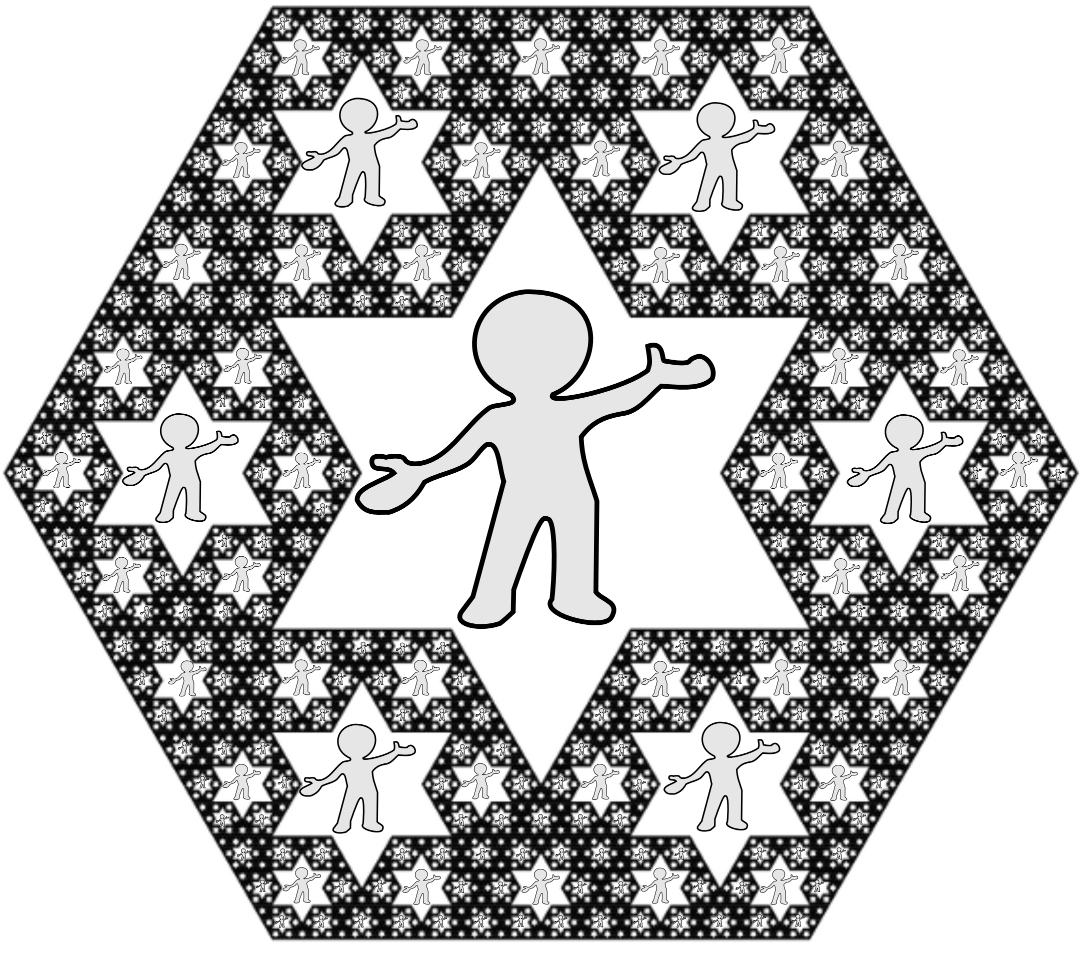
n. 2

Pisa, 16-18 febbraio 2016

Matematica
il giornalino degli OpenDays

...notizie, giochi e pillole di matematica



Dipartimento di Matematica

Su indicazione della Commissione Orientamento.
Realizzazione a cura degli studenti Counseling 2016:

Agnese Barbensi

Dario Villanis Ziani

Coordinamento: Prof. Giovanni Gaiffi

Per informazioni:
counseling.dm.unipi@gmail.com

Introduzione

In occasione di alcuni importanti eventi di orientamento presso l'Università di Pisa, il Dipartimento di Matematica pubblica questo giornalino, preparato dagli studenti del Corso di Laurea in Matematica e rivolto agli studenti delle scuole secondarie.

Come i primi numeri, che potete consultare all'indirizzo
https://www.dm.unipi.it/webnew/it/orientamento/il-giornalino-degli-open-days,
anche questo volumetto offre un piccolo 'panorama matematico', e propone, speriamo, molti spunti di riflessione.

Per cominciare potrete trovare un breve resoconto di un'indagine su cosa fanno i laureati in matematica al termine degli studi. Qualche statistica potrà essere utile per orientarsi nella scelta universitaria.

Poi potrete leggere due articoli in cui si illustrano due importanti (e molto recenti) applicazioni della matematica: il primo riguarda alcune possibili applicazioni della statistica alla lotta contro i tumori, il secondo riguarda gli algoritmi che i motori di ricerca utilizzano per stabilire la rilevanza delle pagine web.

Non poteva certo mancare una sezione di esercizi (troverete online anche le soluzioni), e, per stimolarvi a 'entrare in gioco', una sezione dedicata al gioco Chomp, di cui verranno spiegate le regole ed il significato matematico.

Per coloro che desiderano approfondire la conoscenza della matematica, suggeriremo infine alcune letture e alcuni link interessanti da consultare. Buona lettura!

Indice

Il Corso di Laurea in Matematica

Il Corso di Laurea in Matematica è essenzialmente suddiviso in due parti: un Corso di Laurea triennale, in cui vi verranno forniti i fondamenti della Matematica, ed al termine del quale conseguirete un primo titolo; ed un Corso di Laurea Magistrale, della durata di due ulteriori anni, che vi fornirà l'opportunità di ampliare le vostre conoscenze, approfondendo il settore che più vi piace. Sebbene il biennio magistrale sia una naturale prosecuzione dei primi tre anni, potrete anche decidere di terminare i vostri studi dopo la sola triennale, o eventualmente migrare verso altri corsi magistrali affini. Nel caso in cui conseguiate anche il titolo Magistrale, avrete l'opportunità di accedere al mondo della ricerca mediante un Dottorato di Ricerca.
All'interno del percorso della Laurea Triennale avrete la possibilità di scegliere tra due possibili curriculum, uno fondamentale ed uno computazionale; il primo è volto ad approfondire le discipline più astratte, come l'Algebra, la Geometria, l'Analisi o la Logica, mentre il secondo è caratterizzato da un aspetto più applicativo e modellistico. In ogni caso, poichè come già detto il corso Triennale vuole fornire una preparazione di base, molti esami sono in comune ai due curriculum ed obbligatori per tutti ed in particolare gli esami dei primi due anni sono quasi interamente prefissati; inoltre la scelta di un curriculum nel corso di laurea triennale non impedisce in alcun modo di cambiare radicalmente rotta in un'eventuale Magistrale.
Illustriamo dunque concretamente alcuni possibili percorsi di studi, sempre relativi al primo triennio, individuando prima gli esami in comune e poi le possibili scelte.

- Al primo anno, in comune ai due curriculum, affronterete gli esami annuali di Analisi e di Geometria, oltre ai corsi semestrali di Fondamenti di Programmazione con Laboratorio, Fisica 1 con Laboratorio e Aritmetica ed ad un Laboratorio di Comunicazione mediante Calcolatore.

- Anche il secondo anno prevede poca diversificazione: avrete ancora un esame annuale di Analisi ed uno di Geometria, ed i corsi semestrali di Algebra 1, Analisi Numerica con Laboratorio, Elementi di Probabilità e Statistica oltre ad un corso di Inglese Scientifico e ad un Laboratorio Didattico di Matematica Computazionale; inoltre coloro che avranno scelto il curriculum computazionale affronteranno un esame di Algoritmi e Strutture Dati, mentre chi avrà preferito il curriculum fondamentale avrà un esame a scelta, che tipicamente viene selezionato tra Elementi di Teoria degli Insiemi e Algebra 2.

- Al terzo anno si avverte maggiormente la diversificazione tra i due curriculum, dal momento che l'unico esame in comune è Sistemi Dinamici; per il resto, il curriculum fondamentale prevede Fisica 2, Fisica 3, un Laboratorio Sperimentale di Matematica Computazionale e quattro esami a scelta, che vengono spesso selezionati tra i corsi più teorici, come ad esempio Logica Matematica, Geometria e Topologia Differenziale, Elementi di Geometria Algebrica, Analisi Matematica III; invece all'interno del curriculum computazionale sono inseriti gli esami di Calcolo Scientifico, Linguaggi di Programmazione con Laboratorio, Ricerca Operativa, un Laboratorio Computazionale e tre esami a scelta, come ad esempio Statistica Matematica, Elementi di Meccanica Celeste, Finanza Matematica.

Sbocchi lavorativi

È opinione diffusa che un laureato in Matematica sia destinato esclusivamente all'insegnamento o alla carriera accademica, ma in realtà si aprono varie altre possibilità. Uno studente che abbia terminato il Corso di Laurea in Matematica è una persona in grado di affrontare in modo efficace e generale una vasta gamma di problemi, sia a livello di conoscenze che a livello di competenze, ed è principalmente la forma mentale acquisita durante gli studi in Matematica ad essere particolarmente apprezzata in ambito lavorativo.

Qui di seguito riportiamo alcuni dati occupazionali, rilevati ad un anno dalla laurea, riguardanti i laureati Triennali e Magistrali nel 2014. Come potete vedere, la situazione occupazionale che ne risulta è molto positiva.

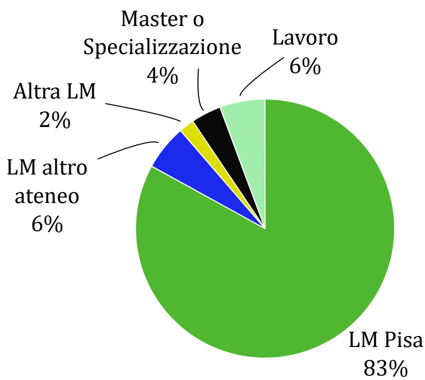


Figura 1: Laureati Triennali nel 2014

In particolare, la maggior parte dei laureati Triennali prosegue gli studi in Matematica, principalmente nel nostro ateneo, e solo una piccola parte decide di completare la formazione in altro modo oppure di entrare nel mondo del lavoro.

Per quanto riguarda i laureati Magistrali, la maggior parte sceglie di proseguire la propria formazione con un dottorato di

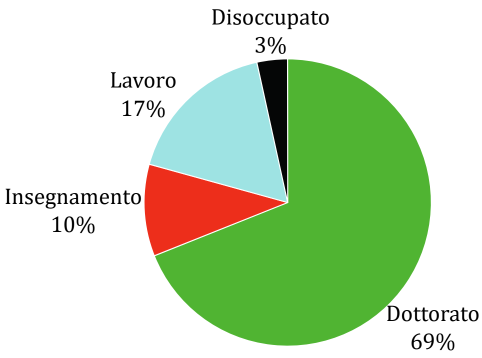


Figura 2: Laureati Magistrali nel 2014

ricerca, circa un quarto entra nel mondo del lavoro o dell'insegnamento e soltanto una piccola percentuale risulta, a dodici mesi dal conseguimento della laurea, inoccupata.

Automi cellulari e biologia dei tumori

Matematica ed oncologia

L'incontro con la malattia ed il dolore è tra le esperienze più importanti della vita e per questo molte persone elaborano il desiderio di contribuire alla cura o all'ideazione di cure per le malattie. La strada maestra senza dubbio è diventare medico ed operare direttamente nel settore sanitario ma, pensando anche al versante dello sviluppo di nuove cure, lo studio della chimica e della biologia, o di discipline più recenti come la bioingegneria, può portare su quella strada. Persino la fisica ha sviluppato una specializzazione che si avvicina alla medicina: la figura professionale del fisico medico.

La matematica invece, per quanto sia a fondamento della fisica, della chimica e dell'ingegneria, ha un rapporto meno diretto con le scienze della vita ed in particolare con la medicina. La connessione più ovvia è attraverso la statistica: tutti sentiamo esporre in forma di percentuali alcuni fatti relative alla salute dell'uomo, percentuali ottenute tramite indagini statistiche - è recente ad esempio l'affermazione che il consumo di carni trasformate - salumi ecc. - aumenta del 18% il rischio relativo di contrarre un tumore.

La matematica però ha tra le sue abilità quella di sviluppare modelli. La fisica è basata su questo. I modelli possono adattarsi a situazioni anche distanti dalla fisica, come ad esempio la finanza e l'economia. Allora perché non cercare di sviluppare modelli anche su temi specifici di medicina?

Sostanzialmente la resistenza a questa direzione di ricerca sta nella complessità dei fenomeni coinvolti. La fisica ha saputo scomporre i problemi in blocchi elementari, isolati dalla complessità dei fenomeni usuali; la matematica ha poi permesso di descrivere accuratamente il singolo tassello e tornare successivamente a costruzioni più complesse. In medicina, o nella vicina biologia, questo ancora non è accaduto.

Pensiamo ad esempio ad un un tessuto umano: esso è composto di una miriade di cellule ed altre strutture. Anche un gas, studiato dalla fisica, è composto da una miriade di atomi o molecole. Ma gli atomi sono oggetti relativamente semplici, tutti uguali se della stessa materia, hanno gli uni con gli altri dei rapporti (interazioni) piuttosto elementari, mentre ogni cellula è un organismo di complessità eccezionale, compie una miriade di operazioni diverse, interagisce in modo estremamente complesso con le altre cellule. Ogni semplice modello matematico del comportamento di un tessuto non potrà tener conto di questa complessità. Anche un atomo di un gas internamente ha una sua complessità ma minore e di solito essa non interviene sul comportamento del gas nel suo complesso così fortemente quanto invece succede per l'esempio del tessuto e delle cellule.

Sembra quindi una strada senza speranza o per lo meno lontana ma credo che non ci si debba perdere d'animo e si debbano invece compiere con fiducia dei primi passi, ben sapendo che la strada è molto più lunga rispetto ad altre discipline. D'altra parte, se si potesse contribuire anche solo in minima parte alla cura, ad esempio, dei tumori, o delle malattie di tipo circolatorio, il contributo al bene dell'umanità sarebbe enorme.

Descrivere matematicamente la crescita di un tumore

Che forma può avere un modello matematico in oncologia? Focalizziamo su pochi aspetti, per arrivare rapidamente a qualche risultato. Poniamo la nostra attenzione sulla *crescita di una massa tumorale.*

Premessa su alcuni elementi di biologia cellulare dei tumori

Prima di sviluppare la matematica, è bene ricordare alcuni fatti essenziali di biologia dei tumori. All'origine c'è una prima cellula normale di un tessuto, di solito epiteliale (pelle, superfici interne al corpo), che ha subito un certo numero di modificazioni genetiche, piuttosto precise. Le cellule epiteliali, essendo soggette ad un continuo ricambio ed essendo maggiormente esposte ad agenti esterni come radiazioni, fumo, inquinamento, eccessivo utilizzo di certi cibi, hanno maggior probabilità di sviluppare generiche modificazioni genetiche; si trasformano in cellule tumorali se le modificazioni sono di tipo specifico. La caratteristica principale della nuova cellula modificata è quella di poter attivare la divisione cellulare liberamente, senza bisogno di stimoli precisi come avviene per le cellule sane, e di trasmettere questa caratteristica alle due cellule che si generano dalla sua divisione cellulare. Nulla può allora arrestare questo meccanismo proliferativo: da una cellula se ne formano due in breve tempo (la durata del ciclo cellulare varia molto da cellula a cellula; per molte però dura circa 16 ore; per semplificare, nel seguito supponiamo che serva un giorno per una duplicazione), poi da ciascuna delle due se ne formano altre due e così via. Se il processo fosse così rigido e preciso, supponendo per semplicità che serva un giorno per una duplicazione cellulare, dopo 10 giorni da una cellula ne avremmo $2^{10} = 1024$, e dopo un solo mese $2^{30} \sim 10^9$. Questo numero, 10^9, è l'ordine di grandezza giusto del numero di cellule in un tumore maturo, di quelli percepibili con analisi cliniche. Però in genere, per fortuna, serve più di un mese per arrivarci, perché le cellule tumorali rallentano il loro ritmo di duplicazione quando sono eccessivamente circondate da altre cellule tumorali, a causa della carenza di ossigeno e nutrimenti che si crea per l'eccessivo affollamento. Continuano a duplicare secondo la norma solo le cellule della parte più esterna del tumore, mentre quasi si fermano quelle più interne.

Nel frattempo, tra le modificazioni genetiche tipiche delle cellule tumorali, c'è quella che permette il movimento. Tutte le cellule hanno una certa capacità di movimento che però di solito è impedita dall'adesione tra le membrane cellulari: due cellule in contatto attivano, nella parte di membrana con cui sono in contatto, delle proteine che, come fossero dei gancetti, le tengono unite. La loro capacità di movimento si restringe quindi a poche mosse di assestamento, ad esempio necessarie in fase di proliferazione, per lasciar spazio alle nuove nate. Ma le cellule tumorali, magari non sin dall'inizio ma dopo un po' di maturazione della popolazione cellulare, sviluppano attraverso ulteriori trasformazioni genetiche la capacità di sciogliere quel legame tra membrane ed essere quindi libere di muoversi nell'ambiente circostante. Tale ambiente, che chiameremo genericamente *matrice extracellulare*, è composto da altre cellule, varie sostanze chimiche, varie proteine, strutture filamentose e colloidali che conferiscono una certa stabilità, vasi sanguigni ed altro. Lentamente ed in modo erratico, una cellula tumorale può spostarsi nella matrice extracellulare, di solito con l'obiettivo di raggiungere zone più ricche di ossigeno e nutrimento ed in particolare avvicinandosi così ai vasi sanguigni, da cui viene diffuso l'ossigeno. Il contatto, purtroppo, tra cellule tumorali e vasi sanguigni è la maggior causa di pericolosità dei tumori perché, se il loro effetto sul corpo vivente fosse solo di aumentare di dimensione nella regione in cui è nata la prima cellula tumorale, intanto provocherebbero un danno relativo, solo in quella regione, poi sarebbero dopo un po' facilmente identificabili e contrastabili chirurgicamente; il corpo stesso, per le ragioni di difficoltà di approvvigionamento di ossigeno alla zona interna del tumore, porrebbe un freno alla crescita. Invece, grazie al movimento delle cellule tumorali che le porta a contatto coi vasi sanguigni

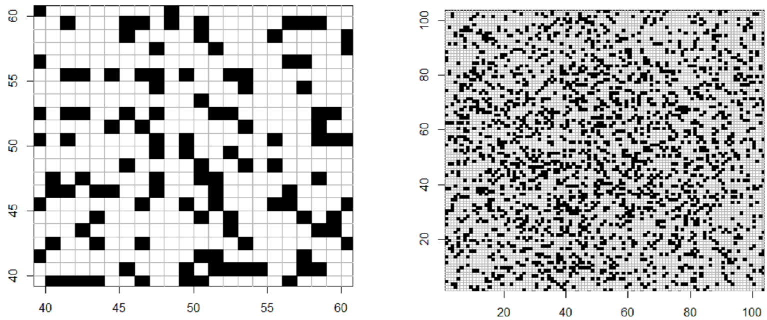
- e grazie ad un altro fenomeno simmetrico, l'angiogenesi, di cui però ora non parleremo per non allargare troppo il discorso - le cellule possono entrare nel circolo ematico e raggiungere altre parti del corpo vivente, sviluppando colonie in molti punti e impedendo così di contrastarne efficacemente la crescita. Non che questi viaggi siano una passeggiata: si pensi che mediamente una sola cellula tumorale su 10.000 sopravvive al viaggio nel sangue e riesce a impiantarsi in un altro punto del corpo; ma i numeri di cellule in gioco sono elevatissimi, come già detto sopra.

Gli automi cellulari

Descriviamo ora un modello matematico molto versatile che si può adattare a varie situazioni della realtà. La coincidenza di nome, cellulare, non deve far pensare che sia stato ideato specificamente per la descrizione delle cellule tumorali.

Si pensi ad una scacchiera, magari immensa, in teoria anche infinita, magari tridimensionale - come è in realtà qualsiasi tessuto di un corpo; nel seguito, per varie ragioni, la supporremo bidimensionale, proprio come una scacchiera usuale. Ogni casella ha un colore non predeterminato ma variabile, secondo le regole del gioco che verranno specificate. Pensiamo al caso più semplice, in cui le caselle possono essere solo bianche e nere, ma nulla impedisce di sviluppare giochi più complicati con varie colorazioni.

Supponiamo che nero significhi occupato, mentre bianco sia vuoto. La scacchiera con certe caselle bianche ed altre nere ci mostra quali zone del piano sono occupate e quali libere. Si deve pensare anche ad una visione dall'alto, da lontano, con lo zoom, su una grandissima scacchiera: emergeranno delle masse nere su fondo bianco, composte di una miriade di caselle occupate. Ecco due figure che illustrano un dettaglio della scacchiera vista da vicino ed una visione più da lontano:



Nel nostro esempio dei tumori, si può pensare che per occupato intendiamo occupato da una cellula tumorale invece che da cellule sane o altre strutture della matrice extracellulare. La visione da vicino ci mostra il dettaglio delle singole cellule, come sono disposte le une rispetto alle altre; la visione da lontano mostra la massa tumorale nel suo complesso, rispetto allo spazio vuoto (non occupato dal tumore).

Ciò di cui abbiamo parlato per ora sono le cosiddette *configurazioni* del sistema. Una configurazione è la specificazione di quali caselle sono nere e quali bianche. Se ci sono in tutto N caselle nella scacchiera, le configurazioni possibili sono 2^N: numerando le caselle e percorrendole una dopo l'altra, la prima casella potrebbe trovarsi in due condizioni - bianco o nero - e data la sua condizione, la seconda casella potrebbe trovarsi anch'essa in due e così via (2 per 2 per 2 ..., N volte). L'insieme delle configurazioni possibili è molto ampio; si pensi che una piccola scacchiera quadrata avente ogni lato di lunghezza 100 (come nella maggior parte delle figure mostrate qui), ha un totale di $N = 100 \times 100$ caselle; ma allora il numero di configurazioni possibili sarà $2^{100\times 100}$, numero nemmeno immaginabile nei termini dei numeri usualmente accessibili. In questo senso, pur essendo l'automa cellulare un *modello discreto*, esso può approssimare abbastanza bene la complessità di un *modello continuo*.

Ora dobbiamo parlare della dinamica di un automa cellulare, dobbiamo specificare le mosse con cui esso modifica la propria configurazione. In teoria, si potrebbero specificare mosse globali, che toccano contemporaneamente tutte le caselle della scacchiera, ma tali mosse avrebbero una complessità troppo elevata, metterebbero in gioco i numeri troppo grossi visti sopra. Conviene allora specificare mosse locali. Con questo si intende che viene identificata una casella, con qualche criterio - ad esempio del tutto a caso - e viene prescritto cosa avviene ad essa ed eventualmente a quelle vicine. Una mossa globale, sempre che abbia senso o sia utile pensarla, è in un certo senso il risultato di una miriade di mosse locali successive.

Così facendo, agendo cioè per mosse locali, da un lato possiamo più facilmente prescrivere delle regole che ricalchino la realtà che vogliamo descrivere, dall'altro abbiamo uno schema più facile da implementare con un codice numerico al calcolatore. Un'ultima osservazione generale: nei modelli usuali di automi cellulari il caso viene immesso nell'algoritmo a vari livelli. Abbiamo già accennato al fatto che la scelta della casella in cui effettuare una mossa locale è spesso effettuata a caso, dando ad ogni casella la stessa probabiità di essere selezionata; oltre a questo, viene spesso scelto a caso il tempo che deve intercorrere tra una mossa e l'altra; scelta la casella, nell'operare la mossa viene spesso inserito il caso per scegliere tra varie mosse possibili (in questo caso pesando opportunamente le varie mosse in modo da far accadere più spesso quelle più frequenti).

Un esempio di automa cellulare per la crescita di un tumore

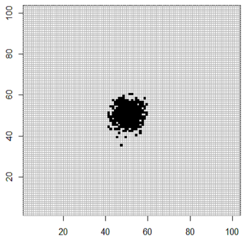
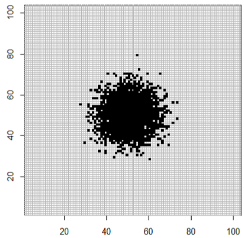
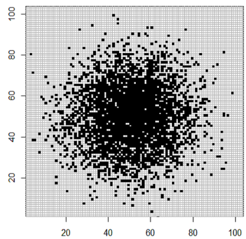
Immaginiamo quindi di prescrivere come condizione iniziale della nostra dinamica cellulare la configurazione formata da una casella occupata e tutte le altre libere - è il momento iniziale in cui si è formata la prima cellula tumorale. Prescriviamo poi che il sistema scelga a caso tra due possibilità, movimento o proliferazione, magari con probabilità non uguali, $1-p$ e p. Se l'algoritmo sceglie il movimento, poi deve scegliere una a caso delle 4 caselle circostanti per spostare lì la cellula tumorale. Se sceglie la proliferazione, deve analogamente scegliere a caso una delle 4 caselle circostanti per generare lì una delle due nuove cellule; l'altra nuova generata, occuperà la posizione del genitore.

Dopo la prima duplicazione, abbiamo due caselle occupate, due cellule tumorali. Il sistema sceglie a caso tra le due: questa è la fase iniziale casuale della scelta della casella in cui operare la mossa locale. Non si sceglie a caso tra tutte le caselle ma solo tra quelle occupate, perché non serve fare mosse nelle caselle vuote, in questo modello molto semplificato - si possono invece immaginare situazioni più complicate in cui invece degli spazi vuoti ci sono spazi occupati da varie componenti della matrice extracellulare, che quindi sono coinvolte nella dinamica.

Scelta una delle due, l'algoritmo opera come in precedenza: decide a caso se operare un movimento o una proliferazione e poi sceglie la casella limitrofa in cui muoversi o proliferare. L'unico accorgimento è che la scelta della casella limitrofa è limitata a quelle libere, perché non tutte lo sono. A questo punto l'algoritmo procede nello stesso modo, con le stesse regole, anche quando la scacchiera contiene molte cellule tumorali: ad ogni passo viene scelta una cellula tumorale, viene deciso se muoverla o proliferare, viene scelta una casella limitrofa in ciò avviene.

Può accadere che non ci siano caselle limitrofe libere: in questo caso la mossa viene cancellata e si ricomincia scegliendo a caso una casella. In questo modo, automaticamente, sia il movimento sia soprattutto la proliferazione vengono inibite nelle zone troppo piene, rispettando una caratteristica reale.

Le seguenti tre figure mostrano uno stadio iniziale, uno più avanzato in cui il tumore ha ancora un carattere prevalentemente proliferativo (qui il parametro p è quasi uguale ad 1) ed uno ancora più avanzato in cui viene maggiormente espresso il *fenotipo* diffusivo (qui il parametro p è quasi uguale ad 0):

L'algoritmo ora descritto a parole è suscettibile di una ben precisa formulazione matematica tramite il concetto di *catena di Markov* e la *teoria dei processi stocastici*. Il calcolo delle probabilità permette di formalizzare le varie mosse casuali descritte sopra e descrivere l'intera procedura tramite uno schema iterativo casuale. La teoria matematica consente ad esempio di esaminare rigorosamente cosa accade quando si prende una scacchiera infinita e la si guarda con lo zoom: si studia così il *limite al continuo*, in cui al posto del reticolo della scacchiera discreta c'è il piano ed al posto di una configurazione discreta di caselle nere c'è una *densità* continua di cellule tumorali. Mentre la crescita nel tempo del tumore a livello discreto è basata su uno schema iterativo, nel limite continuo fa uso delle cosiddette equazioni differenziali alle derivate parziali, in modo simile a come si descrivono in fisica vari fenomeni che cambiano nel tempo ed hanno variazioni spaziali. Tramite opportuni codici di calcolo numerico, queste dinamiche vengono poi simulate al calcolatore. Facendo varie simulazioni, anche al variare dei diversi parametri simili a p, si possono esaminare diversi comportamenti e fare stime ad esempio sul tempo necessario al tumore per raggiungere i vasi sanguigni.

L'utilizzo medico di questo tipo di modelli è ancora allo stato embrionale, come dicevamo all'inizio. La strada è lunga ed incerta. Ma la prospettiva di poter contribuire alla ricerca ed applicazione medica in un futuro non troppo lontano è sufficiente a motivare ogni sforzo possibile per rendere sempre più realistici questi modelli.

Franco Flandoli, Professore Ordinario
Dipartimento di Matematica, Università di Pisa

La matematica dell'importanza

Il problema del ranking

In questo articolo vogliamo descrivere un problema che ha un'importanza e un'utilità pratica sempre maggiori negli ultimi anni caratterizzati dalla crescita di internet e dalla disponibilità sempre maggiore di dati da analizzare.

Supponete di fare una ricerca su internet per trovare siti che parlano dell'ultimo film della Disney. Ci sono diversi risultati pertinenti: per esempio, la pagina ufficiale del film, le recensioni sui giornali, e le pagine dei fan che pubblicano opinioni, immagini e commenti. Quando digitiamo il titolo del film, i motori di ricerca moderni riescono con un'efficacia sorprendente a distinguere quali pagine sono più importanti per noi (per esempio, il sito ufficiale) e quali sono secondarie (per esempio, le pagine dei fan). Vi siete mai chiesti come fanno?

Per gli argomenti di maggiore rilevanza, come un film famoso, è possibile chiedere a una persona di selezionare manualmente le pagine; però, farlo per tutte le possibili ricerche è inimmaginabile. Per questo è necessario un metodo automatico, un algoritmo.

Proviamo a ragionare come un matematico, e cerchiamo di formalizzare il problema: **dato un insieme di pagine web che parlano di un certo argomento, trovare quali sono più importanti.** Dobbiamo innanzitutto dire cos'è una "pagina web". Sicuramente, essa contiene del testo e delle immagini. C'è un altro elemento però: i collegamenti (link) tra una pagina e l'altra. Se indichiamo con delle frecce i collegamenti da una pagina all'altra, otteniamo una struttura come quella della Figura 3, che i matematici chiamano *grafo*.

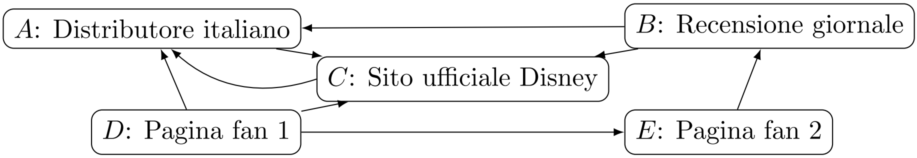


Figura 3: Collegamenti tra una pagina e l'altra, esempio.

Diverse soluzioni

Il secondo problema, più difficile, è quello di decidere cosa vuol dire "la più importante". Qui il matematico si ferma, e riconosce che a questa domanda non c'è una risposta univoca, automatica. **Definire l'importanza è una scelta arbitraria.** Una volta fatto, possiamo cercare dei modi di calcolarla, ma prima dobbiamo accordarci su una definizione.

Le prime possibilità che ci possono venire in mente sono le seguenti:

1. Le pagine 'migliori' sono quelle che contengono il nome del film nel titolo, oppure scritto più in grande.

2. Le pagine 'migliori' sono quelle che contengono più spesso il nome del film.

Riuscite ad individuare il grosso problema con criteri di questo tipo? Sono *facilmente manipolabili*: se un fan modifica il testo della sua pagina, riesce facilmente a conquistare il primo posto, scalzando i siti ufficiali. I motori di ricerca della prima generazione, negli anni '90, utilizzavano criteri di questo tipo, e abusi come questo erano all'ordine del giorno. Spesso pagine di 'spam' contenevano solo il nome di un film, ripetuto moltissime volte. Un criterio un po' migliore è il seguente:

3. Le pagine 'migliori' sono quelle verso le quali ci sono molti collegamenti.

Abbiamo introdotto un elemento nuovo: **sfruttare la struttura dei link** per ottenere un rating più affidabile. Manipolare questo criterio è un po' più scomodo, ma ancora fattibile: un fan deve semplicemente creare molte pagine fittizie, che nessuno mai guarderà, ma che contengono un link al suo sito.

Voti più e meno importanti

Ci serve invece un criterio che tenga conto che non tutti i collegamenti hanno lo stesso valore: un link proveniente dal sito della Disney 'vale' molto di più che non uno proveniente dal blog di un fan. In uno slogan: *una pagina è importante se viene linkata da pagine importanti*. Questa definizione è ricorsiva: apparentemente, per calcolare l'importanza di una pagina dobbiamo sapere l'importanza di tutte le pagine da cui partono collegamenti a lei, e così via. In realtà, vedremo che per calcolare questo punteggio basta risolvere un sistema di equazioni lineari. Vediamo di formalizzare la cosa: chiamiamo $x_A, x_B, \dots, x_E$ l'"importanza' delle pagine in Figura 3. Il sito B contiene collegamenti ad A e C; possiamo immaginare che esso 'voti' per A e C, o che 'distribuisca' la sua importanza equamente tra di essi:

$$x_C = \tfrac{1}{2}x_B + \dots \text{altri termini} \dots, \quad x_A = \tfrac{1}{2}x_B + \dots \text{altri termini} \dots$$

Analogamente, il D dividerà equamente la sua 'importanza' tra i tre siti verso cui ha link, e A la darà tutta al sito C. Complessivamente, le equazioni sono

$$\begin{cases} x_A = \frac{1}{2}x_B + x_C + \frac{1}{3}x_D, \\ x_B = x_E, \\ x_C = x_A + \frac{1}{2}x_B + \frac{1}{3}x_D, \\ x_D = 0, \\ x_E = \frac{1}{3}x_D. \end{cases} \tag{1}$$

Da sole, esse non sono ancora sufficienti a determinare univocamente l'importanza: dati $x_A, x_B, \ldots, x_E$ che soddisfano il sistema, possiamo moltiplicare tutto per due (o per tre, per quattro...) e ottenere un'altra soluzione altrettanto valida. Aggiungiamo una condizione per fissare il totale:

$$x_A + x_B + x_C + x_D + x_E = 1. \tag{2}$$

Il sistema formato da (1) e (2) ha una sola soluzione: $x_A = x_C = \frac{1}{2}, x_B = x_D = x_E = 0$. Cioè, *le uniche pagine importanti sono A e C.*

Possibili modifiche

Vi soddisfa questo risultato? Se la risposta è no, dobbiamo **cambiare la definizione di importanza**. Per esempio, per evitare zeri nella soluzione, possiamo assumere che solo una certa percentuale dell'importanza, diciamo l'85%, venga distribuita in questo modo; il restante 15% viene suddiviso in parti uguali, in modo che nessuno resti a zero. Le equazioni così modificate diventano

$$\begin{cases} x_A = (1-\alpha)\frac{1}{5} + \alpha\left(\frac{1}{2}x_B + x_C + \frac{1}{3}x_D\right), \\ x_B = (1-\alpha)\frac{1}{5} + \alpha x_E, \\ x_C = (1-\alpha)\frac{1}{5} + \alpha\left(x_A + \frac{1}{2}x_B + \frac{1}{3}x_D\right), , \\ x_D = (1-\alpha)\frac{1}{5}, \\ x_E = (1-\alpha)\frac{1}{5} + \alpha\frac{1}{3}x_D. \end{cases} \tag{3}$$

dove abbiamo lasciato scritto esplicitamente il parametro $\alpha = 0.85$. Il sistema formato da (2) e (3) ha soluzione $x_A = 0.43$, $x_B = 0.063$, $x_C = 0.43$, $x_D = 0.030$, $x_E = 0.038$. Ora la soluzione sembra molto più ragionevole! Cattura il fatto che il fan 2 è più importante del fan 1, e che la recensione è più importante di entrambi. Ora siamo pronti per usare questo metodo su problemi più grandi che non uno con cinque sole pagine. Un momento, però...

Metodi di soluzione

Come avete risolto i sistemi qui sopra (se avete provato a farlo da soli)? Molto probabilmente avete eliminato una variabile dopo l'altra, oppure avete sommato e sottratto tra loro le equazioni. Questi metodi richiedono un numero di operazioni che cresce come *il cubo del numero di pagine.* Se avessimo 10 pagine invece di 5, calcolarne l'importanza non richiederebbe il doppio del tempo, ma *otto volte tanto.* I numeri diventano difficili da gestire molto presto, anche per un calcolatore. Un computer moderno richiede circa un millisecondo per risolvere un sistema di 100 equazioni lineari in 100 incognite. Anche a questa velocità, però, calcolare l'importanza di *tutte* le pagine che parlano di un certo argomento è difficile: se abbiamo 100.000 pagine, servono 11 giorni. E se invece ne abbiamo 1.000.000?

Un metodo diverso, approssimato ma più efficiente, viene dal dare a queste equazioni un'interpretazione diversa. Supponiamo di guardare un utente annoiato che naviga tra le pagine seguendo ogni volta un link scelto a caso, tutti con la stessa probabilità. Dopo che ha seguito questa procedura per molto tempo, qual è la probabilità di trovarlo su una pagina piuttosto che su un'altra? Per esempio, chiamiamo x_C la probabilità di trovarlo su C. Perché questo succeda, ci sono tre possibilità: o la pagina precedente era la A, che succede con probabilità x_A; o era nella pagina B (e in questo caso però visto che B ha due link uscenti c'è solo $\frac{1}{2}$ di possibilità che finisca in C), oppure era nella pagina D (e in questo caso ha $\frac{1}{3}$ di possibilità di finire in C). Quindi, abbiamo l'equazione $x_C = x_A + \frac{1}{2}x_B + \frac{1}{3}x_D$. Ma questa equazione l'abbiamo già incontrata in (1), e così quelle delle altre pagine: **le quantità $x_A, x_B, \ldots, x_E$ calcolate più**

sopra rappresentano le probabilità di trovare un 'utente casuale' in ognuna delle pagine. L'equazione (2) ci dice per l'appunto che queste probabilità devono sommare a uno. In altre parole, più una pagina è importante, più è facile finirci seguendo collegamenti a caso, il che è ragionevole.

Questo suggerisce un altro metodo di soluzione: simuliamo il comportamento di questo navigatore casuale. Partiamo da una delle pagine, per esempio D. Poi, con un generatore di numeri casuali, scegliamo un link tra quelli presenti sulla pagina, per esempio quello verso E, e seguiamolo. Poi scegliamo un link a caso tra quelli presenti su E (in questo caso uno solo), e continuiamo così. Teniamo traccia della percentuale del tempo che abbiamo passato su ognuna delle pagine, e al crescere di k, questa quantità tenderà con grande probabilità alla soluzione del sistema (1)+(2). Un metodo leggermente più efficiente è questo: partiamo al passo 0 con probabilità uguali di essere su ogni pagina: $[p_A(0), p_B(0), p_C(0), p_D(0), p_E(0)] = [\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}]$. Al passo 1, non è complicato calcolare che seguendo i collegamenti le probabilità di essere su ogni pagina sono $[p_A(1), p_B(1), p_C(1), p_D(1), p_E(1)] = [\frac{11}{30}, \frac{1}{5}, \frac{11}{30}, 0, \frac{1}{15}]$. Notate che non è possibile essere nella pagina D, perché non ci sono collegamenti che puntano ad essa. Al passo 2, le probabilità sono $[\frac{7}{15}, \frac{1}{15}, \frac{7}{15}, 0, 0]$, e al passo successivo arriviamo alla soluzione vera del sistema (1)+(2).

Teletrasporto

Riusciamo ad attribuire un significato simile alle equazioni (3)? Sì, e anche questa volta la soluzione è molto elegante: esse corrispondono alle probabilità di incontrare in una certa pagina un navigatore che con probabilità α segue un collegamento, oppure altrimenti (con probabilità $1 - \alpha$) si trasferisce su una pagina scelta a caso uniformemente tra tutte. Provate a scrivere le equazioni corrispondenti a questo comportamento e a verificarlo.

Costo

Quante operazioni sono necessarie per calcolare le probabilità al passo $t+1$, date quelle al passo t? Le equazioni sono identiche a quelle di (1) (o (3)), solo che le quantità a sinistra dell'uguale si riferiscono al passo $t+1$, quelle a destra al passo t: per esempio,

$$p_C(t+1) = p_A(t) + \tfrac{1}{2}p_B(t) + \tfrac{1}{3}p_D(t).$$

C'è un addendo per ogni link *entrante* nella pagina C. Se ci sono k addendi, dobbiamo fare k moltiplicazioni e $k-1$ addizioni. Quindi il numero di operazioni in un 'passo' di questo algoritmo è $(2\ell-1)n$, dove n è il numero di pagine e ℓ è il numero medio di link entranti in ogni pagina. Per esempio, nel grafo in Figura 3, in A e C arrivano tre link, in B ed E uno, in D nessuno, quindi $\ell = 8/5 = 1.6$ collegamenti.

Potrebbe preoccuparvi questa quantità ignota ℓ: ci sono pagine che ricevono molti collegamenti, per esempio non è irragionevole che la pagina di un film famoso venga raggiunta da centinaia di migliaia di link. Però questa elegante osservazione, che un matematico chiamerebbe 'teorema', vi dimostra che la media non può essere troppo alta:

Teorema 1. *Il numero medio di link entranti è uguale al numero medio di link uscenti da una pagina.*

Dimostrazione. Invece di pensare in termini di medie, pensiamo in termini di totali: ogni freccia ha una 'coda' e una 'punta', quindi il numero totale di frecce che escono da una pagina è uguale al numero totale di frecce che entrano in una pagina. □

Nel nostro esempio, da A, C, E esce un link solo, da B due, e da D tre: in media, 1.6 collegamenti, esattamente come previsto dal teorema. Anche per pagine molto lunghe, il numero di collegamenti presenti su una pagina non supererà qualche decina, quindi ℓ non può essere troppo grande.

Un ultimo dettaglio: quanti passi di questo algoritmo sono necessari per ottenere una buona approssimazione delle 'importanze' vere? È difficile rispondere, anche per un matematico, ma per matrici grandi algoritmi di questo tipo di solito sono più efficienti di quelli 'diretti' che richiedono n^3 operazioni. Nel caso del sistema (2)+(3), per esempio, già al passo 3 abbiamo quattro cifre significative esatte.

passo	A	B	C	D	E
1	0.3417	0.2000	0.3417	0.0300	0.0867
2	0.4139	0.1037	0.4139	0.0300	0.0385
3	0.4344	0.0627	0.4344	0.0300	0.0385
4	0.4344	0.0627	0.4344	0.0300	0.0385
5	0.4344	0.0627	0.4344	0.0300	0.0385

Hubs and authorities

Una variante di questo metodo è un modello chiamato *hubs and authorities*: esso riconosce che ci sono anche pagine che pur non essendo 'buone' contengono link a pagine di buona qualità. Per esempio, una pagina di recensioni conterrà link a buoni film, mentre un buon film magari conterrà solo link a film dello stesso produttore, non necessariamente buoni. Ad ogni pagina quindi associamo due punteggi, che chiamiamo y e z (quindi $y_A, z_A, y_B, z_B, \dots$). I punteggi y indicano quanto una pagina è buona come 'recensione', e i punteggi z indicano quanto sono buoni i suoi contenuti. Una pagina ha un punteggio y più alto quando contiene link a pagine con z alto, e ha un punteggio z alto quando viene puntata da pagine con y alto. Sapreste scrivere delle equazioni per calcolare questi punteggi?

L'importanza dell'importanza

Algoritmi come questi sono diventati molto popolari negli ultimi anni; il primo indice a farne uso nell'ambito dei motori di ricerca è stato il *pagerank* di Google, che è stato uno dei fattori che ne hanno determinato il successo. Analisi automatiche di questo tipo compaiono in sempre più siti: i siti di e-commerce vogliono trovare i prodotti più adatti da consigliare ai clienti, i produttori di pubblicità vogliono trovare gli *ad* più rilevanti da mostrarci. . .

Il ruolo del matematico

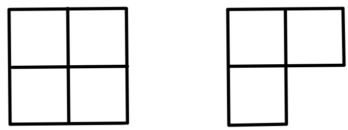
Trovare il modo migliore di calcolare questi indici è un problema che interessa matematici e informatici a diversi livelli: bisogna trovare buone definizioni di concetti come 'importanza', saperli calcolare in modo efficiente, e dimostrarne le proprietà. Per esempio, un problema che va individuato e risolto è capire come vanno modificati gli algoritmi precedenti se da una pagina non esce nessun link. Oppure: provate a vedere cosa succede se rimpiazzate il grafo di Figura 3 con uno in cui ci sono due insiemi di pagine separati, senza collegamenti tra l'uno e l'altro. Dovreste notare che il sistema (1)+(2) ha infinite soluzioni, mentre il sistema (2)+(3) continua a funzionare. Un matematico è in grado di dimostrare questi risultati e spiegare il comportamento del metodo. Infine, non è poi facile dimostrare che l'algoritmo che abbiamo descritto qui sopra produce sempre una soluzione con tutti gli x positivi o nulli, che è fondamentale per poter assegnare loro il significato di 'importanza'.

Insomma, anche nella scienza del web e dei *big data* (algoritmi per lavorare con grandi quantità di dati) il ruolo del matematico è fondamentale.

Federico Poloni, Ricercatore
Dipartimento di Informatica, Università di Pisa
`federico.poloni@unipi.it`

Esercizi

Esercizio 1. *Vogliamo tassellare una scacchiera rettangolare* 7×6 *con delle tessere che hanno le seguenti due forme: tessere quadrate* 2×2 *oppure tessere 'ad angolo' costituite da tre caselle, come in figura.*



La tassellazione deve essere tale da coprire completamente la scacchiera e due tessere non devono mai sovrapporsi.

Data una tassellazione, chiamiamo x *il numero di tessere quadrate che vengono usate.*

Dimostrare che se $x \neq 0, 3, 6$ *non è possibile avere una tassellazione. Esibire delle tassellazioni per le quali* $x = 0, 3, 6$.

Esercizio 2. *Si considerino nel piano cartesiano le disposizioni di 35 punti* $\{p_1, ..., p_{35}\}$ *tali che le loro coordinate siano intere e che, presi due qualunque di questi punti, la distanza fra di essi sia un numero intero. Si consideri l'insieme delle coppie*

$$C = \{(p_i, p_j) \mid 1 \leq i < j \leq 35\}$$

Esiste una disposizione con le caratteristiche richieste sopra e tale che per più del 60 per cento degli elementi di C *i punti che formano la coppia abbiano distanza pari ? Stessa domanda con dispari al posto di pari.*

Esercizio 3. *Dato un numero reale positivo* a*, dimostrare che, per ogni intero positivo* n*, almeno uno fra i numeri* $a, 2a, 3a, \ldots, (n-1)a$ *differisce da un intero per meno di* $\frac{1}{n}$.

Esercizio 4. *a) Consideriamo 7 punti nel piano, con la proprietà che ogni retta che passa per due di tali punti ne incontra anche almeno un altro. Dimostrare che i 7 punti sono allineati.*
b) In generale, se consideriamo n punti nel piano ($n \geq 3$) che soddisfano la stessa proprietà si può concludere che sono allineati ?

Esercizio 5. *Si consideri il seguente gioco-passatempo per un solo giocatore. All'inizio del gioco, su una scacchiera infinita ($\mathbb{N} \times \mathbb{N}$), ci sono $k > 1$ pedine sulla casella $(0, 0)$. Il giocatore può fare quante mosse vuole, tutte di questo tipo: può levare una pedina che sta sulla casella (a, b) e sostituirla con due pedine, una piazzata sulla casella $(a + 1, b)$ e l'altra sulla casella $(a, b + 1)$. Una pedina, se è da sola su una casella, diciamo (c, d), fa guadagnare al giocatore $\frac{1}{2^{c+d}}$ euro, se non è da sola non porta alcun guadagno. Lo scopo del giocatore è guadagnare più possibile; all'inizio del gioco possiede 0 euro visto che le pedine sono sovrapposte.*
a) Se $k = 3$ il giocatore può arrivare a guadagnare 3 euro?
b) Se $k = 4$ il giocatore può arrivare a guadagnare 4 euro?

Le soluzioni di questi esercizi compariranno sulla pagina web dell'orientamento dedicata al giornalino

http://www.dm.unipi.it/webnew/it/orientamento/il-giornalino-degli-open-days

verso fine febbraio. Per avere suggerimenti subito scriveteci all'indirizzo e-mail *counseling.dm.unipi@gmail.com*

Giocare col cioccolato: Chomp!

Descrizione e regole

Il gioco che presentiamo riguarda due amici e una tavoletta di cioccolato con un ultimo quadratino avvelenato! Lo scopo sarà quello di evitare di mangiare questo quadratino velenoso e costringere il vostro amico a farlo. Il gioco in questione è stato denominato Chomp da Martin Gardner (si veda [2]) ma la sua invenzione viene attribuita a David Gale.

Cominciamo con la descrizione del gioco. I due giocatori hanno a disposizione, come in figura, una tavoletta rettangolare con m quadratini su un lato e n sull'altro, in cui il quadratino in basso a sinistra è avvelenato.

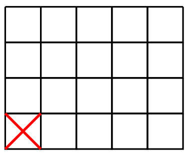


Figura 4: Chomp 5×4

Ogni giocatore ha una *bocca rettangolare*, cioè, scelto un quadratino, è costretto a mangiare tutti quelli che stanno più in alto o più a destra di esso. Ogni *morso* sarà quindi un rettangolo di cioccolata il cui vertice in basso a sinistra è il quadratino scelto dal giocatore. Il primo giocatore dunque avrà a disposizione l'intera tavoletta di cioccolata, e ne mangerà un rettangolino. A questo punto il turno passa al secondo giocatore, che a sua volta dovrà scegliere un rettangolo da mangiare (non è consentita la mossa nulla), e così di seguito. Perde il giocatore a cui rimane da mangiare soltanto il quadratino avvelenato.

In figura è indicata una possibile partita a Chomp con una tavoletta di dimensioni 5×4; con (1) indichiamo il primo giocatore e con (2) il secondo; il rettangolo in grigio è il morso che viene dato dal giocatore.

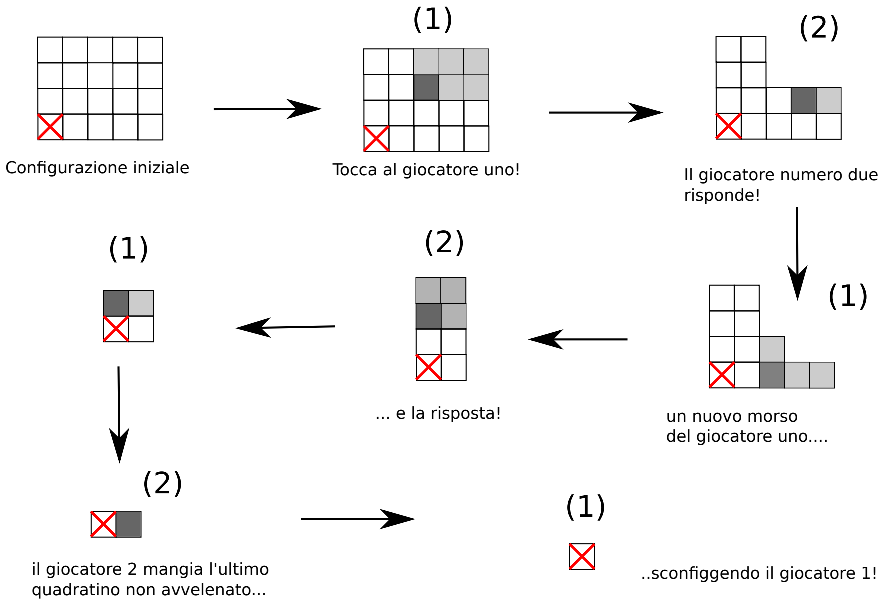


Figura 5: Un esempio di partita con un Chomp 5×4

Provate a fare delle partite! Secondo voi, il giocatore che comincia ha un qualche tipo di vantaggio? Esiste cioè, una strategia che il primo giocatore può seguire che gli assicuri la vittoria?

Chomp è un *gioco combinatorio finito*, è cioè *deterministico*, ha un numero finito di *configurazioni*, ognuna delle quali può comparire solo una volta durante una stessa partita, non esiste la *mossa nulla*, non è possibile pareggiare ed entrambi i giocatori conoscono sempre la configurazione in cui si trova il gioco.

Queste caratteristiche assicurano l'esistenza di una strategia vincente. Per capirne qualcosa in più, cominiciamo col dare qualche definizione.

Chiameremo *configurazione* del gioco ognuna delle forme che la tavoletta di cioccolata può assumere durante una partita. Chiameremo poi *grafo associato al gioco* il grafo contenente tutte le configurazioni possibili, in cui due configurazioni sono collegate da una freccia se con una mossa si può passare dalla prima alla seconda.

Per poter essere sicuri di avere una strategia vincente, è necessario analizzare ogni tipo di configurazione che può venire a crearsi. Tuttavia, il numero delle configurazioni possibili diviene molto complicato all'aumentare di m ed n.

Cominciamo intanto con l'analizzare completamente un esempio pratico: una tavoletta di cioccolata di dimesioni 2×3. In questo caso ci sono soltanto 9 configurazioni possibili, raffigurate nel grafo in figura.

Associamo a ogni configurazione un'eticchetta assegnata nel seguente modo:

- chiamiamo anzitutto *perdente* (P) la configurazione con il solo quadratino avvelenato: il giocatore che si trova in questa situazione purtroppo ha perso;

- chiamiamo *vincente* (V) una configurazione se, da essa, il giocatore può sempre mangiare in modo da riportare il gioco in configurazione perdente;

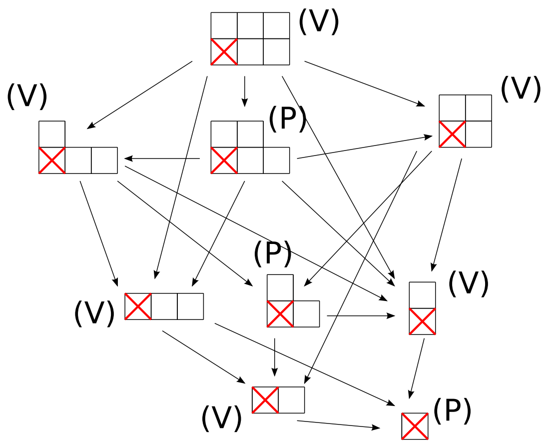


Figura 6: Il grafo delle configurazioni e delle mosse possibili nel caso 3×2

- chiamiamo ancora *perdente* (P) una configurazione se qualunque mossa venga fatta porti l'avversario a trovare il gioco in configurazione vincente.

In questo particolare caso siamo riusciti a contrassegnare ogni configurazione come perdente o vincente, ed inoltre la prima configurazione risulta vincente. Il primo giocatore risulta quindi notevolmente avvantaggiato: giocando bene, ha infatti la vittoria assicurata, a prescindere dall'abilità del suo avversario.

Giochi combinatori finiti

Chiedersi quale dei due giocatori abbia un vantaggio, e se esiste una strategia vincente sono domande naturali quando si parla di un gioco, in quanto ogni giocatore cercherà di fare il possibile per strappare la vittoria all'avversario!

Per quanto riguarda i giochi combinatori finiti, ognuna di queste domande ha una risposta. Qua di seguito enunceremo e dimostreremo dei risultati di carattere generale su questi argomenti.

Teorema. *In un gioco combinatorio finito ogni configurazione risulta vincente o perdente.*

Dimostrazione. Nel grafo di un gioco combinatorio finito devono esistere necessariamente delle configurazioni *finali*, dalle quali cioè non sia possibile fare ulteriori mosse. Se così non fosse, esisterebbero dei *loop* e sarebbe possibile continuare il gioco all'infinito. Inoltre, poichè non è possibile pareggiare, ognuna delle configurazioni finali sarà vincente o perdente.

Consideriamo adesso le altre configurazioni possibili. Chiamiamo $S(1)$ l'insieme delle configurazioni dalle quali, qualunque mossa il giocatore faccia, in un passo si arriva a una configurazione finale. Analogamente, chiamiamo $S(d)$ l'insieme delle configurazioni per le quali esiste almeno un percorso lungo d che arriva ad una finale, ed ogni altro percorso richiede meno di d mosse per arrivare a uno stato finale.

Dato che il grafo è finito, scelto n sufficientemente grande, ogni configurazione appartiene a uno degli $S(1), S(2), \ldots, S(n)$, che sono a due a due disgiunti. Cerchiamo adesso di etichettare ogni configurazione, partendo da quelle appartenenti ad $S(1)$. Scegliamo $s \in S(1)$: se da s, con un'unica mossa, è possibile raggiungere uno stato finale perdente, la etichettiamo come vincente (lo è infatti per il giocatore che deve muovere, poichè può mandare il giocatore successivo in una configurazione finale perdente), altrimenti come perdente. Procediamo allo stesso modo per ogni configurazione in $S(1)$.

Consideriamo adesso $t \in S(2)$. Per come lo abbiamo definito, ogni mossa che parte da t porta a una configurazione finale o in $S(1)$: in ogni caso, in arrivo abbiamo solo configurazioni già etichettate! Se fra tutti i possibili risultati della mossa ne esiste almeno una che sia perdente, etichettiamo t come vincente, altrimenti, come perdente.

In questo modo siamo riusciti a dichiarare vincente o perdente ogni configurazione in $S(2)$. Procedendo per induzione (provate a formalizzare), si può dimostrare che è possibile fare la stessa cosa per ogni configurazione in $S(3), S(4), \ldots$ fino ad etichettarle tutte. $\square$

Questo teorema assicura che in un gioco combinatorio finito come Chomp ogni configurazione è completamente nota. È inoltre possibile provare che nel Chomp il primo giocatore ha a disposizione una strategia vincente. Per tale risultato abbiamo bisogno di alcune definizioni: chiameremo $\mathcal{I}$ l'insieme delle configurazioni del gioco che è possibile raggiungere con una sola mossa dalla configurazione iniziale, e *configurazione pivot* una configurazione $M \in \mathcal{I}$ dalla quale, con un'unica mossa, sia possibile raggiungere solo configurazioni che stanno ancora in $\mathcal{I}$.

Teorema. *In un gioco combinatorio finito che abbia un pivot, il primo giocatore ha sempre una strategia vincente.*

Dimostrazione. Dal momento che, per il teorema precedente, ogni configurazione è vincente o perdente, in particolare lo sarà la configurazione iniziale: se risultasse vincente, allora il primo giocatore, giocando bene, avrebbe la vittoria assicurata, in caso contrario il primo a cominciare potrà solo sperare nella scarsa esperienza dell'avversario. Adesso immaginiamo che il primo giocatore, partendo dalla configurazione iniziale, porti il gioco in una configurazione pivot. A questo punto il secondo giocatore potrà muoversi soltanto verso una configurazione che sta ancora in $\mathcal{I}$, ossia una configurazione che poteva essere raggiunta direttamente dal primo giocatore! In particolare, se la mossa del primo giocatore porta ad una configurazione pivot perdente, esso ha adottato una strategia vincente; viceversa, se la sua mossa porta ad una configurazione pivot vincente, il secondo giocatore potrà, con la sua mossa, portare il gioco in configurazione perdente che però sarà ancora una configurazione appartenente ad $\mathcal{I}$, ossia una configurazione che poteva essere raggiunta anche direttamente dal primo giocatore. Questo dimostra il teorema. □

Questo vale per il Chomp?

Teorema. *Nel gioco Chomp c'è un'unica configurazione pivot.*

Grazie a questi risultati possiamo concludere che in Chomp esiste sempre una strategia vincente e che il primo giocatore è decisamente avvantaggiato: se gioca bene, ha la vittoria assicurata!

Alcuni casi particolari

Abbiamo visto che per il Chomp è possibile sapere se una qualsiasi configurazione sia vincente o perdente, ma costruire il grafo associato ad una determinata configurazione può essere molto complicato, anche con una tavoletta di cioccolato non molto grande.

Possiamo allora trovare esplicitamente la strategia vincente in alcuni casi particolari, a cui sarà in taluni casi possibile ricondursi con mosse appropriate. Vi proponiamo un paio di situazioni, provate a trovare la strada giusta per vincere!

Domanda. *Se avete una tavoletta di cioccolato quadrata, ossia del tipo $n \times n$, sapete individuare una strategia vincente, immaginando di essere il primo giocatore?*

Domanda. *Se avete una tavoletta di cioccolato $2 \times n$, sapete individuare una strategia vincente, immaginando di essere il primo giocatore?*

Per chi volesse provare a giocare contro il computer, segnaliamo la pagina (https://www.math.ucla.edu/ tom/Games/chomp.html. Suggeriamo anche di visitare la pagina personale di Andries Brouwer, dove potrete trovare ulteriori informazioni su questo gioco: (http://www.win.tue.nl/ aeb/games/chomp.html.

Riferimenti bibliografici

[1] E. Delucchi, G. Gaiffi, L. Pernazza, *Giochi e Percorsi Matematici*, Springer, 2012

[2] M. Gardner, *Mathematical Games*, Scientific American, 1973

Agnese Barbensi, Dario Villanis Ziani
laureati triennali in Matematica

Alcuni libri consigliati e pagine web utili

Ecco una breve lista di libri che vi suggeriamo per un ulteriore approfondimento: alcuni contengono delle vere e proprie pagine di matematica scritte con tutti i dettagli, ma in modo da poter essere gustate pienamente da uno studente di scuola superiore, altri sono delle biografie di famosi matematici, la cui vicenda personale ha lasciato un segno nella storia importante quanto il loro contributo scientifico.

- R. Courant, H. Robbins, *Che cos'è la matematica*, Bollati Boringhieri: uno dei libri fondamentali di divulgazione matematica, lo consigliamo, per approfondire e appassionarsi;

- M. Gardner, *Enigmi e giochi matematici*, BUR: un classico, da un grande autore dell'intrattenimento matematico;

- G.H. Hardy, *Apologia di un matematico*, Garzanti: biografia di uno dei maggiori teorici dei numeri del secolo scorso, con uno spaccato della vita del famoso matematico indiano Ramanujan;

- O. A. Ivanov, *Facile come pi greco*, Bollati Boringhieri: problemi ed approfondimenti alla portata di chi ha una preparazione al livello della scuola superiore;

- A. Parlangeli, *Uno spirito puro: Ennio De Giorgi*, Milella: racconto della vita di Ennio De Giorgi, uno dei più grandi matematici italiani, a 20 anni dalla scomparsa, attraverso le testimonianze di chi ha avuto la fortuna di conoscerlo;

- E. Sinibaldi, *IL FIBONACCI. Breve viaggio fra curiosità matematiche*, UMI: raccolta dei bellissimi poster a cura di Franco Conti, pieni di esercizi interessanti, a cui l'autore ha aggiunto le soluzioni;

- A. Weil, *Ricordi di apprendistato. Vita di un matematico*, Einaudi: la biografia di Andrè Weil, uno dei più grandi matematici del secolo scorso.

Per non confondere le idee ci siamo limitati a proporre una bibliografia essenziale. Di lettura in lettura sarete forse voi stessi ad aggiungere altri titoli e a scoprire altri libri a cui rimarrete affezionati.

Per finire, ecco un breve elenco di siti web che vi consigliamo di visitare e dove potrete trovare informazioni, notizie ed esercizi utili:

- Versione on-line del giornalino:
  https://www.dm.unipi.it/webnew/it/orientamento/
  il-giornalino-degli-open-days

- Sito del Dipartimento di Matematica di Pisa:
  http://www.dm.unipi.it/webnew/

- Sito delle olimpiadi di matematica:
  http://olimpiadi.dm.unibo.it/

- Sito della Scuola Normale Superiore di Pisa:
  http://www.sns.it/

- Sito degli studenti di matematica di Pisa:
  http://poisson.dm.unipi.it/

Per ogni ulteriore informazione, come pure per scaricare la versione elettronica di questo giornalino e dei numeri precedenti, vi invitiamo a visitare il sito (http://www.dm.unipi.it/webnew/it /orientamento/home-orientamento).

STAMPATO IN PROPRIO