeAMI MQI SQLILe Ni ReSSeLa a ZQaUe UiOiMe Ai AViThela Ua MQa ASLeTaMa aQBaNia

- I con o:

SiTo a ZQaTBo NaMiUiTa aRQAeLai BeUUa ToAMLa VaDieTDa VeL ZQaTMO MeOVO aTNoLa NoBeAMo MQo SQLoLe Ni ReSSeLa a ZQaUe UiOiMe Ai AViThela Ua MQa ASLeTaMa aQBaNia

A questo punto si può notare che le lettere L, M e T hanno le stesse frequenze nel testo cifrato.

Si prova a sostituire L con ‘n’ e poi si procede con le sostituzioni successive.

- L con n:

SiTo a ZQaTBo NaMiUiTa aRQAenai BeUUa ToAMna VaDieTDa Ven ZQaTMO MeOVO aTNona NoBeAMo MQo SQnone Ni ReSSena a ZQaUe UiOiMe Ai AViThena Ua MQa ASneTaMa aQBaNia

- M con l:

SiTo a ZQaTBo NaliUiTa aRQAenai BeUUa ToAlna VaDieTDa Ven ZQaTlo leOVO aTNona NoBeAlo lQo SQnone Ni ReSSena a ZQaUe UiOile Ai AViThena Ua lQa ASneTala aQBaNia

- T con r:

Siro a ZQarBo NaliUira aRQAenai BeUUa roAlna VaDierDa Ven ZQarlo leOVO arNona NoBeAlo lQo SQnone Ni ReSSena a ZQaUe UiOile Ai AVirHena Ua lQa ASnerala aQBaNia

- Q con t:

Siro a ZtarBo NaliUira aRtAenai BeUUa roAlna VaDierDa Ven Ztarlo leOVO arNona NoBeAlo lto
Stnone Ni ReSSena a ZtaUe UiOile Ai AVirHena Ua lta ASnerala atBaNia

- A con s:

Siro a ZtarBo NaliUira aRtsenai BeUUa roslna VaDierDa Ven Ztarlo leOVO arNona NoBeslo lto
Stnone Ni ReSSena a ZtaUe UiOile si sVirHena Ua lta sSnerala atBaNia

- U con c:

Siro a ZtarBo Nalicira aRtsenai Becca roslna VaDierDa Ven Ztarlo leOVO arNona NoBeslo lto Stnone
Ni ReSSena a Ztace ciOile si sVirHena ca lta sSnerala atBaNia

Anche qui si vede che ci sono parole che, sebbene non ancora completamente decifrate, non sembrano appartenere al vocabolario italiano.

Quindi si riparte dal punto in cui è stato osservato che le lettere L, M, T hanno la stessa frequenza e si procede con le seguenti sostituzioni: L con 'r' e T con 'n'.

- T con n:

Sino a ZQanBo NaMiUina aRQAeLai BeUUa noAMLa VaDienDa VeL ZQanMo MeOVO anNoLa
NoBeAMo MQo SQLoLe Ni ReSSeLa a ZQaUe UiOiMe Ai AVinHeLa Ua MQa ASLenaMa aQBaNia

- M con l:

Sino a ZQanBo NaliUina aRQAeLai BeUUa noAlLa VaDienDa VeL ZQanlo leOVO anNoLa NoBeAlo lQo
SQLoLe Ni ReSSeLa a ZQaUe UiOile Ai AVinHeLa Ua lQa ASLenala aQBaNia

- L con r:

Sino a ZQanBo NaliUina aRQAerai BeUUa noAlra VaDienDa Ver ZQanlo leOVO anNora NoBeAlo lQo
SQrore Ni ReSSera a ZQaUe UiOile Ai AVinHera Ua lQa ASrenala aQBaNia

- Q con t:

Sino a ZtanBo NaliUina aRtAerai BeUUa noAlra VaDienDa Ver Ztanlo leOVO anNora NoBeAlo lto
Strore Ni ReSSera a ZtaUe UiOile Ai AVinHera Ua lta ASrenala atBaNia

- A con s:

Sino a ZtanBo NaliUina aRtserai BeUUa noslra VaDienDa Ver Ztanlo leOVO anNora NoBeslo lto
Strore Ni ReSSera a ZtaUe UiOile si sVinHera Ua lta sSrenala atBaNia

Qui l'intuizione può aiutarci osservando, ad esempio, la settima parola. Si ipotizza sia "nostra" e non "noslra" e quindi si sostituisce M con 't' e non più con 'l'.

Sostituendo M con 't' si ottiene la seguente frase:

Sino a ZQanBo NatiUina aRQAeLai BeUUa noAtLa VaDienDa VeL ZQanto teOVO anNoLa NoBeAto
tQo SQLoLe Ni ReSSeLa a ZQaUe UiOite Ai AVinHeLa Ua tQa ASLenata aQBaNia

Poi, si può procedere con altre sostituzioni.

- L con r

Sino a ZQanBo NatiUina aRQAerai BeUUa noAtra VaDienDa Ver ZQanto teOVO anNora NoBeAto tQo SQrore Ni ReSSera a ZQaUe UiOite Ai AVinHera Ua tQa ASrenata aQBaNia

Adesso Q non si può più sostituire con 't' poiché abbiamo già sostituito M con 't'.

Riprendiamo il testo ottenuto dopo la sostituzione di L con 'r' e sostituiamo A con 's' e otteniamo la frase seguente:

Sino a ZQanBo NatiUina aRQserai BeUUa nostra VaDienDa Ver ZQanto teOVO anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe UiOite si sVinHera Ua tQa sSrenata aQBaNia

Adesso l'intuizione diventa sempre più efficace e tale da sostituire, almeno parzialmente, l'analisi delle frequenze. Le parole VaDienDa e Ver suggeriscono di sostituire V con 'p' e D con 'z'.

- V con p:

Sino a ZQanBo NatiUina aRQserai BeUUa nostra paDienDa per ZQanto teOpo anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe UiOite si spinHera Ua tQa sSrenata aQBaNia

- D con z:

Sino a ZQanBo NatiUina aRQserai BeUUa nostra pazienza per ZQanto teOpo anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe UiOite si spinHera Ua tQa sSrenata aQBaNia

Un ulteriore suggerimento, ad esempio, è fornito dall'undicesima parola: teOpo ci fa pensare a sostituire O con 'm'.

- O con m

Sino a ZQanBo NatiUina aRQserai BeUUa nostra pazienza per ZQanto tempo anNora NoBesto tQo SQrore Ni ReSSera a ZQaUe Uimite si spinHera Ua tQa sSrenata aQBaNia

La dodicesima parola suggerisce di sostituire N con 'c':

Sino a ZQanBo catiUina aRQserai BeUUa nostra pazienza per ZQanto tempo ancora coBesto tQo SQrore ci ReSSera a ZQaUe Uimite si spinHera Ua tQa sSrenata aQBacia

La sesta parola suggerisce di sostituire B con 'd' e U con 'l':

Sino a ZQando catilina aRQserai della nostra pazienza per ZQanto tempo ancora codesto tQo SQrore ci ReSSera a ZQale limite si spinHera la tQa sSrenata aQdacia

Poi, per completare, dalla terza parola si intuisce che si possono effettuare le seguenti sostituzioni: Q con 'u' e Z con 'q' e si ottiene la frase seguente:

Sino a quando Catilina abuserai della nostra pazienza per quanto tempo ancora codesto tuo furore ci befferà a quale limite si spingerà la tua sfrenata audacia.

Adesso la frase è praticamente in chiaro. Solo per completare lo schema di cifratura appena scoperto possiamo sostituire S con 'f', R con 'b' e H con 'g'. Si ottiene quindi il testo in chiaro:

Fino a quando Catilina abuserai della nostra pazienza per quanto tempo ancora codesto tuo furore ci befferà a quale limite si spingerà la tua sfrenata audacia.¹⁷

5. CONCLUSIONE

Il crittosistema di Cesare viene di solito presentato come la prima applicazione della matematica alla crittografia. Le successive modifiche che abbiamo visto hanno permesso di ottenere schemi crittografici più sicuri e contemporaneamente hanno dimostrato che serve qualche conoscenza ulteriore in ambito matematico per essere in grado di forzare il crittosistema in tempi "ragionevoli".

Nel secondo crittosistema che abbiamo incontrato, oltre ad aver descritto una possibile strategia di decrittazione facilmente inseribile in un percorso didattico, è stato mostrato che sono sufficienti anche pochi concetti base di statistica descrittiva per agevolare il lavoro del crittoanalista, permettendogli di risalire al testo in chiaro in tempi brevi.

Questa necessità di forzare rapidamente un crittosistema è il compito principale di chi vuole conoscere un segreto senza essere in possesso della chiave: tale attività, come abbiamo visto è agevolata dalla conoscenza della matematica. Quindi, con l'inserimento opportuno della crittografia in percorsi di apprendimento si permette anche di superare quella visione della matematica secondo cui essa è qualcosa di astratto in cui è difficile poter vedere un'applicazione pratica.

BIBLIOGRAFIA

BERARDI L., BEUTELSPACHER A.
1996, *Crittologia. Come proteggere le informazioni riservate*, Milano, Franco Angeli.

¹⁷ Cfr. CICERONE 1962 (frase tradotta e considerata senza punteggiatura né lettere accentate per agevolare il lavoro di decrittazione).

BORRELLI M., FIORETTO A., SGARRO A., ZUCCHERI L.

2002, *Cryptography and Statistics: A Didactical Project*, Proceedings of the 2nd International Conference on the Teaching of Mathematics (at the undergraduate level), Hernissos, Crete. 1-6 luglio 2002, IRAKLIO, CRETE. John Wiley & Sons Inc., pp. 1-6.

CICERONE M. T.

1962, *Orationes in Catilinam*, Roma, A. Signorelli.

CAZZOLA M., GRAZIAN V.

2021, *Giochi di Crittografia elementare per la scuola primaria*, in: R. BONINO, D. MAROCCHI, M. RINAUDO, M. SERIO (a cura di), *Apprendimento laboratoriale in Matematica e Fisica in presenza e a distanza* (Torino, 11-12-13 ottobre 2021 – online), Torino, Università degli Studi di Torino, pp. 350-357.

D'AMORE B.

1999, *Elementi di Didattica della Matematica*, Bologna, Pitagora Editrice.

DI MARTINO P., ZAN R.

2020, *Problemi per crescere Matematica senza paura*, Firenze, Giunti Scuola.

KERCKHOFFS A.

1883, «La cryptographie militaire», *Journal des sciences militaires*, Vol IX, pp. 5-38.

MARCEDDU M. C.

2002, *Il gioco dell'agente segreto - II Parte*, in: «Matematica dei ragazzi: Scambi di esperienze tra coetanei – antologia 2000 – 2002», Trieste, EUT – Edizioni Università Trieste, pp. 36-38.

PERES E.

1986, *Giochi matematici*, Roma, Editori Riuniti.

SABENA C., FERRI F., MARTIGNONE F., ROBOTTI E.

2019, *Insegnare e apprendere matematica*, Milano, Mondadori.

SINGH S.

2001, *Codici e segreti*, Milano, BUR.

SUETONIUS TRANQUILLUS G.

2016, *De vita Cæsarum*, Libro 1, Capitolo LVI, Chieti, Vestigium.

ZUCCHERI L.

1992, «Crittografia e statistica nella scuola elementare», *Insegnamento della matematica e delle scienze integrate*, 15, n. 1, pp. 19-38.

Seconda parte

Note e aggiornamenti critici*

ALLEGRI R., GARBARINO O.

2018, Vernazza. *Le trasformazioni nel tempo dell'insediamento e del territorio*, Novara, Libreria Geografica, 174 pp.**

Questo lavoro si inserisce in un consolidato ambito di ricerca geografica dedicato agli *studi regionali* ed è focalizzato sul centro abitato di Vernazza localizzato nelle Cinque Terre (Levante ligure) ma si estende opportunamente all'ambiente in cui l'insediamento è incastonato.

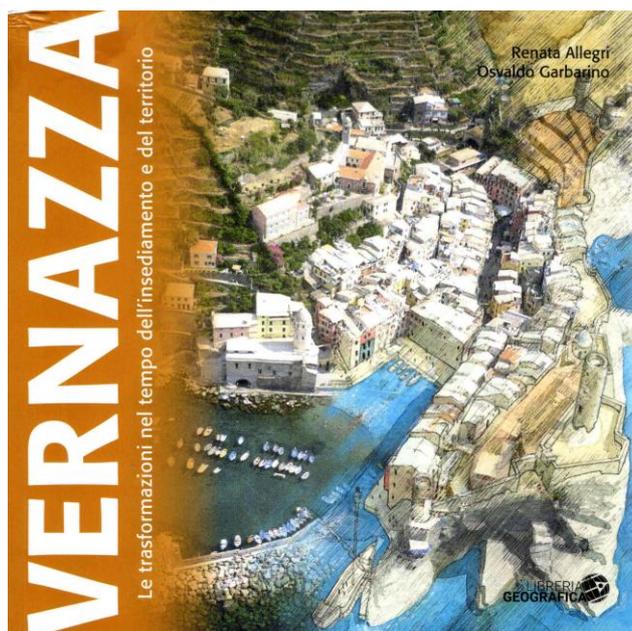


Figura 1. La copertina del volume recensito.

Si tratta di un sito angusto, caratterizzato da ripidi versanti che si affacciano sul mare, di quando in quando intervallati da modesti bacini idrografici come quello del Torrente Vernazza, e che sono divenuti oggetto, nel corso dei secoli, dell'opera

* Title: *Critical notes and upgrades.*

** ISBN: 978-8869851704.

incessante dell'uomo, volta a strappare pervicacemente spazi da utilizzare ai fini delle attività rurali di sussistenza, con la creazione di spettacolari opere di terrazzamento. Preceduta da un'introduzione, a firma di R. Allegri, che offre le prime suggestioni per introdurre alle peculiari chiavi di lettura geo-storiche utilizzate per analizzare il peculiare paesaggio in esame, la raffinata monografia si articola in quattro capitoli. I testi elaborati dagli autori sono corredati da un cospicuo apparato iconografico che comprende suggestive foto panoramiche (anche d'epoca), immagini satellitari, carte storiche, geotematiche e archeologiche, piante e analisi archeologiche di edifici e di siti urbanizzati, il tutto impreziosito da bellissimi diorami, da interessanti ricostruzioni grafiche dell'assetto del territorio e da disegni originali di pregevole fattura.



Figura 2. I terrazzamenti che caratterizzano i versanti a nord-ovest del centro abitato.
(Fonte: Google Maps, <<https://maps.google.com>>)

Nel primo capitolo, R. Allegri offre un'accurata descrizione geografica del territorio e delle sue trasformazioni geostoriche, indulgiando in particolare sui terrazzamenti e illustrandone dettagliatamente caratteristiche e forme di recente valorizzazione, grazie all'istituzione del Parco Nazionale delle Cinque Terre.

In questo interessante capitolo, l'autrice si sofferma sulle tecniche e sulle soluzioni utilizzate ai fini della costruzione dei terrazzamenti e della loro manutenzione, sugli interventi complementari di sistemazione idraulico-forestale volti alla messa in sicurezza dei versanti ma anche alla disciplina, al governo e all'utilizzo razionale delle acque, sia selvagge sia incanalate, nonché sulle tecniche di coltivazione degli oliveti e dei predominanti vigneti piantumati sui ripiani in precedenza ricavati e, infine, sulle dinamiche conseguenti all'abbandono dei terrazzamenti stessi.

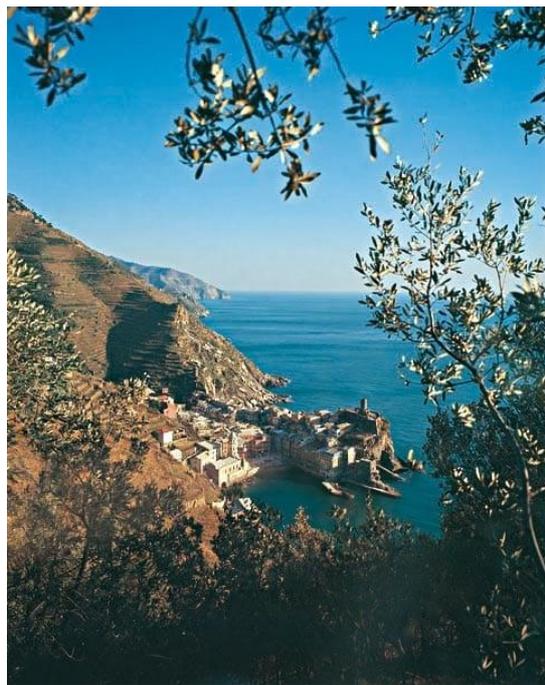


Figura 3. Il centro abitato di Vernazza inquadrato da nord-ovest.
(Fonte: <<https://borghipiubelliditalia.it/borgo/vernazza/>>)

Nel secondo capitolo, O. Garbarino esamina il borgo e i suoi monumenti da un punto di vista storico-archeologico, dedicando una finestra di approfondimento all'architettura eulitica e descrivendone minuziosamente caratteristiche e diffusione in termini cronospaziali nell'ambito euri-mediterraneo, nonché individuandone e cartografandone le

testimonianze esistenti a Vernazza che affiorano da sotto gli intonaci più recenti. L'autore offre inoltre accurati esempi di analisi archeologica di edifici paradigmatici del borgo (es. antico Hospitale, Chiesa di Santa Margherita), mettendone in luce la stratificazione delle fasi costruttive.



Figura 4. Il sito (sopra) e la posizione (sotto) di Vernazza, ossia il territorio in cui il sito è inserito. (Fonte: Google Maps, <<https://maps.google.com>>)

Nel terzo capitolo, O. Garbarino, delinea un'accurata ricostruzione della storia del borgo e del territorio vernazzese sulla base di quanto desumibile da fonti storiche di diversa tipologia, punteggiando la trattazione con puntuali ricostruzioni diacroniche ipotetiche sia del sito urbanizzato, sia di insediamenti disseminati nel territorio

vernazzese (es. il Casale di Ponzò), sia di peculiari edifici (es. la Chiesa-Santuario di Nostra Signora di Reggio), sia del territorio vernazzese nel suo insieme, anche con l'attribuzione e la localizzazione dei toponimi fondiari di antica origine.

Per quanto concerne il centro urbano, l'autore dedica un interessante focus al palinsesto archeologico della Chiesa di Santa Margherita d'Antiochia ma meritano certamente considerazione anche la ricostruzione delle successive trasformazioni che hanno interessato l'isolato Survelortu nonché l'Hospitale, così come l'analisi stratigrafica di una porzione d'argine insistente sul Torrente Vernazzola che evidenzia ben cinque fasi di rifacimento delle opere murarie.

Nel quarto capitolo, R. Allegri espande la trattazione, per includervi un'originale analisi geostorica delle relazioni extraterritoriali di Vernazza, concentrandosi sulla coltura del vino. Si tratta di aspetti che derivano dall'intrecciarsi della storia del borgo con quella di Genova, impegnata a espandere la propria sfera d'influenza e i propri scambi commerciali nel bacino del Mediterraneo. Merita menzione, in proposito, il focus dedicato a Chios e alla produzione di vino in epoca antica, alla fama dei vini greci e alla tradizione del vitigno importato proveniente dall'isola egea.



Figura 5. Il territorio del Parco Nazionale delle Cinque Terre.
(Fonte: <https://www.parconazionale5terre.it/map.php>)

L'appendice offre un'interessante tabella di sintesi, curata da O. Garbarino, che ricapitola le successive fasi di antropizzazione del territorio di Vernazza a partire dall'epoca

romana, per ciascuna delle quali, accanto alla descrizione del contesto, vengono richiamate le principali fonti (scritte, materiali, toponomastiche, orali/iconografiche).

L'opera propone, infine, uno spazio opportunamente dedicato alla cartografia, ove sono riportati estratti in scala 1:25.000 della *Carta escursionistica delle Cinque Terre, Golfo della Spezia* per i tipi delle edizioni Geo4Map, utili a quanti intendessero progettare una visita di studio al territorio in questione o un più ampio viaggio di istruzione.

Dal punto di vista didattico la monografia offre utili spunti sul piano metodologico per impostare in termini euristici il lavoro scolastico attraverso la realizzazione di laboratori interdisciplinari dedicati all'esame di territori paradigmatici, in particolare per quanto attiene lo studio degli insediamenti e l'analisi comparativa di paesaggi rurali terrazzati.

PER APPROFONDIRE

ALLEGRI R.

2014, «Osservazione analisi e lettura del territorio di Vernazza: 1. come uno studio di geografia storica può indicare un metodo», *Liguria Geografia*, XVI, n. 12, Supplemento, pp. 77-82, scaricabile dall'indirizzo: <<http://www.aiig.altervista.org/notiziario/2014/Lig-Geo,%20Supplem.%20al%20n.%2012%20del%202014.GG.pdf>>.

GARIBALDI G.

2008, *L'estremo Levante ligure e l'area apuana. Ambiente, popolazione, economia dei comuni rivieraschi tra Moneglia e Montignoso e delle valli del Vara e del Magra*, Imperia, Associazione Italiana Insegnanti di Geografia – Sezione provinciale Imperia-Sanremo (il volume può essere richiesto, si veda in proposito l'indirizzo web: <<http://www.aiig.altervista.org/pubbl.html>>).

SITI WEB

FAI. I LUOGHI DEL CUORE. IL CENSIMENTO DEI LUOGHI ITALIANI DA NON DIMENTICARE

Terrazzamenti delle Cinque Terre. Vernazza, La Spezia,

<<https://fondoambiente.it/luoghi/terrazzamenti-delle-cinque-terre-vernazza?ldc>>, sito consultato il 20.10.2023.

I BORCHI PIÙ BELLI D'ITALIA, IMPERIA.

Vernazza nell'incanto delle cinque terre,

<<https://borghipiubelliditalia.it/borgo/vernazza/>>, sito consultato il 20.10.2023.

PARCO NAZIONALE – AREA MARINA CINQUE TERRE

<<https://www.parconazionale5terre.it/>>, sito consultato il 20.10.2023.

MICHELE STOPPA
Dipartimento MIGe
Università di Trieste
mstoppa@units.it