Dipartimento di Matematica - Università di Pisa

PISA, 16-18 GENNAIO 2019

Matematica
il giornalino degli open days

IN SUPREMÆ DIGNITATIS · 1343

notizie, giochi
e pillole
di matematica

Su indicazione della Commissione Orientamento.

Realizzato con la collaborazione degli studenti counselling:

Luca Bruni
Chiara Giraudo
Filippo Testa
Alessandra Tullini

Coordinamento: Alessandra Caraceni, Giovanni Gaiffi

Grafica: Alessandra Caraceni

Introduzione

Il presente giornalino prende forma grazie alla collaborazione di studenti del Corso di Laurea in Matematica e professori dell'omonimo Dipartimento dell'Università di Pisa, in previsione delle consuete attività di orientamento. Questa ottava edizione, come le sue precedenti, si rivolge non solo a chi mostra già un discreto interesse nei confronti della Matematica, ma anche a chi è incuriosito da una così peculiare disciplina.

A tal proposito, cominceremo con una presentazione del Corso di Laurea in Matematica, accompagnata da una breve discussione riguardo le possibilità per il futuro di chi intraprende questo percorso di studi. In questa occasione citeremo alcune indagini statistiche e le relative fonti.

In seguito troverete alcuni articoli divulgativi il cui scopo è quello di permetterci di dare uno sguardo all'ambiente matematico.

Il primo ci condurrà alla scoperta di tecniche e dimostrazioni fra le più ingegnose in probabilità e combinatoria, fino a rispondere alla domanda: *Quanto ci vuole per mescolare un mazzo di carte?*

Il secondo, *Problemi classici e moderni in teoria dei numeri*, fornisce un'ampia panoramica di risultati aritmetici e di questioni esplorate a fondo dai matematici, ma tuttora traboccanti di mistero: dalle congetture sui numeri primi alle equazioni diofantee, dagli algoritmi di fattorizzazione alla crittografia.

In questo numero, come di consueto, presenteremo brevemente un gioco di carattere matematico; questa volta si tratta di "Ordine e Caos", una sorta di successore (più interessante!) del tris al quale vi invitiamo a sfidare i vostri amici con carta e penna.

A seguire troverete la rubrica *I problemi del giornalino* inaugurata nella scorsa edizione. Qui vi proponiamo alcuni quesiti su cui riflettere, dei quali siamo curiosi di leggere le vostre soluzioni! Potete inviarcele all'indirizzo `LezioniAperteMatematica@gmail.com`.

Infine, per i più curiosi, abbiamo raccolto una lista di pagine web, libri e film per darvi ulteriori spunti interessanti e qualche approfondimento.

Indice

Il corso di laurea in Matematica

«Cosa studi all'Università?»

«Matematica!»

Tutti gli studenti di Matematica, dopo aver sostenuto questa breve conversazione, devono affrontare qualche secondo di tombale silenzio, mentre il loro interlocutore li osserva con sguardo spaventato. La conversazione, solitamente, prosegue in una delle seguenti direzioni:

«Ah, Matematica! Io sono sempre stato imbranato in Matematica.»

oppure

«Wow, devi essere un genio!»

per poi concludersi con dei complimenti. In effetti, si è soliti pensare che il Corso di Laurea in Matematica sia riservato soltanto a chi, da più o meno tempo, mostra grandi abilità nel campo dei numeri e della logica.

È indubbio che tale percorso voglia stimolare lo sviluppo di competenze di questo tipo, ma la vastità della materia permette di descrivere la mente matematica con tanti aggettivi, nessuno dei quali necessariamente accompagnato dall'appellativo di *genio*. È altrettanto vero che aver allenato la mente a condurre certi tipi di ragionamento non può che giovare all'apprendimento. Ciononostante, le caratteristiche essenziali per il successo dello studente sono la determinazione e la forza di volontà; ed in questo si trovano d'accordo sia gli studenti di Matematica che le persone più estranee a questo mondo.

Fortunatamente, possiamo condividere con il lettore il premio che segue tanto duro lavoro: da più indagini risulta che i laureati in Matematica siano soddisfatti della scelta fatta e godano di un ampio spettro di possibilità lavorative. In particolare:

- Secondo l'Occupational Information Network (sito patrocinato dal ministero del lavoro americano, vedi [3]) i matematici si meritano la medaglia d'argento per il lavoro in cui "si guadagna tanto e ci si stressa poco"; infatti sia che lavorino nei centri di ricerca sia che prestino servizio nelle grandi aziende, la media salariale dei matematici è di 88mila euro all'anno, e un indice di stress di 57/100. Da considerare che sono indicati come lavori differenti gli statistici (5° posto, quindi comunque molto alto), i sistemisti (17°) e gli sviluppatori informatici (18°), tutti lavori accessibili dal CdL in matematica.

- I dati di Almalaurea (vedi [2]) riportano che il 94,6% dei laureati magistrali in matematica risulta occupato a tre anni dalla laurea, contro il 94% dei laureati in ingegneria.

- Sempre dal sito di Almalaurea [1] si possono trovare i seguenti dati:
  - In media il 95% degli studenti di matematica di Pisa sceglierebbero di nuovo lo stesso CdL (la percentuale è dell'80% per gli altri CdL).
  - In media il tempo necessario per completare la laurea triennale in matematica a Pisa è di 4,1 anni.
  - Oltre il 95% di coloro che hanno studiato matematica a Pisa negli ultimi anni vogliono proseguire gli studi (contro il 75% degli altri atenei e CdL italiani nello stesso periodo).

Per avere maggiori informazioni su cosa possa fare un laureato in matematica rimandiamo alla pagina web Mestieri dei matematici e ai materiali raggiungibili dalla pagina Matematici al lavoro:

`https://www.mestierideimatematici.it`
`https://www.dm.unipi.it/webnew/it/orientamento/matematici-al-lavoro-0`

Il corso di Laurea a Pisa

Il corso di Laurea in Matematica si divide formalmente in Laurea Triennale e Laurea Magistrale. La prima corrisponde al titolo internazionale *Bachelor's degree* e prevede il conseguimento di 180 Crediti Formativi Universitari (CFU) in tre anni accademici; la seconda, invece, è internazionalmente identificata con la *Master's degree*. Ogni CFU corrisponde orientativamente a 25 ore tra lezioni e studio individuale.

Come in altri ambiti, condizione necessaria per raggiungere un livello avanzato è la specializzazione delle proprie conoscenze, tant'è che alcuni sistemi universitari prevedono che lo studente si concentri su uno specifico settore da subito. L'offerta del Dipartimento di Matematica di Pisa cerca di accompagnare lo studente verso una scelta consapevole del proprio ramo d'interesse, garantendo delle ampie basi attraverso una serie di esami obbligatori, ma lasciando anche la possibilità di saziare il proprio gusto personale con alcuni esami a scelta. In particolare, si ha fin dall'inizio la possibilità di inquadrare i propri studi in uno dei due curricula seguenti, a seconda che si preferisca un percorso classico od uno volto ad applicazioni computazionali:

- il curriculum fondamentale;
- il curriculum computazionale.

Sfruttando la collaborazione con altri Dipartimenti, il primo integra lo studio della matematica con approfondimenti di fisica, mentre il secondo con argomenti di matematica computazionale e informatica. Al momento dell'immatricolazione vi verrà chiesto di scegliere il curriculum a cui iscrivervi. Non preoccupatevi se siete indecisi: come potete notare dalla Tabella 1, i corsi sono identici fino al secondo semestre del secondo anno, quindi avrete tempo per capire cosa vi piace e scegliere di conseguenza.

Naturale prosecuzione è il biennio magistrale, alla fine del quale si acquisisce il titolo omonimo. Esso è diviso in cinque diversi curricula: didattico, modellistico, applicativo, generale e teorico; tramite questi sarà possibile approfondire gli argomenti che più vi hanno appassionati durante i primi tre anni di studio.

Nella Tabella 1 trovate l'elenco degli esami da sostenere durante la laurea triennale, molti dei quali riguardano argomenti non trattati a scuola. Per iniziare a capire cosa studiano queste discipline, quali problemi cercano di risolvere e quali sono alcune delle tecniche usate, consigliamo il libro:

R. Courant, H. Robbins, *Che cos'è la matematica*, Bollati Boringhieri.

Si tratta sicuramente di uno dei migliori libri introduttivi alla matematica, dal carattere divulgativo ma contenente vari teoremi con dimostrazioni vere e proprie che costituiscono dei primi esempi di "vera" matematica.

Fondamentale	**Computazionale**
I anno	
Aritmetica (9 CFU)	
Fondamenti di programmazione con laboratorio (9 CFU)	
Laboratorio di comunicazione mediante calcolatore (3 CFU)	
Analisi matematica 1 (15 CFU)	
Geometria 1 (15 CFU)	
Fisica I con laboratorio (9 CFU)	
II anno	
Algebra 1 (6 CFU)	
Analisi numerica con laboratorio (9 CFU)	
Inglese scientifico (6 CFU)	
Analisi matematica 2 (12 CFU)	
Geometria 2 (12 CFU)	
Elementi di probabilità e statistica (6 CFU)	
Laboratorio didattico di matematica computazionale (3 CFU)	
Esame a scelta (6 CFU)	Algoritmi e strutture dati (6 CFU)
III anno	
Meccanica razionale (6 CFU)	
Fisica II (9 CFU)	Calcolo scientifico (6 CFU)
Fisica III (6 CFU)	Laboratorio computazionale (6 CFU)
Laboratorio sperimentale di matematica computazionale (6 CFU)	Linguaggi di programmazione con laboratorio (9 CFU)
	Ricerca operativa (6 CFU)
4 Esami a scelta (24 CFU)	*3 Esami a scelta* (18 CFU)
Prova finale (9 CFU)	

Tabella 1: Gli esami della Laurea triennale.

Bibliografia

[1] `https://www2.almalaurea.it/cgi-php/universita/statistiche/framescheda.php?anno=2016&corstipo=L&ateneo=70024&facolta=1391&gruppo=1&pa=70024&classe=10032&corso=tutti&postcorso=tutti&isstella=0&disaggregazione=&LANG=it&CONFIG=profilo`

[2] `http://www.ilsole24ore.com/art/notizie/2017-08-16/-statistica-chimica-lauree-che-danno-lavoro-9-studenti-10-131340.shtml?uuid=AEqplaDC`

[3] `http://www.alleyoop.ilsole24ore.com/2017/10/25/ecco-24-lavori-perfetti-ad-alto-tasso-di-guadagno-e-a-basso-tasso-di-stress/`

[4] `https://www2.almalaurea.it`

Quanto tempo ci vuole per mescolare un mazzo di carte?

*di **Alessandra Caraceni**, ricercatrice presso l'Università di Bath, UK*

...e quanto per completare una collezione di figurine? Quanto per mandare in rovina un giocatore d'azzardo?

Lo studio di questi problemi trova le proprie radici nelle origini settecentesche della teoria della probabilità, ma porta ancora oggi sempre nuovi frutti, nella forma di ulteriori risultati, tecniche e applicazioni.

Alla ricerca di risposte, ci imbatteremo in alcuni degli strumenti più efficaci che i matematici abbiano elaborato per catturare con i loro modelli un fenomeno del reale che a lungo li aveva elusi: il caso.

1 Le catene di Markov

Consideriamo un problema che all'apparenza ha poco a che fare col mescolare mazzi di carte: sfidiamoci a testa o croce. A turno, lanciamo una moneta; diciamo che, con grande generosità, lascerò che cominciate voi. Se esce testa, sarò costretta a darvi un euro; se esce croce, sarete voi a darmi un euro. Poi tocca a me, che lancerò una moneta per determinare se sarete voi a dovermi dare un euro o viceversa, e avanti così. Quando uno di noi rimane a corto di euro, il gioco termina.

Come "modellizzare" questo gioco in modo da poterlo analizzare? La quantità di cui ci interessa tenere traccia nel tempo è, chiaramente, quanti euro ciascuno di noi due possieda; in verità, assumeremo che il numero totale di euro rimanga invariato (diciamo che è n all'inizio, di cui voi ne possedete a e io $n-a$); è così sufficiente conoscere il totale e tenere traccia dei vostri euro per poter ricostruire quanti ne abbia io. Ad ogni "mossa" del gioco, supponendo che voi abbiate $0 < k < n$ euro, il lancio della moneta

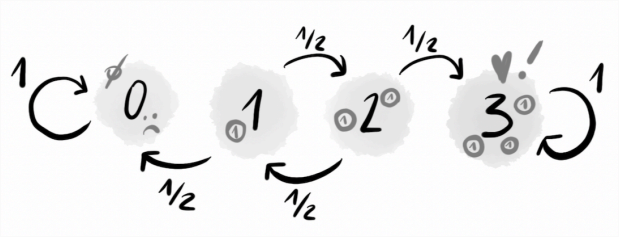


Figura 1: I quattro stati della catena X_t nel caso $n = 3$ e le relative probabilità di transizione.

(che sia il mio turno o il vostro) determina se al turno successivo avrete $k + 1$ o $k - 1$ euro; se la moneta non è truccata, questi due eventi hanno uguale probabilità. Se però avete 0 o n euro, il gioco è finito: rimarrete in questo stato per sempre.

Detto X_t il numero di euro che possedete dopo t turni (dove, diciamo, $X_0 = a$), si tratta per $t > 0$ di una quantità aleatoria che dipende dalla sequenza dei risultati dei lanci che abbiamo effettuato, con la proprietà che, se conosciamo il valore di X_t, sappiamo precisamente quali siano le probabilità per X_{t+1} di assumere ciascuno dei valori possibili; ovvero, il valore di X_{t+1} dipende da quello di X_t, ma non da come il gioco si sia svolto in precedenza: non importa che io sia stata fortunata per tutta la partita; il $t+1$-esimo lancio della moneta determinerà in maniera del tutto imparziale se il vostro capitale aumenta o diminuisce.

Un processo X_t con questa proprietà prende il nome di *catena di Markov*; l'insieme dei valori possibili per X_t, che nel nostro caso è $\{0, 1, \ldots, n\}$, è detto *insieme degli stati*. L'intero processo può essere descritto semplicemente dichiarando, per ogni coppia di stati, la probabilità che, se la catena si trova nel primo stato al tempo t (o al "passo" t, o dopo la t-esima mossa), si trovi nel secondo stato al tempo $t+1$[1]. Un modo molto naturale di rappresentare una catena di Markov è quello in Figura 1.

Una branca molto vivace della probabilità si occupa di studiare il comportamento asintotico, cioè dopo tempi molto grandi, delle catene di Markov. E' chiaro che, per la nostra catena X_t, quello che immaginiamo debba succedere dopo tanto tempo è trovarci nello stato n o nello stato 0. Questo ha a che fare con il fatto che questi due stati sono raggiungibili dal nostro stato iniziale a e che, se la catena parte da 0 o da n, vi rimarrà indefinitamente (si chiamano, incidentalmente, stati *di assorbimento*).

[1] per la verità questa probabilità potrebbe anche dipendere da t, ma noi considereremo solo catene in cui questo non avviene, cosiddette *omogenee* nel tempo.

Una domanda naturale da porsi è quale sia la probabilità di andare prima o poi a finire in 0 a seconda del vostro capitale a di partenza e di quello totale n. La risposta ha una forma sorprendentemente semplice e vi invito a tentare di scoprirla, ma non è questo il genere di questione del quale ci occuperemo in questo articolo. Quello che vogliamo invece studiare è il tempo necessario perché la catena raggiunga uno fra lo stato 0 e lo stato n (si dice *diventi stazionaria* o *raggiunga l'equilibrio*): voglio almeno una stima di quanto tempo dovrò dedicarvi per questo gioco; basta il tempo di un caffè? Devo cancellare i miei piani per la prossima settimana?

Torneremo presto ad analizzare il gioco proposto, ma prima vediamo perché quel che abbiamo detto finora si presti effettivamente ad insegnarci qualcosa sui mazzi di carte e sui modi di mescolarli: perché è possibile parlare di mescolate in termini di catene di Markov?

Beh, immaginate un mazzo di n carte; questo sarà inizialmente ordinato in un certo modo, uno di $n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$ modi diversi[2]. Per mescolarlo sarà necessario compiere una sequenza di mosse in qualche modo "casuali" che modifichino questo ordinamento. Immaginare questa sequenza come una catena di Markov è molto naturale: il nuovo ordinamento del mazzo dopo una mossa di mescolamento dipenderà sicuramente solo da quello immediatamente precedente e non dalle mosse fatte in passato. In più, una mossa di mescolamento dovrebbe essere "abbastanza casuale" perché, se il mazzo è già mescolato (cioè non conosciamo il suo ordinamento, che può essere uniformemente uno degli $n!$ possibili), questo rimanga ben mescolato dopo un'ulteriore mossa; e "abbastanza efficace" perché, dato un qualunque ordinamento iniziale, sia possibile con una sequenza di mosse raggiungerne qualunque altro (altrimenti il mazzo potrebbe non mescolarsi mai!).

Queste richieste sulle mosse sono essenzialmente sufficienti perché, se consideriamo la catena Y_t il cui insieme degli stati è quello degli $n!$ ordinamenti delle carte e i cui passaggi da uno stato all'altro sono dati dalle mosse di mescolamento, alla lunga il mazzo risulti ben mescolato qualunque sia Y_0; ovvero ci aspettiamo che, per t abbastanza grande, la probabilità che Y_t sia un qualunque ordinamento fissato sia, almeno all'incirca, $1/n!$.

Ma quale sarebbe un esempio di "catena di mescolamento" per un mazzo di carte? Una singola mossa potrebbe consistere, per esempio, nel prendere la carta in cima al mazzo e reinserirla "a caso" al suo interno.

[2] Quello che voglio dire è che ci sono n possibilità per la carta che sta in cima al mazzo; scelta quella, $n-1$ possibilità per la successiva, e così via fino all'ultima. Quindi il numero di ordinamenti possibili è il prodotto dei numeri da 1 a n, detto n *fattoriale* o $n!$.

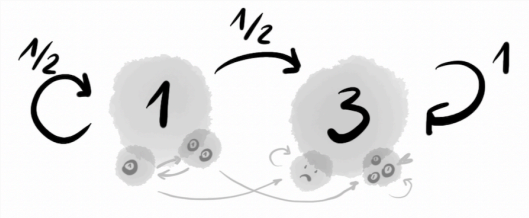


Figura 2: I due stati della catena $\tilde{X}_t$; ciascuno "ingloba" due diversi stati della catena originale X_t.

Questo metodo, spesso chiamato per ovvi motivi *top-to-random*, è il primo che analizzeremo; non sarà particolarmente efficiente, ma è semplice da modellizzare e dovrebbe essere – almeno intuitivamente – chiaro che alla lunga sarà efficace nel mescolare il mazzo. In ogni caso dimostreremo quest'ultimo fatto e analizzeremo in dettaglio il tempo necessario per raggiungere la stazionarietà nella Sezione 4. Nella Sezione 5 concluderemo con l'analisi di un metodo di mescolamento molto più simile a uno reale, che Bayer e Diaconis sono riusciti a trattare in un articolo nel 1992 che fece notizia perfino nel mondo non matematico, guadagnando loro un posto in prima pagina sul New York Times!

Ma per adesso torniamo al nostro gioco di testa o croce; per semplicità, analizzeremo solo il caso $n = 3$, che è il primo caso interessante. La catena X_t che stiamo considerando è, c'è da dire, un po' particolare per il fatto di avere *due* possibili esiti "stazionari": potete vincere tutti i 3 euro o perderli tutti. Un po' perché è in generale preferibile studiare catene che abbiano un'unica distribuzione stazionaria e un po' per mostrare come sia possibile, scegliendo un insieme degli stati diverso, attaccare lo stesso problema analizzando catene diverse, modifichiamo leggermente il nostro modello.

Se quella che ci interessa è la durata del gioco e non la probabilità che questo termini con una vostra vittoria o sconfitta, potremmo decidere di prendere come stato il valore assoluto della differenza fra il mio e il vostro numero di euro. Ma questo può valere solamente 1 o 3, e quando diviene 3 il gioco termina! La nuova catena $\tilde{X}_t$, che stiamo definendo come $|X_t - (3 - X_t)| = |2X_t - 3|$, è quindi semplicissima: la probabilità di passare da 3 a 3 è 1 e quella di passare da 3 a 1 è zero (una volta raggiunta differenza 3, non possiamo più giocare e questa si conserva per sempre). D'altra parte, se uno di noi ha un euro e l'altro due, a ogni mossa c'è una probabilità di 1/2 che chi ha meno euro ne guadagni uno (lo stato rimane lo stesso) e

1/2 che lo perda (si passa allo stato 3), vedi Figura 2.

Supponiamo di partire dallo stato $\tilde{X}_0 = 1$ e sia T il primo tempo tale che $\tilde{X}_T = 3$; allora la probabilità che T valga k, che chiamerò $\mathbb{P}(T = k)$, è la probabilità di rimanere nello stato 1 per $k - 1$ volte e di uscirne alla k-esima, cioè $\frac{1}{2^k}$.

Come calcolare dunque la durata media del gioco, cioè il valore medio di T? Si tratta, appunto, di calcolare una media dei valori possibili per T che dia a ciascuno un peso proporzionale alla sua probabilità di verificarsi. Se T prendesse un certo numero finito di possibili valori, ciascuno con la medesima probabilità, il suo valor medio, anche detto "valore atteso" di T e spesso denotato con $\mathbb{E}(T)$, sarebbe semplicemente la media aritmetica dei valori possibili. In questo caso vogliamo dare a ciascun valore k un peso dato dalla probabilità che si abbia $T = k$, ovvero moltiplicarlo per 2^{-k}.

In pratica, vorremmo conoscere o approssimare la quantità

$$1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + \ldots,$$

ma disgraziatamente non è ovvio come riuscirci!

Quello che conviene fare è usare seguente trucco: si può calcolare $\mathbb{E}(T)$, anziché come somma dei valori $k\,\mathbb{P}(T = k)$ per $k \geq 1$, come somma dei valori $\mathbb{P}(T > s)$ per $s \geq 0$, ovvero come somma delle probabilità che T assuma un valore strettamente maggiore di ciascun numero naturale fissato. Le due somme danno lo stesso risultato perché, dato che $\mathbb{P}(T > s)$ è la somma di $\mathbb{P}(T = k)$ per $k > s$, in entrambe le versioni l'addendo $\mathbb{P}(T = k)$ compare esattamente k volte (nella prima è scritto esplicitamente e nella seconda è "nascosto" dentro ai termini $\mathbb{P}(T > 0)$, $\mathbb{P}(T > 1)$, …, $\mathbb{P}(T > k - 1)$, che sono esattamente k).

Nel nostro caso, visto che $\mathbb{P}(T > k)$ è la probabilità che per (almeno) k volte chi ha meno euro vinca il lancio di moneta e vale quindi $\frac{1}{2^k}$, abbiamo così

$$\mathbb{E}(T) = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 1 + 1 = 2.$$

Il caso generale in cui abbiamo n euro in totale è più difficile da trattare e più adatto a essere risolto con metodi un po' diversi; tuttavia, è ancora possibile ottenere un valore preciso per $\mathbb{E}(T)$. Nel caso siate curiosi, si ha che $\mathbb{E}(T) = a(n - a)$, dove a è il vostro capitale iniziale e $n - a$ il mio (non a caso abbiamo trovato $\mathbb{E}(T) = 2$ e siamo partiti con uno e due euro). Questo significa che, se partiamo con capitali analoghi, ci aspettiamo che

il gioco duri in media un numero di mosse dell'ordine del quadrato degli euro coinvolti.

Tutto ciò che abbiamo imparato in questa sezione ci servirà per stimare il tempo necessario per mescolare un mazzo di carte con la mossa top-to-random. Ma prima ancora ci aspetta un nuovo problema, simile sotto vari aspetti a quello degli euro giocati a testa o croce; dimenticatevi la partita a carte per adesso: è ora di andare a fare la spesa!

2 Il collezionista

> *Manuel è un avido collezionista. Ogni volta che spende 25 euro al supermercato OGrande, ottiene in regalo una statuina raffigurante uno di venti personaggi della saga di Star Wars. Ovviamente si tratta ogni volta di un personaggio casuale, e ancor più ovviamente Manuel è determinato a completare la sua collezione. Assumendo che non possa procurarsi i personaggi in altro modo, quanto dovrà spendere in media per riuscire a collezionarli tutti?*

A prima vista, questo problema non sembra aver molto a che fare con gli algoritmi per mescolare mazzi di carte. Ma abbiate pazienza, miei giovani Padawan! Risulterà che ne abbia eccome.

Anzitutto, non vi sorprenderà il fatto che per modellizzare l'impresa di Manuel si possano tirare in ballo le catene di Markov; e nemmeno, immagino, che il "tempo" della nostra catena sia naturalmente scandito a colpi di 25 euro di spesa.

Potremmo stabilire che gli stati della catena siano tutte le possibili collezioni di Manuel e che il valore X_t della catena al tempo t sia proprio la collezione di Manuel al momento in cui raggiunge i $25t$ euro di spesa, ovvero le t statuine; ad esempio si potrebbe avere:

$X_0 = \emptyset, X_1 =$  , $X_2 =$  , $X_3 =$  , $X_4 =$  , $\ldots$

Questo ci consentirebbe senz'altro di analizzare il problema della stima del tempo T tale che X_T sia la prima collezione a contenere 20 personaggi distinti, ma è un modello inutilmente ricco (e l'insieme degli stati non è nemmeno finito!).

Possiamo invece ignorare completamente ogni aspetto delle collezioni "parziali" di Manuel eccetto il numero di personaggi distinti che possiede, in modo da considerare una catena i cui stati sono gli elementi

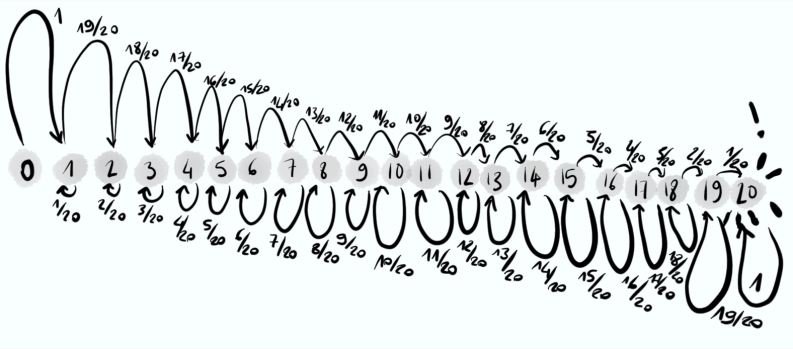


Figura 3: La catena associata all'evoluzione della collezione di Manuel.

dell'insieme $\{0, 1, \ldots, 20\}$. In effetti, dallo stato i (Manuel ha i personaggi distinti) è possibile solamente rimanere allo stato i (Manuel spende 25 euro e ottiene una statuina che possiede già) o allo stato $i+1$ (i prossimi 25 euro di spesa procurano a Manuel un nuovo personaggio).

Supponiamo che si abbia $X_t = i$; a prescindere da quante statuine abbia Manuel in totale (cioè da quanto valga t), dato che i personaggi sono equiprobabili, la probabilità che al prossimo passaggio ne ottenga uno degli i che ha già è $\frac{i}{20}$; la probabilità che ne ottenga uno nuovo è $\frac{20-i}{20}$. Possiamo quindi riassumere la situazione, come nella sezione precedente, con lo schema in Figura 3.

A prescindere dallo stato iniziale X_0 (nel nostro caso $X_0 = 0$), dopo un tempo eventualmente molto lungo che chiameremo T_{20} la catena raggiungerà lo stato 20 e in esso rimarrà per sempre. La domanda che ci poniamo è: quanto vale in media T_{20} (ovvero, se volete, quanto vale $\mathbb{E}(T_{20})$)?

Andiamo per gradi: se chiamiamo T_1 il tempo necessario per raggiungere lo stato 1, abbiamo che $T_1 = 1$. Ma che dire di T_2, cioè il tempo necessario per avere due personaggi distinti? Per ogni k, non è difficile calcolare la probabilità che si abbia $T_2 - T_1 > k$: è la probabilità che si abbia

$$X_{T_1} = X_{T_1+1} = X_{T_1+2} = \ldots = X_{T_1+k} = 1,$$

ovvero che Manuel ottenga sempre lo stesso personaggio per k volte consecutive, cioè 20^{-k}.

Quello che stiamo facendo non è molto diverso dal calcolo della durata media del gioco della sezione precedente; applicando lo stesso trucco possiamo calcolare il valore medio $\mathbb{E}(T_2 - T_1)$ come la somma di $\mathbb{P}(T_2 -$

$T_1 > k$) per $k \geq 0$, cioè

$$1 + \frac{1}{20} + \frac{1}{20^2} + \frac{1}{20^3} + \dots .$$

La somma che si otteneva alla fine della sezione precedente (la somma degli inversi delle potenze di 2) è talmente famosa che ne abbiamo scritto direttamente il valore numerico (cioè 2) senza fare commenti. In questo caso, come possiamo sommare fra di loro le potenze di $1/20$? I termini della somma infinita che vorremmo calcolare sono tutti positivi e diventano rapidamente molto molto piccoli. Se credete nella forza e nel fatto che esista per la somma un valore ben definito S, sarete forse disposti a credere che, siccome possiamo riscriverla come

$$1 + \frac{1}{20}\left(1 + \frac{1}{20} + \frac{1}{20^2} + \frac{1}{20^3} + \dots\right),$$

debba necessariamente aversi $S = 1 + \frac{1}{20}S$, e che il valore che cerchiamo sia $S = \frac{20}{19}$.

Se siete scettici... congratulazioni! Avete la stoffa del vero matematico: non vi resta che calcolare le somme parziali (cioè le approssimazioni finite ottenute sommando fino a ogni valore fissato di k) e convincervi che si tratti di approssimazioni sempre migliori di $\frac{20}{19}$ (cosa che, trattandosi di somme di successioni geometriche, non dovrebbe risultarvi difficile).

Ma a questo punto ci siamo quasi! Calcolare quanto valga in media il tempo $T_{r+1} - T_r$, cioè il tempo trascorso dalla catena nello stato r, non è davvero più difficile. Abbiamo

$$\mathbb{P}(T_{r+1} - T_r > k) = \left(\frac{r}{20}\right)^k$$

(probabilità per Manuel di ottenere uno degli r personaggi che ha già per k volte consecutive) e quindi, sostituendo $\frac{r}{20}$ a $\frac{1}{20}$ nel ragionamento di prima (provate!),

$$\mathbb{E}(T_{r+1} - T_r) = \left(\frac{r}{20}\right)^0 + \left(\frac{r}{20}\right)^1 + \left(\frac{r}{20}\right)^2 + \dots = \frac{20}{20 - r}.$$

Possiamo ora scrivere

$$T_{20} = T_1 + (T_2 - T_1) + (T_3 - T_2) + \dots + (T_{20} - T_{19}),$$

dove ogni addendo è un numero aleatorio che rappresenta il tempo necessario, una volta ottenuti r personaggi diversi, per procurarsene uno

nuovo. E adesso... beh, il valore di T_{20} in media non sarà che la somma delle medie che abbiamo trovato! Ovvero

$$\frac{20}{19} + \frac{20}{18} + \ldots + \frac{20}{3} + \frac{20}{2} + \frac{20}{1},$$

che fa poco meno di 71. In altre parole, in media Manuel dovrà sborsare la bellezza di quasi 1775 euro al supermercato OGrande per completare la sua collezione!

3 Qualche approfondimento sul problema del collezionista

Il problema che abbiamo appena risolto è tipico nell'ambito dello studio della convergenza all'equilibrio delle catene di Markov e aiuta a formalizzare un'intera classe di altri problemi apparentemente diversi. È perciò utile porsi il problema del collezionista più in generale per collezioni con n tipologie distinte di oggetti e chiedersi come vari il tempo necessario per completare la collezione in funzione di n, specialmente quando n è molto grande; a studiare questo problema fra i primi furono figure classiche della matematica quali De Moivre, Eulero e Laplace, il che fa di questa domanda uno dei pilastri della teoria della probabilità.

Il medesimo ragionamento della sezione precedente comporta che, nel caso di n oggetti, il tempo necessario per completare la collezione sia in media uguale a

$$n\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1} + \frac{1}{n}\right).$$

La quantità fra parentesi è un numero che, dato un valore specifico di n, potremmo calcolare (si chiama l'n-esimo numero armonico); se però ci accontentiamo di un suo valore approssimato che ci consenta di valutare il tipo di crescita che ha in funzione di n, potremmo dimostrare (e chi di voi conosce i primi rudimenti di analisi matematica ne sarebbe probabilmente capace) che il suo valore non è troppo diverso da $\log n$.

In particolare possiamo dire che il tempo stimato per completare una collezione di n elementi cresce all'incirca come $n \log n$ (più velocemente di n, ma ben più lentamente di n^2); con qualche strumento matematico in più, è possibile rendere questa affermazione ancora più precisa: si può mostrare che il tempo necessario per completare la collezione, sebbene

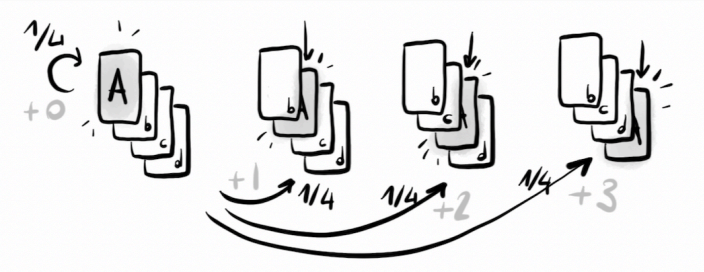


Figura 4

dipenda da quanto siamo fortunati, tenderà con probabilità enorme ad aggirarsi in un certo intervallo intorno a $n \log n$ quando n è molto grande: difficilmente possiamo sperare di fare di meglio, né abbiamo molto da temere.

4 L'algoritmo top-to-random

E' giunto finalmente il momento di occuparci di metodi per mescolare mazzi di carte e in particolare dell'algoritmo top-to-random descritto nella sezione 1, che consiste, dato un mazzo di n carte, nel prendere ripetutamente la prima carta e reinserirla in una posizione casuale.

In altre parole, ad ogni "mossa" scegliamo uniformemente un numero k compreso fra 1 e n (quindi ogni scelta ha la stessa probabilità $1/n$) e spostiamo la carta in cima al mazzo in avanti inserendola dopo le $k-1$ carte immediatamente successive (se scegliamo 1, lasciamo l'ordinamento del mazzo così com'è). Per vedere questo algoritmo come una catena di Markov, è sufficiente considerare come insieme degli stati l'insieme di tutti i possibili $n!$ ordinamenti del mazzo; la nostra catena partirà da un certo stato X_0 e, ad ogni iterazione dell'algoritmo, passerà da X_t a un ordinamento X_{t+1} che è identico al precedente se dal mazzo viene rimossa la carta che si trova in cima secondo l'ordinamento X_t. In Figura 4 vedete rappresentate le transizioni di questa catena da uno stato fissato nel caso $n = 4$.

Come detto in precedenza, al crescere del tempo t ci aspettiamo che il mazzo sia "sempre più mescolato", ovvero che la distribuzione data dalla catena si avvicini sempre di più a quella uniforme sull'insieme degli stati.

Ma quando possiamo dirci soddisfatti e dichiarare che il mazzo "è quasi sicuramente ben mescolato"?

Una prima idea potrebbe essere quella di ripetere l'operazione top-to-random n volte, quante le carte del mazzo; ma è molto probabile che, visto che non tutte le carte vengono reinserite verso il fondo, ci troveremo a spostare alcune carte più di una volta, e quasi certamente non arriveremo a modificare l'ordine relativo delle ultime carte del mazzo entro n iterazioni. No, n proprio non va. Ma $2n$? n^2? 2^n? In quanto tempo siamo ragionevolmente sicuri di farcela?

A partire dal ragionamento fallito sulle n mescolate possiamo farci venire un'idea su come essere sicuri di aver mescolato per bene.

Supponiamo di segnarci il momento in cui per la prima volta scegliamo il numero n, cioè mettiamo la prima carta in fondo al mazzo. Diciamo che questo avvenga al tempo t_1; continuiamo imperterriti a mescolare. Adesso segniamoci il tempo $t_2 > t_1$ in cui per la prima volta (dopo t_1) scegliamo n o $n-1$. Siete d'accordo sul fatto che, al tempo t_2, l'ordine delle ultime due carte sia casuale? Dovreste! Infatti, posto che avremo appena fatto la mossa corrispondente a n o a $n-1$, queste due mosse hanno sempre la stessa probabilità di essere effettuate; dunque la carta che al tempo $t_2 - 1$ era in cima al mazzo si trova ora subito dopo o subito prima della carta che al tempo $t_2 - 1$ era in fondo al mazzo, con uguale probabilità.

Oramai avrete capito l'antifona: definiamo t_3 come il primo momento dopo t_2 in cui effettuiamo una fra la mossa n, la mossa $n-1$ e la mossa $n-2$; al tempo $t_3 - 1$ le ultime due carte del mazzo erano "mescolate", i tre possibili punti di inserzione della carta della cima sono equiprobabili, e dunque le ultime tre carte sono in un ordine casuale al tempo t_3. Una volta definiti $t_4 < t_5 < \ldots < t_n$ allo stesso modo... voilà! Possiamo affermare con certezza che al tempo t_n il mazzo "sia mescolato" (e lo è per sempre dopo quel momento).

Ma i tempi che abbiamo definito vi dicono niente?

Visto che abbiamo definito t_n come il primo istante dopo t_{n-1} in cui venga effettuata la mossa corrispondente a n o a $n-1$ o a $n-2$... o a 2 o a 1 (cioè una mossa qualunque), abbiamo semplicemente $t_n - t_{n-1} = 1$. Per quanto riguarda $t_{n-1} - t_{n-2}$, è il tempo che è necessario attendere perché, scegliendo a caso ad ogni iterazione un numero in $\{1, \ldots, n\}$, ne otteniamo uno compreso fra 2 e n, o più semplicemente diverso da 1; allo stesso modo, la quantità $t_{n-2} - t_{n-3}$ è il tempo da attendere per pescare un numero diverso da 1 e da 2... Infine, per far concludere l'intervallo di tempo t_1 va bene soltanto "pescare" il numero n. Se consideriamo questi

intervalli di tempo in ordine inverso, quindi, non stiamo facendo altro che tentare di completare una collezione di n statuine!

Abbiamo perciò, esattamente come nel caso generale del problema del collezionista,

$$\mathbb{E}(t_n) = n\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right) \approx n \log n.$$

Si noti che, nel caso in cui $n = 52$, si ha che $\mathbb{E}(t_n)$ vale all'incirca 236, ossia siamo riusciti a mostrare che un mazzo di 52 carte è sicuramente completamente mescolato prima di un evento che richiede in media 236 mosse top-to-random per verificarsi.

Cosa? 236 mosse? Che barba!, dirà qualcuno.

Alt!, risponderanno i più attenti fra di voi, *non è mica detto che il mazzo non fosse mescolato prima*.

È in effetti possibile che il mazzo fosse già perfettamente mescolato e che non fosse indispensabile attendere il verificarsi degli eventi con i quali abbiamo definito $t_1, t_2, \ldots, t_n$. Tuttavia, per valori grandi di n si dimostra che non possiamo aspettarci di mescolare il mazzo in meno di $n \log n$ mosse all'incirca: quello che si può scoprire è che, se effettuiamo un numero sensibilmente minore di mosse top-to-random, con grande probabilità la carta che all'inizio era l'ultima del mazzo si troverà ancora fra le ultime $n/\log n$ carte, cosa che invece in un mazzo ben mescolato accade con probabilità $\frac{1}{n} \cdot \frac{n}{\log n} = \frac{1}{\log n}$, cioè molto molto piccola.

D'altra parte, forse c'è un motivo se nessuno, di fronte a un vero mazzo di carte, si metterebbe ad applicare la mossa top-to-random...

5 Il riffle shuffle

A conclusione di questo articolo ci getteremo un'impresa più difficile, quella di analizzare un modello di algoritmo di mescolamento molto più realistico e quindi un po' più complicato. La dimostrazione che presentiamo non è quella piuttosto sofisticata di Bayer e Diaconis, ma una versione un po' più debole ed elementare elaborata da Diaconis e Aldous, che tuttavia ne cattura gli aspetti fondamentali.

Per mescolare un vero mazzo di carte, una procedura molto comune consiste nel dividerlo in due mazzetti e, tenendo uno dei due mazzetti nella mano sinistra e uno nella destra, ricomporre il mazzo con un gesto di scorrimento dei pollici sulle carte il cui effetto è quello di "compenetrare" un mazzetto nell'altro; le carte di un mazzetto si frappongono a quelle

dell'altro in posizioni casuali, ma l'ordine di ogni coppia di carte all'interno di un singolo mazzetto si mantiene.

Modellizzeremo questo algoritmo di mescolamento, spesso chiamato "riffle shuffle", nel seguente modo. Supponiamo di voler mescolare un mazzo di n carte e di partire da un certo ordinamento. Una singola mossa del riffle shuffle consisterà nello scegliere uniformemente a caso una stringa di zeri e uni di lunghezza n; poiché le stringhe possibili sono 2^n, ciascuna avrà probabilità $\frac{1}{2^n}$. Supponiamo che la stringa abbia k zeri e $n-k$ uni; allora separiamo le prime k carte del mazzo dalle ultime $n-k$ e le riordiniamo inserendo le k carte del mazzetto "superiore" in corrispondenza degli zeri della stringa e quelle del mazzetto "inferiore" in corrispondenza degli uni della stringa, mantenendo l'ordine interno di ciascun mazzetto.

Ad esempio, supponiamo di avere $n = 8$, etichettiamo le carte con numeri interi positivi e partiamo dall'ordinamento $1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8$; la mossa corrispondente alla stringa 01101000 consiste nel separare i mazzetti $1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$ (corrispondente ai 5 zeri) e $6 \cdot 7 \cdot 8$ (corrispondente ai 3 uni) e ricomporli come

$$\begin{matrix} 0 \cdot 1 \cdot 1 \cdot 0 \cdot 1 \cdot 0 \cdot 0 \cdot 0 \\ 1 \cdot 6 \cdot 7 \cdot 2 \cdot 8 \cdot 3 \cdot 4 \cdot 5 \end{matrix} .$$

Intuitivamente ci aspettiamo che il riffle shuffle sia un algoritmo più efficiente di quello dato dalla mossa top-to-random e che un numero abbastanza piccolo di mosse (se non altro rispetto a 236...) sia sufficiente nella realtà per essere ragionevolmente sicuri di aver raggiunto un buon livello di mescolamento. Ma *quanto* piccolo?

Per rispondere useremo alcune tecniche che sono centrali nell'area della matematica che si occupa di stimare i "tempi di mescolamento", ovvero i tempi di convergenza delle catene di Markov verso l'equilibrio.

Anzitutto, troveremo più comodo, anziché stimare in quanto tempo il mazzo risulti mescolato via mosse del riffle shuffle, stimare in quanto tempo lo si riesca a mescolare con mosse "inverse": dati un ordinamento e una stringa di zeri e uni, costruiremo l'ordinamento successivo come quello che si ottiene prendendo le carte in corrispondenza degli zeri (nell'ordine in cui si trovano) e mettendole in cima al mazzo. Dovrebbe essere chiaro che questa mossa sia l'inversa di quella descritta in precedenza, ma anche che il problema di stimare il tempo di mescolamento del riffle shuffle sia esattamente equivalente a quello di stimare il tempo di mescolamento di questo riffle shuffle "inverso", che chiameremo, ehm... *elffir shuffle* da ora in poi.

A questo punto, introduciamo l'idea chiave della nostra stima – o meglio, della stima di Diaconis e Aldous! – che è quella di costruire un cosiddetto "accoppiamento". La strategia è questa: supponete di avere il vostro mazzo da mescolare con un certo ordinamento fissato delle n carte, ma supponete anche di avere a disposizione un mazzo premescolato, il cui ordinamento è uniformemente casuale fra gli $n!$ possibili. Adesso immaginate che io vi dia una ricetta per tradurre ogni mossa dell'elffir shuffle da effettuare sul primo mazzo in una singola mossa di elffir shuffle corrispondente da effettuare sul mazzo premescolato. Purché questa ricetta sia fatta in modo che le mosse da effettuare sul mazzo premescolato siano uniformi, l'ordinamento del mazzo premescolato cambierà, ma rimarrà uniformemente casuale (dopotutto, effettuare una mossa del riffle shuffle o dell'elffir shuffle su un mazzo mescolato lo mantiene mescolato).

Ma adesso, se io sono in grado di fornirvi una ricetta tale che, dopo un certo tempo, si possa dire che il primo mazzo ha precisamente lo stesso ordinamento del secondo mazzo, a quel punto anche il primo mazzo sarà per forza mescolato!

Ed è proprio questo che intendo fare. Avete il vostro mazzo di n carte da mescolare e accanto un secondo mazzo di carte immaginario (o più immaginario del primo!) premescolato. Supponete di voler effettuare la mossa corrispondente a una certa stringa di zeri e uni sul primo mazzo; la stringa da far corrispondere per il secondo mazzo è ottenuta così: sbirciate quale sia la prima carta del secondo mazzo e andate a cercarla nel primo; se questa corrisponde a uno zero nella stringa scelta (si sposta nella parte superiore del mazzo) le facciamo corrispondere uno zero anche nella stringa che costruiamo per il secondo mazzo (la stringa comincerà per zero); viceversa, se le corrisponde un uno, scriviamo un uno anche nella stringa della mossa "immaginaria". Facciamo lo stesso per tutte le carte del secondo mazzo: associamo loro la stessa cifra – zero o uno – che è loro associata nella stringa applicata al primo mazzo, indipendentemente dal fatto che avranno posizioni diverse nei due mazzi.

Questa ricetta fa corrispondere a ciascuna stringa di zeri e uni una stringa diversa e quindi assegna probabilità $\frac{1}{2^n}$ a ogni stringa da applicare al secondo mazzo, che rimane perciò correttamente mescolato (ogni ordinamento mantiene la stessa probabilità).

Supponiamo di effettuare una serie di mosse di elffir shuffle sul primo mazzo (e le corrispondenti sul secondo). A ogni singola carta (ad esempio alla donna di picche) toccherà per ogni mossa uno zero o un uno, a seconda che questa vada a far parte della sezione superiore o inferiore nel mazzo; notate che, per come abbiamo costruito la ricetta per la corri-

spondenza delle stringhe, ad ogni iterazione le verrà assegnata la stessa cifra nel primo e nel secondo mazzo (e in effetti potreste sospettare che questa ricetta fosse elaborata apposta per ottenere questa proprietà...).

Per ogni carta, segnatevi la successione di zeri e uni che le viene assegnata mano a mano che iterate mosse di elffir shuffle. Naturalmente è possibile che, dopo per esempio cinque mosse, sia la donna di picche che il re di fiori abbiano avuto l'assegnazione delle cifre $0, 0, 1, 0, 1$, in sequenza; in quel caso non è possibile dire a priori se venga prima la donna o il re dopo le cinque mosse di elffir shuffle: dipende dal loro ordine nel mazzo iniziale. Ma se le sequenze di zeri e uni associate alle due carte sono diverse (ad esempio: la donna ha avuto $1, 0, 1, 0, 0$ e il re ha avuto $0, 0, 0, 1, 0$) sappiamo dire quale si trovi sopra e quale sotto dopo le cinque smazzate, a prescindere dal loro ordine iniziale; nel nostro esempio, sappiamo che dopo la penultima smazzata la donna si è trovata sopra al re (stavano in due sezioni diverse del mazzo); poiché la quinta ha posto le due carte all'interno dello stesso gruppo (quello da mettere sopra, ma questo non è importante), il loro ordine reciproco non è cambiato. Più in generale, se due carte hanno una sequenza di zeri e uni diversa, la carta che si trova più in alto è quella che ha uno zero come ultima cifra "diversa"; ma soprattutto, le due carte risulteranno necessariamente nello stesso ordine nel mazzo "reale" e nel perpetuamente mescolato mazzo "immaginario"!

Ma un momento: questo significa, per quanto detto prima, che non appena le sequenze di zeri e uni associate a ogni singola carta del mazzo sono tutte diverse, siamo sicuri che l'ordinamento del mazzo sia lo stesso del mazzo immaginario, e quindi uniformemente casuale!

Non ci resta che stimare dopo quante smazzate questo evento si verifichi, ma questo non è difficile. Ogni mossa di elffir shuffle, dato che produce una stringa casuale di zeri e uni, associa indipendentemente ad ogni carta una cifra uniformemente casuale; date due carte fissate, qual è la probabilità che, dopo t mosse, le sequenze associate siano uguali? Si tratta della probabilità che le t cifre assegnate alla seconda carta replichino perfettamente quelle della prima, cioè 2^{-t}.

D'altra parte, vi è un modo molto semplice di stimare la probabilità che al tempo t almeno una coppia di carte abbia la stessa sequenza di zeri e uni; tale probabilità non può essere maggiore della somma delle probabilità che questo avvenga per ciascuna coppia. Dato che le coppie sono

$n(n-1)/2$ (perché?[3]), stiamo dicendo che

$$\mathbb{P}\left(\begin{array}{c}\text{dopo } t \text{ iterazioni ci sono}\\ \text{due carte con la stessa sequenza}\end{array}\right) \leq n(n-1)2^{-t-1}.$$

Detto T il numero di iterazioni necessarie perché per la prima volta tutte le sequenze siano distinte si ha quindi

$$\mathbb{P}(T > t) \leq n(n-1)2^{-t-1} \leq n^2 2^{-t-1}.$$

Ma abbiamo mostrato che al tempo T il mazzo "reale" ha precisamente lo stesso ordine del premescolato mazzo "immaginario"; ovvero, qualunque mazzo sarà certamente ben mescolato dopo T iterazioni dell'elffir shuffle, o, equivalentemente, del riffle shuffle.

Possiamo ora scegliere (in funzione di n) un valore di t che ci dia una probabilità che consideriamo soddisfacente di aver mescolato il mazzo dopo t iterazioni. Ad esempio, dopo $2\log_2 n$ smazzate, abbiamo che il mazzo è ben mescolato con probabilità almeno $1 - n^2 2^{-2\log_2 n - 1}$, cioè $1/2$; dopo $2\log_2 n + 2$ smazzate, la probabilità sarà almeno $7/8$ e dopo $2\log_2 n + 6$ più del 99%. Nel caso di un mazzo di 52 carte, per esempio, abbiamo dimostrato che dopo 14 iterazioni del riffle shuffle il mazzo sarà ben mescolato con probabilità maggiore del 90%. Una bella differenza rispetto alle centinaia di iterazioni richieste dall'algoritmo top-to-random!

Bibliografia

[1] D. Aldous, J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs*, 2002, disponibile all'indirizzo `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[2] D. Bayer, P. Diaconis, *Trailing the Dovetail Shuffle to its Lair*, 1992, The Annals of Applied Probability, n. 2, p. 294–313.

[3] Scegliamo la prima carta fra le n possibili, la seconda fra le $n-1$ diverse dalla prima; dividiamo per due perché la coppia (donna di picche, re di fiori) è in effetti la stessa di (re di fiori, donna di picche).

Problemi classici e moderni in teoria dei numeri

di ***Roberto Dvornicich****, professore presso il Dipartimento di Matematica di Pisa*

1 Introduzione

Lo scopo di questo lavoro è di presentare lo stato dell'arte relativamente ad alcuni problemi classici della teoria dei numeri.

È molto difficile descrivere esattamente cos'è la teoria dei numeri, perché la ricerca in questo settore si è allargata in un gran numero di filoni diversi, seppure spesso con interazioni reciproche. Credo perciò che sia più utile evidenziare alcuni di questi filoni:

1. lo studio dei *numeri primi*, delle loro proprietà e della loro *distribuzione*;

2. lo studio delle *equazioni diofantee* (cioè le equazioni per cui non si ricercano tutte le soluzioni reali, ma solo quelle con numeri interi);

3. lo studio dell'*approssimazione diofantea* (cioè della possibilità di approssimare un numero non razionale, per esempio π, mediante frazioni);

4. lo studio delle *proprietà aritmetiche di insiemi di numeri più complessi dei numeri interi*, ma che hanno caratteristiche simili, sia finiti che infiniti, e l'estensione dei problemi diofantei a questi insiemi;

5. le *applicazioni* dell'aritmetica ai *sistemi di trasmissione digitale di dati* (codici e crittografia).

Nel seguito ci occuperemo di toccare alcuni di questi argomenti nella maniera più elementare possibile. Dovendo necessariamente fare delle scelte, abbiamo scelto di trattare:

- l'ipotesi di Riemann sulla distribuzione dei numeri primi
- le congetture "famose" (primi gemelli e congettura di Goldbach)
- i numeri "famosi" (e e π)
- equazioni diofantee: l'ultimo teorema di Fermat ed altro
- test di primalità e algoritmi di fattorizzazione
- applicazioni

2 Numeri primi

L'ipotesi di Riemann (ma sarebbe meglio dire *congettura* di Riemann) è probabilmente il più grande problema aperto della teoria dei numeri. Questa congettura, dovuta appunto a Riemann, risale al 1859.

Per capire di cosa si tratta, facciamo un passo indietro.

Quanti sono i numeri primi? Tutti sanno che sono infiniti. Ma la domanda ha ancora un senso se viene posta in maniera più precisa. Per ogni numero reale positivo x, definiamo $\pi(x)$ il numero dei primi che sono compresi fra 1 e x.

Come varia $\pi(x)$ in funzione di x?

Per esempio, abbiamo la seguente tabella:

x	$\pi(x)$	$x/\pi(x)$
1000	168	6.0
1000000	78498	12.7
1000000000	50847534	19.7

Si può notare che:

(a) la *percentuale* dei numeri primi fra 1 e x decresce con x;

(b) approssimativamente, il rapporto $x/\pi(x)$ è una funzione lineare del numero di cifre di x.

In effetti, esiste un certo argomento euristico, basato su un modello naturale di probabilità (essenzialmente, si assume che, per un generico intero k, gli eventi *"k è divisibile per m"* e *"k è divisibile per n"*, dove m ed n sono primi fra loro, siano eventi indipendenti), che induce alla seguente

Congettura 1. *La probabilità che un numero n sia primo è $\frac{1}{\log n}$, dove il logaritmo è il logaritmo naturale, ossia fatto rispetto alla base e (costante di Nepero).*

Questo fatto è stato dimostrato, ed è il succo del cosiddetto **Teorema dei numeri primi**.

Teorema 1.

$$\pi(x) \sim \int_2^x \frac{1}{\log t} dt \sim \frac{x}{\log x}$$

Il simbolo $\sim$ (asintotico a) sta a denotare un'approssimazione, ma un'approssimazione molto precisa.

Il teorema dei numeri primi si può enunciare in una forma che si dimostra essere del tutto equivalente alla prededente "pesando" ogni numero primo p con un peso uguale a $\log p$:

Teorema 2. $\theta(x) := \sum_{p \leq x} \log p \sim x$.

L'ipotesi di Riemann riguarda la bontà di questa approssimazione.

Essa dice: se il modello probabilistico che abbiamo inventato funziona, esso dovrebbe seguire le leggi della probabilità. La probabilità (legge dei grandi numeri) dice per esempio che se facciamo una serie successiva di n lanci di monete (testa o croce) non solo ci aspettiamo che circa la metà dei lanci diano testa e metà croce, ma anche che lo scostamento rispetto a questo valore atteso sia piccolo. Lo scostamento previsto è circa $\sqrt{n}$.

Sarà vero anche nel nostro caso? È vero, cioè, che nella nostra formula approssimata l'errore che facciamo è al massimo quello che dovrebbe essere, ossia all'incirca $\sqrt{\frac{x}{\log x}}$? O, nella formulazione equivalente, $\sqrt{x}$?

La validità di questa tesi sembra molto plausibile, ed avrebbe conseguenze rilevanti per la conoscenza di un gran numero di problemi collegati ai numeri primi.

Purtroppo, a tutt'oggi, non conosciamo la verità. Non sappiamo se la congettura di Riemann sia vera o falsa, anche se ci sono vari indizi a favore. Il primo indizio è di carattere "filosofico": un modello probabilistico che

si discosta da quello "naturale" dovrebbe essere motivato da fenomeni distorsivi speciali, che nessuno ha mai riscontrato.

Ma ci sono indizi più convincenti dal punto di vista numerico.

Ci sono vari modi equivalenti di formulare la congettura di Riemann: il più famoso di essi è che un insieme infinito di punti del piano (gli zeri non banali della funzione zeta di Riemann) sia in realtà costituito da punti che giacciono tutti *su una medesima retta verticale*.

Allora si può verificare se almeno qualcuno di questi punti giace effettivamente su questa retta. Si sono fatti dei calcoli:

I primi 24.000.000.000.000 punti controllati giacciono effettivamente sulla retta!

Per molti osservatori esterni questa evidenza numerica costituisce una prova oggettiva, ma purtroppo non è così per i matematici. Un'eccezione è sempre possibile, e non sarebbe il primo caso nella storia quello in cui si è trovata un'eccezione non prevista anche quando tutti credevano che ormai non fosse più ragionevolmente possibile pensare che ci fosse.

I numerosissimi tentativi di dimostrazione della congettura hanno dato solo risultati parziali, come per esempio quello di garantire che almeno il 40% dei punti giace sulla retta (specificando cosa vuol dire il 40% di *infiniti* punti).

Se vi volete cimentare a dimostrarla voi, per dir così, in casa, ecco una formulazione alternativa semplice ma assolutamente esatta dell'ipotesi di Riemann, che tutti possono capire. Consideriamo il minimo comune multiplo di tutti di numeri fra 1 ed n e chiamiamolo $M(n)$. Si sa, dal teorema dei numeri primi, che

$$M(n) \sim e^n$$

ossia che

$$\log M(n) \sim n.$$

È vero che l'approssimazione scritta sopra rispetta le leggi normali della probabilità, ossia che la differenza fra $\log M(n)$ *ed* n *è al massimo (all'incirca)* $\sqrt{n}$*?*

3 I primi gemelli

Il problema dei primi gemelli è quello di stabilire se esistano infinite coppie di numeri primi della forma $(n, n+2)$. Per esempio, le coppie

$$(3,5), (5,7), (11,13), (17,19), (29,31), (41,43)$$

sono coppie di primi gemelli.

Purtoppo, anche questo è un problema aperto. È interessante comunque notare che, se il modello probabilistico dei numeri naturali delineato prima fosse adeguato, allora si potrebbe non solo dimostrare che esistono infinite coppie di primi gemelli, ma anche dire "quante" sono.

Denotiamo con $\pi_2(x)$ il numero di coppie di primi gemelli $(n, n+2)$ fatte con numeri minori o uguali a x. Allora si ha la formula euristica (congetturale)

$$\pi_2(x) = 1,320326... \times \frac{x}{\log^2 x}.$$

Anche qui si sono fatti alcuni calcoli; l'ultimo risultato disponibile riguarda un'analisi di tutti i numeri che hanno fino a 15 cifre decimali. L'errore percentuale che dà la formula congetturale è inferiore a un milionesimo!!!

4 La congettura di Goldbach

La congettura di Goldbach dice che

Congettura 2. *Ogni numero pari maggiore o uguale a 4 si può scrivere come la somma di due numeri primi.*

Esiste una congettura analoga per i numeri dispari:

Congettura 3. *Ogni numero dispari maggiore o uguale a 7 si può scrivere come somma di tre numeri primi.*

L'attenzione attuale, specialmente da parte dei dilettanti della matematica, è rivolta verso il problema che riguarda i numeri *pari*. Infatti il problema relativo ai numeri dispari è stato recentemente risolto in maniera definitiva (2013).

IL CASO DISPARI: $2n + 1 = p_1 + p_2 + p_3$.

Analizziamo però come mai la soluzione definitiva sia solo così recente. Nel 1937 il matematico russo Vinogradov ha dimostrato il seguente

Teorema 3. *Tutti i numeri dispari "abbastanza grandi" si possono scrivere come somma di tre numeri primi.*

Cosa si intende con "abbastanza grandi"? Analizzando la dimostrazione di Vinogradov, si vede che essa funziona per tutti i numeri maggiori o uguali di una costante incredibilmente grande, $3^{3^{15}}$ (un numero con quasi 7 milioni di cifre decimali). Si potrebbe pensare che, utilizzando il computer, si possano trattare tutti i numeri minori di $3^{3^{15}}$ e quindi arrivare ad un teorema valido per tutti i dispari maggiori o guali a 7.

Sembra facile, ma non lo è. Vediamo perché.

Supponiamo di voler verificare tramite un computer che tutti i numeri dispari che la dimostrazione lascia in sospeso si possono esprimere come somma di tre numeri primi. Se $x = p_1 + p_2 + p_3$ allora è chiaro che almeno uno degli addendi deve essere maggiore o uguale di $x/3$. Questo significa che dobbiamo avere a disposizione una tavola di numeri primi che arrivi almeno fino a $1/3 \times 3^{3^{15}}$. Ammesso che abbiamo a disposizione i mezzi teorici per farlo, c'è un problema: *gli atomi dell'universo sono "solo"* 10^{80}!

Fortunatamente, dopo Vinogradov altri matematici hanno via via migliorato la costante di riferimento, fino ad abbassarla (Helfgott, 2013) ad un livello accettabile per i nostri computer.

IL CASO PARI: $2n = p_1 + p_2$?

Risolto il caso dispari, è su questo caso che si concentra l'attenzione di molti appassionati. Infatti per questo caso non esiste un teorema analogo a quello del caso dispari. I fatti definitivamente dimostrati hanno una validità minore. Ecco due esempi. Il primo riguarda una variazione del problema:

Teorema 4. *Ogni numero pari "abbastanza grande" si può scrivere come somma di due numeri dei quali uno è sicuramente primo e l'altro o è primo oppure è il prodotto di due numeri primi.*

Anche qui "abbastanza grande" è *troppo* grande per poter verificare tutti i casi esclusi dal teorema.

Il secondo esempio riguarda il numero di possibili eccezioni alla validità della congettura. Definiamo

$$E(x) := \#\{n \leq x \,|\, 2n \neq p_1 + p_2\}.$$

Teorema 5. *Per ogni $\varepsilon > 0$ esiste una costante $C = C(\varepsilon)$ tale che*

$$E(x) \leq C x^{\frac{1}{2}+\varepsilon}.$$

Da questo teorema di deduce che "quasi tutti" i numeri pari si possono scrivere come somma di due numeri primi.

5 I numeri famosi

I numeri famosi e e π (ma il discorso vale per molti altri numeri di uso quotidiano in matematica) non si possono scrivere con esattezza usando il sistema decimale, perché si avrebbe bisogno di infinite cifre. Non si può nemmeno specificare una regola che permetta di calcolare tutte le cifre, come per esempio per

$$\frac{1}{11} = 0,090909090909...$$

perché tale regola non esiste.

Il motivo risiede nel fatto che essi sono definiti con processi di *limite* e non semplicemente tramite le usuali quattro operazioni.

La domanda è: si possono definire esattamente questi numeri usando strumenti puramente *algebrici* (come le quattro operazioni, i radicali, eccetera) ma senza strumenti analitici?

Per specificare il problema abbiamo bisogno della seguente definizione:

Definizione 1. *Un numero (reale o complesso) si dice algebrico se è radice di un polinomio non nullo a coefficienti interi.*

Per esempio, $\sqrt{3}$ è radice del polinomio $x^2 - 3$ e un numero α che soddisfi la relazione $\alpha^5 - \alpha - 1 = 0$ è algebrico (anche se non si riesce a scrivere tramite radicali).

Il problema quindi diventa:

Problema 1. *I numeri e e π sono algebrici?*

La risposta è NO per entrambi i casi.

Il risultato non è inatteso, nel senso che si può verificare che, preso un numero "a caso", è quasi certo (la probabilità è uguale a 1) che la risposta sia no per questo numero.

Tuttavia, come è facilmente intuibile, è assai complicato escludere che un certo numero possa essere radice di uno qualsiasi fra gli infiniti polinomi a coefficienti interi.

La soluzione del problema relativo ad e è datata 1873 (Hermite), quella relativa a π è datata 1882 (Lindemann). Quest'ultima ha conseguenze su un problema posto già dagli antichi greci, il problema della *quadratura del cerchio*.

La quadratura del cerchio

Dato un cerchio di raggio 1, è possibile costruire con riga e compasso un quadrato la cui area sia uguale a quella del cerchio dato (e cioè π)?

Si può dimostrare abbastanza facilmente che, in un sistema di riferimento cartesiano, le coordinate di tutti i punti che si riescono a costruire con riga e compasso *sono* soluzioni di un'equazione $f(x) = 0$, dove $f(x)$ è un polinomio a coefficienti interi.

Se si potesse quadrare il cerchio, si potrebbe costruire un quadrato di lato $\sqrt{\pi}$. Ma né π né la sua radice quadrata (questa è una conseguenza relativamente semplice) sono soluzioni di alcuna equazione di questo tipo.

I problemi aperti, tuttavia, sono sempre più di quelli risolti. Come detto, il problema relativo ad e e π è stato risolto, ma, in pratica, solo per loro e per *pochissimi* altri numeri. Per esempio, non si sa risolvere il problema nemmeno per le combinazioni più semplici che si possono fare con questi due numeri, quali $e + \pi$, $e \cdot \pi$, etc.

6 Le equazioni diofantee

6.1 L'equazione di Fermat

L'equazione diofantea più conosciuta è quella di Fermat:

$$x^n + y^n = z^n.$$

Fermat affermava che, se $n \geq 3$, questa equazione non ha alcuna soluzione con numeri interi ad eccezione di quelle "banali", ossia quelle in cui una delle variabili è uguale a zero (per esempio, $0^5 + 3^5 = 3^5$). La storia di questa equazione è molto lunga, e molto si è speculato sul fatto che Fermat avesse in mente una soluzione del problema da lui stesso posto.

È probabile (ma certamente non è sicuro) che Fermat NON avesse una soluzione. Sta di fatto che il problema è stato risolto solo 350 anni dopo la sua proposizione (Wiles, 1995).

Innanzitutto: qual è l'interesse di sapere se un'equazione come quella di Fermat ha soluzioni, ed eventualmente di conoscere quali?

A questa domanda si potrebbe tranquillamente rispondere: nessuno. Come la stessa cosa si può dire di moltissimi, per non dire quasi tutti, i problemi di matematica. Storicamente, i problemi di matematica sono stati studiati in quanto interessanti *per se stessi*, indipendentemente dalle loro applicazioni pratiche. È altresì vero che molti risultati della matematica *hanno poi avuto* applicazioni pratiche, ma molto spesso applicazioni che non rientravano nell'obiettivo di coloro che vi hanno contribuito, e che non erano nemmeno nella loro immaginazione.

Nel caso del cosiddetto Ultimo Teorema di Fermat (la ricerca delle soluzioni dell'equazione diofantea $x^n+y^n = z^n$) l'interesse puramente speculativo del problema è quello che ha mosso migliaia e migliaia di matematici, professionisti o dilettanti, a dedicarcisi. Col senno di poi si può dire che questo ha contribuito a enormi sviluppi del pensiero matematico, alcuni dei quali hanno avuto *anche* ricadute dal punto di vista delle applicazioni.

Come noto, la dimostrazione di Wiles dell'Ultimo Teorema di Fermat è estremamente lunga e tecnica, e non si può raccontare se non ad un pubblico molto esperto. Perciò ci limitiamo a pochissimi cenni.

Innanzitutto, si tratta di una dimostrazione *per assurdo*.

In secondo luogo, essa usa dei risultati profondi di *geometria*. Che cosa ha a che fare la geometria con un problema puramente aritmetico come questo?

Già negli anni '50 il matematico Frey aveva avuto l'idea di legare l'equazione di Fermat all'equazione di una curva. Una curva, nel piano, si può descrivere tramite un'equazione in due variabili: per esempio, l'equazione $x^2 + y^2 = 1$ descrive i punti di una *circonferenza* (di centro l'origine e di raggio 1).

Frey argomentava così: supponiamo, per assurdo, che l'equazione di Fermat abbia una soluzione (non banale), e che a, b, c siano tre numeri positivi tali che $a^p + b^p = c^p$ (qui l'esponente p è un numero primo diverso da 2, ma si può facilmente vedere che questo è il caso cruciale).

Consideriamo l'equazione

$$y^2 = (x - a^p)(x - b^p)(x + c^p).$$

Le soluzioni di questa equazione formano appunto una curva algebrica, di un genere speciale: una *curva ellittica*.

I geometri classificano le curve algebriche secondo il loro "grado di complessità", un invariante chiamato "genere". Le coniche, che sono le

curve più semplici, hanno genere 0; le curve ellittiche, che rappresentano il livello successivo di difficoltà, hanno genere 1.

Sulle curve ellittiche si sa moltissimo: in particolare, si sa quando due diverse equazioni definiscono curve ellittiche dello stesso *tipo* (cioè sono isomorfe), e si sanno *classificare* tutti i tipi possibili di curve ellittiche.

Sapendo i coefficienti dell'equazione che descrive la curva ellittica, se ne può dedurre il "tipo" (cioè la classe di isomorfismo).

La dimostrazione consiste, essenzialmente, nel far vedere che, se effettivamente si potessero trovare a, b, c come sopra e quindi si costruisse la curva ellittica relativa, questa curva *non potrebbe rientrare in nessuno dei tipi possibili*.

L'interazione fra geometria ed aritmetica, sviluppata enormemente a partire dalla seconda metà del secolo scorso, è uno dei grossi risultati che si sono avuti anche per merito dello studio dell'Ultimo Teorema di Fermat. In particolare, oggi le equazioni diofantee non si studiano più *una alla volta*, ma si raggruppano in famiglie che descrivono insiemi geometrici dello stesso tipo. Per esempio, Faltings ha dimostrato nel 1983 il seguente teorema:

Teorema 6. *Sia $f(x, y) = 0$ l'equazione di una curva di genere > 1. Allora esistono solo un numero finito di soluzioni dell'equazione $f(x, y) = 0$ con x, y numeri razionali.*

6.2 L'equazione di Catalan

Un altro spettacolare risultato recente consiste nella soluzione dell'equazione di Catalan. Catalan (1844) considerava i quadrati

$$1, 4, 9, 16, 25, 36, 49, \ldots,$$

i cubi

$$1, 8, 27, 64, 125, \ldots,$$

le quarte potenze

$$1, 16, 81, 256, \ldots$$

e così via, per riunirli in un'unica successione:

$$1, 4, 8, 9, 16, 25, 27, 36, 49, 64, \ldots$$

Catalan notava che in questa successione ci sono due numeri consecutivi, e cioè 8 e 9.

Ci sono altre coppie di numeri consecutivi in questa successione? Catalan pensava di no. Ed effettivamente Mihăilescu, nel 2001, ha dimostrato che Catalan aveva ragione:

Teorema 7. *Se consideriamo tutti i numeri della forma a^b, dove b è un esponente maggiore o uguale a 2, l'unica coppia di numeri consecutivi è costituita da 8 e 9.*

La spettacolarità della dimostrazione consiste nel fatto che invece, questa volta, si tratta di una dimostrazione *puramente aritmetica*, ed in fondo basata su idee dovute a Kummer intorno alla metà del secolo diciannovesimo (lo stesso Kummer aveva fatto i primi importanti progressi nello studio dell'ultimo teorema di Fermat).

Tuttavia, l'aritmetica dei numeri interi *non basta*: bisogna costruire un'aritmetica su strutture più complesse (i cosiddetti *campi ciclotomici*) ed è lì che si può risolvere il problema.

6.3 Risolveremo tutte le equazioni?

Le soluzioni di problemi così antichi in tempi recenti possono far pensare che siamo vicini a risolvere il problema di tutte le equazioni diofantee.

Non è così. Un problema che Hilbert, nel congresso mondiale dei matematici del 1900, aveva posto in una lista di problemi per il ventesimo secolo era il seguente:

Problema n.10 di Hilbert. *È possibile trovare un algoritmo che determini se una data equazione diofantea in n incognite abbia soluzione?*

Matijaševic, nel 1970, ha risposto di NO. Non esiste, né potrà mai esistere, un modo per risolvere *tutte* le equazioni diofantee.

La dimostrazione di Matijaševic si inquadra nell'ambito della *logica matematica*.

Nel 1936 K. Gödel aveva dimostrato che, nell'usuale sistema di assiomi della matematica, ma anche con qualsiasi altro sistema che si potesse inventare, la matematica ha dei limiti: ci sono degli enunciati di cui non potremo mai dimostrare né che sono veri, né che sono falsi. Si tratta degli enunciati che Gödel ha chiamato *indecidibili*.

Matijaševic ha fatto vedere che esistono dei particolari tipi di equazioni diofantee per cui la questione se abbiano o meno soluzioni è indecidibile.

7 Primalità e fattorizzazione

Uno dei problemi basilari della teoria dei numeri consiste nel determinare se un certo numero n è un numero primo; nel caso in cui non lo sia, di determinare la sua scomposizione in fattori primi. Come accenneremo alla fine, questo problema assolutamente teorico e astratto ha incredibili conseguenze pratiche.

È chiaro che, se il numero dato n è relativamente piccolo, chiunque, o a mano o con l'aiuto di un calcolatore, può rispondere alla domanda. Il problema quindi si pone in termini di *complessità*: dato un numero n di k cifre, quante sono le operazioni necessarie per dare una risposta?

In pratica, *quanto tempo* ci vuole?

Gli algoritmi che riguardano i numeri interi vengono classificati in classi che corrispondono a diversi gradi di complessità (tempo necessario per la loro esecuzione).

La classe **P** è la classe dei problemi per i quali il numero di passi necessario per eseguire l'algoritmo è *polinomiale* rispetto al numero di cifre dei numeri interi che si esaminano.

Nel nostro caso, considerando un numero n con k cifre, un algoritmo che decida se n è primo oppure no si dice polinomiale se si può effettuare con un numero di passi non superiore a una *potenza* di k, per esempio k^2 oppure anche k^{100}.

Esaminiamo l'algoritmo più naturale per decidere se un numero n è primo oppure no:

dividiamo n per 2, per 3, per 4, per 5, e così via: se ad un certo punto troveremo una divisione esatta (con resto zero), allora il numero non sarà primo (ed avremo trovato un fattore di n); se invece tutte le divisioni per numeri minori di n (ma in realtà basta fermarsi alla radice quadrata di n) danno resto diverso da zero, allora il numero è primo.

Quindi l'esecuzione dell'algoritmo, almeno nel caso in cui n sia un numero primo, richiede di fare circa $\sqrt{n}$ divisioni. Se n ha k cifre, diciamo che n è dell'ordine di grandezza di 10^k, ci vorranno quindi circa $10^{k/2}$ divisioni. Per k grande, questo numero è molto superiore a una potenza (qualsiasi) di k.

Se ne deduce che l'algoritmo naturale non è polinomiale, ma *esponenziale*.

Sono stati studiati vari altri algoritmi che "accorciano" il tempo di esecuzione: alcuni *deterministici*, ossia che danno la risposta con assoluta certezza, altri *probabilistici*, ossia che hanno una altissima probabilità di dare

la risposta esatta. Questi ultimi algoritmi sono ovviamente più veloci dei primi, ma bisogna accontentarsi di un grado, se pur minimo, di incertezza.

Tutti i tipi di algoritmi deterministici conosciuti fino a pochissimo tempo fa erano di tipo esponenziale (in realtà, appena migliore); una combinazione ingegnosa dei tipi deterministico e probabilistico porta a degli algoritmi che, nella grande maggioranza dei casi, si possono eseguire in tempo polinomiale, ma che lasciano un numero di eccezioni per le quali è necessario un tempo esponenziale.

Tra la sorpresa generale dei matematici, tre indiani, Agrawal, Kayal e Saxena (in seguito AKS), hanno trovato nel 2002 un algoritmo deterministico per stabilire se un numero è primo oppure no che funziona in tempo polinomiale.

Di questi matematici solo il primo aveva una certa notorietà internazionale, ma forse più per i suoi studi informatici che per quelli matematici; gli altri due sono suoi giovanissimi allievi.

Ma il vero motivo di sorpresa è un altro: l'idea che sta alla base della formulazione dell'algoritmo è così semplice che sarebbe potuta venire in mente a un qualsiasi studente del primo biennio di matematica.

Invece, nel corso di secoli, non era venuta in mente a nessuno!

L'IDEA DELL'ALGORITMO AKS

Si parte da due fatti legati fra loro e noti da secoli. Primo fatto:

Teorema 8. *Se p è un numero primo, allora vale la congruenza*

$$(x+y)^p \equiv x^p + y^p \bmod p.$$

Inoltre, la questa congruenza non vale se al posto di p si prende un numero non primo.

Ricordiamo che due numeri si dicono *congrui* modulo p se divisi per p danno lo stesso resto. La congruenza enunciata sopra dice che, *se p è un numero primo*, il polinomio $(x+y)^p$ ha un termine uguale a x^p, un termine uguale a y^p e tutti gli altri suoi termini hanno coefficienti divisibili per p.

Secondo fatto (piccolo teorema di Fermat):

Teorema 9. *Se p è un numero primo, allora per ogni intero m vale la congruenza*

$$m^p \equiv m \bmod p.$$

Da questi due fatti elementari AKS deducono il loro teorema, che è un semplice esercizio per un normale studente:

Teorema 10. ***(AKS)*** *Siano n ed a due numeri interi senza fattori comuni. Allora vale la congruenza fra polinomi*

$$(x+a)^n \equiv x^n + a \bmod n$$

SE E SOLO SE n è un numero primo.

La difficoltà di ottenere questo teorema, come detto, non consiste affatto nella sua dimostrazione, ma nella sua *invenzione*: bisogna infatti *immaginare* l'enunciato e le sue possibili applicazioni.

Dal teorema AKS è abbastanza chiaro quello che bisogna fare: dato n, provare a vedere se la cosa è vera, per esempio, per $a = 1$. Detto così, questo richiede ancora un tempo troppo elevato, perché bisognerebbe calcolare tutti i coefficienti del polinomio $(x+a)^n$.

Ma, contando su idee presenti in algoritmi precedentemente sviluppati, si vede che in realtà non occorre considerare i coefficienti uno per uno, ma solo un numero assai più limitato di combinazioni fra di loro, e provare a vedere che cosa succede di queste combinazioni se le si divide per numeri piccoli.

Questo porta ad un algoritmo polinomiale.

Nonostante il risultato teorico sia straordinario, l'algoritmo AKS non viene ancora usato nella pratica. Come mai?

Il fatto è che, per testare se un numero con k cifre è primo oppure no, ci vuole un numero di passi che è circa $C \cdot k^{7,5}$, dove C è una costante molto grande. Anche se il numero di passi necessario per gli altri algoritmi è dato da una formula che è sicuramente peggiore per k molto grande, questa dà un risultato migliore, per via della costante C, quando k è relativamente piccolo (un punto di riferimento attuale è $k = 200$).

7.1 Gli algoritmi di fattorizzazione

Quando si usa un test di primalità del tipo di AKS, e si ottiene la risposta "n non è un numero primo", non si individua necessariamente la fattorizzazione di n. Si sa solo che n non soddisfa le proprietà che sono proprie dei numeri primi.

Per avere un algoritmo di fattorizzazione bisogna fare un passo in più. Quello che in realtà serve, se si scopre che un numero n non è primo, è di individuare un suo divisore proprio (cioè diverso da 1 e da n). Infatti, se a è un divisore proprio di n, e dunque $n = ab$ per qualche intero b, si può ripetere l'algoritmo per i numeri a e b al fine di scoprire se essi sono primi

o hanno dei divisori propri. Ripetendo questo ragionamento, con numeri via via più piccoli, si riesce a determinare la scomposizione di n in fattori primi.

Gli algoritmi di fattorizzazione oggi disponibili sono sicuramente molto più efficienti dell'algoritmo "naturale" descritto precedentemente. È forse interessante notare che, tra gli algoritmi più efficienti conosciuti, uno fa un uso sistematico delle curve ellittiche, che abbiamo già incontrato nella discussione a proposito dell'ultimo teorema di Fermat.

Tuttavia, se si eccettuano gli algoritmi *ad hoc* che funzionano solo per numeri di una forma molto speciale, la complessità di tutti gli algoritmi noti è sempre *subesponenziale*, del tipo C^{k^α}, dov C e α sono costanti, con $\alpha < 1$.

7.2 Le applicazioni

Anche se non si può avere una prova sicura che qualcuno non scopra, prima o poi, un algoritmo di fattorizzazione di complessità polinomiale, al giorno d'oggi la fattorizzazione di un numero rimane uno dei problemi più complessi (nel senso di "*time-consuming*"), ed è su questa convinzione che si basa una delle applicazioni della teoria dei numeri più diffusa, la crittografia.

La crittografia si occupa di trovare dei metodi efficienti per trasmettere dei messaggi, o comunque delle informazioni, in modo codificato, in modo tale che una persona che sia in possesso di uno strumento (chiave di lettura) per decodificare le informazioni le possa decifrare, ma una persona che non conosca la chiave di lettura no.

L'uso della crittografia è storicamente provato fin dal tempo degli antichi romani, per scopi militari.

Oggi se ne fanno diversi usi: insieme a quello militare e di spionaggio, quelli preponderanti sono per le transazioni di carattere economico, per garantire la privacy, per un controllo di sicurezza dell'identità degli individui ammessi a certi servizi.

Una carta bancomat, un acquisto on-line con una carta di credito, l'uso della password nell'aprire un computer, per scaricare files, per leggere la posta elettronica o per entrare in alcuni siti internet sono esempi quotidiani dell'uso della crittografia.

Un sistema crittografico *efficiente* deve rispondere ai seguenti requisiti:

1. rendere *facile* l'uso del sistema da parte dei suoi utenti autorizzati; in particolare, per il mittente di un messaggio deve essere facile codificarlo, per il ricevente deve essere facile decodificarlo;

2. rendere *estremamente difficile*, per non dire impossibile, decodificare dei messaggi se non si conosce la chiave di interpretazione.

Il primo obiettivo si risolve facilmente trasformando le parole in numeri di formato limitato, usando le congruenze, ed usando le normali operazioni su di esse.

Per raggiungere il secondo obiettivo, pure con molte varianti, la scelta è quella di usare la difficoltà della fattorizzazione dei numeri interi. In pratica, seppure con molte varianti, sia chi codifica che chi decodifica (conoscendo la chiave) deve fare delle semplici operazioni di moltiplicazione. Ma per decodificare è necessaria una chiave, che può essere scoperta solo se si riescono a fattorizzare numeri molto grandi.

A questo criterio è ispirato il primo sistema crittografico a chiave pubblica (detto RSA dagli inventori Rivest, Shamir e Adleman, 1978), che ha ispirato un gran numero di varianti che sono state usate in maniera massiccia negli ultimi quarant'anni.

Ultimamente spuntano all'orizzonte degli algoritmi quantistici, ma questo è un campo che esula da questa trattazione elementare.

Bibliografia

[1] L. Childs, *Algebra, un'introduzione concreta*, ETS Editrice.

[2] H. Davenport, *Aritmetica superiore: un'introduzione alla teoria dei numeri*, Zanichelli.

[3] R.L. Graham, D.E. Knuth, O. Patashnik, *Matematica discreta (Principi matematici per l'informatica)*, Hoepli.

[4] G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*, Oxford University Press.

[5] N. Koblitz, *A course in number theory and cryptography*, Springer Verlag.

[6] P. Ribemboim, *The new book of prime number records*, Springer Verlag.

Ordine & Caos

di ***Luca Bruni*** *&* ***Chiara Giraudo****,*
studenti presso il Dipartimento di Matematica di Pisa

Per questo numero abbiamo scelto un gioco semplice ma divertente, da fare ovunque con carta e penna. Il gioco si ispira al più famoso *Tris*. Mostreremo una possibile partita e la commenteremo. Non daremo però l'algoritmo per la strategia ottimale, così da incuriosirvi e da farvi giocare per cercare voi stessi le strategie migliori.

1 Un antenato famoso: il gioco del *Tris*

Il *Tris* è un gioco semplice, da fare con carta e penna. Si gioca in due su una griglia 3×3. Prima di cominciare ogni giocatore sceglie il proprio simbolo: solitamente sono una croce ("X") e un cerchio ("O"). A turno i giocatori scelgono una casella vuota, dove segnare il proprio simbolo. Vince il giocatore che riesce a mettere tre dei propri simboli in fila (in riga o in colonna o in diagonale).

Può accadere che nessun giocatore riesca a disporre i propri simboli in linea retta e pertanto la partita può terminare in parità.

Ogni partita in cui entrambi i giocatori giocano con una strategia ottimale finirà in parità. Vediamo le prime mosse di una strategia ottimale.

Supponiamo che il primo e il secondo giocatore abbiano scelto rispettivamente come propri simboli la "X" e il "O".

Come prima mossa il primo giocatore deve disegnare il proprio simbolo nella casella centrale oppure in uno dei quattro angoli.

Il secondo giocatore, se non vuole perdere, ha la sua prima mossa obbligata. Se il primo giocatore ha scelto il centro, allora il secondo giocatore deve disegnare il "O" in un angolo a scelta. Mentre se la "X" si trova in un angolo, il secondo giocatore può disegnare il "O" solo nella casella centrale.

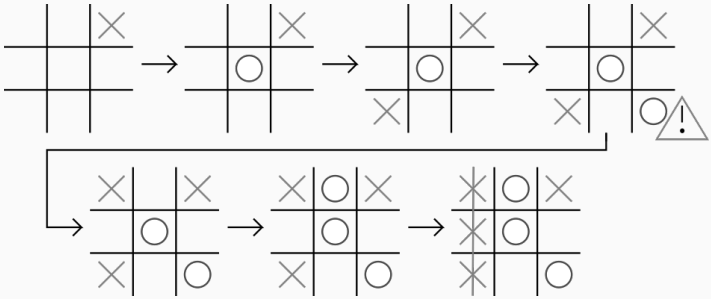


Figura 1: Un esempio di partita a *Tris*

Ad esempio in Figura 1 possiamo vedere come si svolge una partita a *Tris*. Possiamo notare come il secondo giocatore faccia la sua prima mossa corretta per pareggiare la partita, ma la sua seconda è errata e permette la vittoria del suo avversario!

2 Regole del gioco

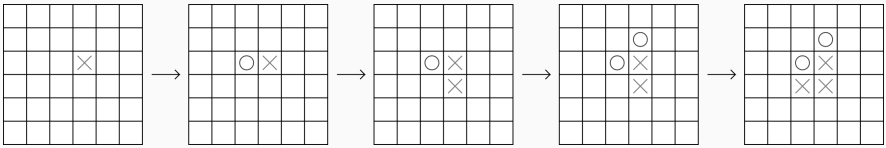
Ordine & Caos è un gioco con carta e matita e può essere considerato una variante del gioco del *Tris*. Questo gioco è stato inventato da Stephen Sniderman ed è stato reso pubblico nel 1981.

Si gioca su una griglia 6×6. Lo scopo del giocatore *Ordine* è quello di creare una linea retta di cinque simboli uguali "X" o "O", il giocatore *Caos* deve impedirne la riuscita. La linea retta può essere o orizzontale o verticale o diagonale. Contrariamente al gioco del *Tris*, ogni giocatore può usare entrambi i simboli ad ogni mossa.

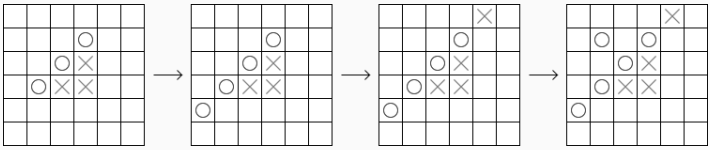
Ordine è sempre il primo a giocare, poi i due giocatori si alternano. Ad ogni turno un giocatore deve disegnare un simbolo qualsiasi tra "X" ed "O", in uno spazio vuoto della tabella. Una volta disegnati i simboli non possono più essere cancellati.

3 Una possibile partita

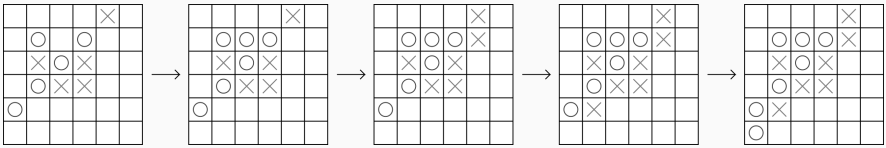
Vediamo adesso un esempio di partita.



Il trovarci di fronte a una scacchiera 6×6 già ci spaventa in quanto le mosse iniziali possibili sono molte. *Ordine* decide di mettere la sua prima "X" in una delle 4 caselle centrali. *Caos* risponde allo stesso stesso modo mettendo un cerchio "O": con questa mossa ha già impedito la realizzazione di 5 segni di fila (sia "X" che "O") nella riga in questione. *Ordine* mette una nuova "X" sempre in una delle 4 caselle centrali e *Caos*, per contrastarlo, blocca la sua colonna con un cerchio. Osserviamo che però il blocco crea due segni "O" in fila sulla diagonale che potrebbero tornare utili al suo avversario. *Ordine* insiste con la sua strategia di "X" al centro e ne inserisce un'altra.

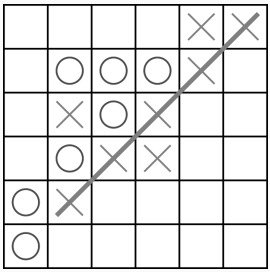


Caos cerca di pararsi e di bloccare sul nascere l'iniziativa di *Ordine* sulla terza riga a partire dal basso apponendo un cerchio. Di nuovo però con questa mossa favorisce il suo avversario! Ci sono infatti 3 cerchi in fila e *Ordine* ne approfitta per forzare l'avversario: ponendo un cerchio in basso nella diagonale, *Caos* è obbligato a piazzare una "X" in alto.



A questo punto *Ordine* è pronto a tessere la propria tela: decide di posizionare il cerchio in un "quasi angolo" strategico. In questo modo attua una duplice minaccia di completamento di due terne. *Caos* corre ai ripari e blocca la colonna con una "X" fra i due cerchi. *Ordine* continua imperterrito con la sua strategia e completa la terna di cerchi. *Caos* è costretto a bloccarlo perché altrimenti alla mossa successiva avrebbe perso. Decide quindi di mettere una "X" alla destra della terna, ma è proprio quello che

Ordine voleva. Si sono formate infatti 3 croci in diagonale e con la quarta messa da *Ordine*, *Caos* non ha più modo di bloccare il suo avversario!



Osserviamo che *Caos* in questa partita non ha giocato al meglio delle sue possibilità; a volte, per bloccare l'avversario, sarebbe stato opportuno effettuare scelte diverse (viste le molteplici possibilità).

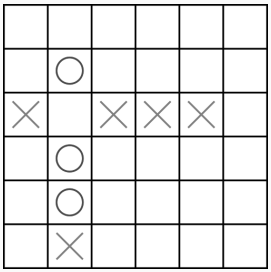
4 Commenti e strategie

Come si può notare dalla partita analizzata il *Ordine & Caos* è molto più complesso del più famoso *Tris*. Una tra le tante difficoltà per i giocatori (soprattutto per *Caos*) è quella della possibilità di inserire entrambi i simboli; a causa di questo fatto *Caos* deve stare molto attento al modo in cui blocca l'avversario: le sue mosse potrebbero ritorcerglisi contro.

Una buona osservazione che si può fare è quella che ci sono delle caselle "più importanti" di altre: le caselle centrali, infatti, fanno parte ciascuna di 6 possibili allineamenti di 5 segni contro i 3 o 4 allineamenti per le caselle più esterne.

Dato che sia *Ordine* che *Caos* possono inserire entrambi i segni, una buona regola per *Caos* è quella di cercare di mantenere il numero di "X" e il numero di "O" uguale durante la partita: se ci sono più "X" in gioco, è, in linea di massima, più probabile un loro allineamento.

Per *Ordine* un modo per poter vincere (supponendo che *Caos* non commetta gravi errori) è quello di riuscire a costruire una situazione in cui con una sola mossa è in grado di creare due allineamenti da 4 (che non saranno pertanto bloccabili dall'avversario). Un altro modo interessante per giungere alla vittoria è riuscire a creare una situazione simile alla seguente (è il turno di *Caos*):



Caos è obbligato a bloccare le "X" di *Ordine* con un cerchio, ma così facendo crea un allineamento di 4 cerchi consegnando la vittoria all'avversario!

Un'ultima osservazione che possiamo fare è che a prima vista il gioco potrebbe sembrare equo: le possibilità di vittoria per *Caos* e *Ordine* sembrano infatti più o meno le stesse. In realtà il gioco ammette una strategia vincente per uno dei due giocatori. Questo risultato è stato provato grazie all'utilizzo di un calcolatore: in particolare sono state ideate alcune strategie per ridurre sensibilmente i casi da analizzare e poi, grazie a un calcolatore, sono state esplorate tutte le casistiche rimanenti. In questo articolo non riportiamo chi tra *Ordine* o *Caos* ha sempre la possibilità di vincere, ma invitiamo il lettore a provare a formulare una propria teoria e strategia divertendosi a giocare con gli amici a questo simpatico gioco!

I problemi del giornalino

una rubrica a cura di ***Davide Lombardo****,*
ricercatore presso il Dipartimento di Matematica di Pisa

Mandateci le vostre soluzioni all'indirizzo `LezioniAperteMatematica@gmail.com`!

1 Divertissement

1.1 Monete alla cieca

Alessandra partecipa ad un gioco a premi: viene portata bendata di fronte ad un tavolo, sul quale – le viene detto – ci sono 100 monete, di cui 80 mostrano "testa" e le altre 20 "croce". Le viene chiesto di manipolare le monete come vuole (ma naturalmente senza togliersi la benda!) e, alla fine delle sue operazioni, separare le monete in due gruppi in modo tale che il numero di "teste" in un gruppo sia uguale al numero di "teste" nell'altro. Ce la farà? Sapreste aiutarla?

1.2 Un po' di geometria

Sia ABC un triangolo con $\widehat{BAC} > \widehat{ACB}$ e sia D il punto di BC tale che $\widehat{BAD} = \widehat{ACB}$. Gli assi di AD e AC si intersecano nel punto E. Provare che l'angolo $\widehat{BAE}$ è retto.

1.3 Una curiosa coincidenza

1. Si può osservare che $1\cdot2\cdot3\cdot4+1 = 25 = 5^2$ e $2\cdot3\cdot4\cdot5+1 = 121 = 11^2$. Si tratta di un fenomeno isolato o queste uguaglianze fanno parte di uno schema generale?

2. Trovare tutti gli interi positivi n tali che $n(n+1)(n+2)(n+3)$ sia un quadrato perfetto (ovvero sia il quadrato di un intero).

1.4 Divisibilità

Sia n un intero positivo. Consideriamo i numeri fra 1 e $2n$: dimostrare che, comunque si scelgano $n+1$ di questi $2n$ interi, se ne sono presi due che non hanno divisori in comune (due interi a e b *non hanno divisori in comune* se l'unico intero positivo che divide sia a che b è 1).

Più difficile: dimostrare anche che comunque si prendano $n+1$ interi nell'insieme $\{1, \ldots, 2n\}$ se ne sono scelti due tali che uno divide l'altro.

2 Qualche apertura verso la matematica non elementare

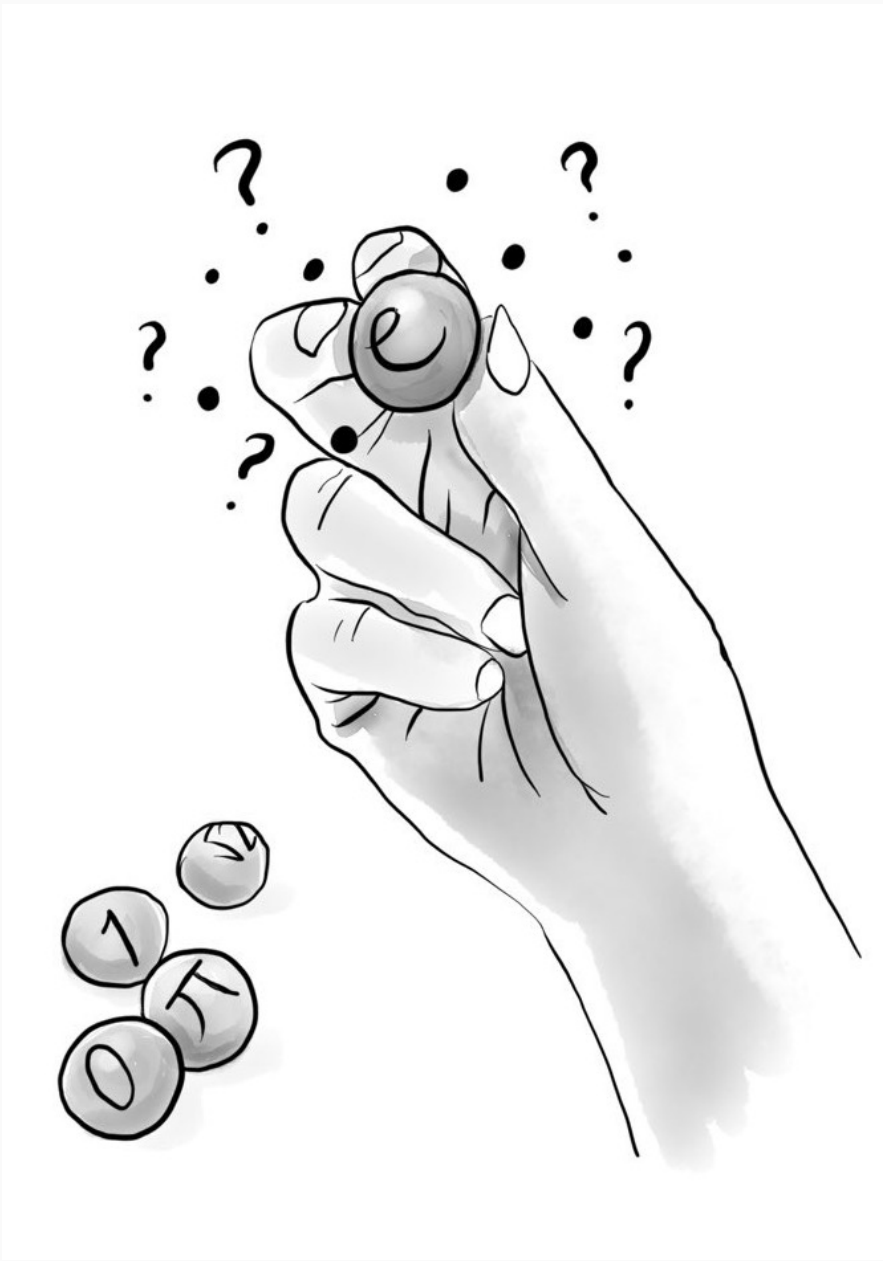
2.1 La lotteria del sultano

Un sacchetto contiene 1001 palline numerate con interi positivi (distinti, ma sulla cui grandezza non sappiamo niente). Il sultano vi sfida al seguente gioco: vi è concesso estrarre una pallina dopo l'altra e leggere il numero che riporta, e dopo ogni estrazione potete decidere se fermarvi (tenendo l'ultima pallina estratta) o andare avanti, estraendo un'altra pallina. Ovviamente nel momento in cui estraete una nuova pallina scartate la precedente, e una volta estratta la milleunesima pallina siete obbligati a fermarvi.

Quando decidete di fermarvi, il sacchetto viene aperto e tutti i numeri rivelati. Se la pallina che avete in mano è quella con il numero più alto in assoluto avete vinto una quantità di kuruş (la moneta dei sultani) pari al vostro peso, altrimenti non avete vinto nulla. Le uniche informazioni che avete sono le regole del gioco e il numero di palline nel sacchetto.

A prima vista, sembra che le vostre probabilità di vittoria siano minime: tanto per incominciare, non avete nessuna idea di quali numeri possano essere scritti sulle palline nel sacchetto! Dimostreremo invece che in realtà esiste una strategia che permette di avere sempre una probabilità di vittoria di almeno $1/4$ – e questo perfino se invece di mille le palline fossero un milione!

Pensateci un po' prima di girare pagina e leggere un indizio... siamo sicuri che, quando ve la diremo, concorderete che si tratta di una strategia molto ragionevole!


e
1
0

Lavoriamo con un numero generico n di palline. L'idea è quella di fissare una soglia $r < n$, guardare (e scartare) le prime r palline estratte – tanto per "farsi un'idea" di quanto siano grandi i numeri "tipici" scritti sulle palline – e poi fermarsi non appena si estrae una pallina con un numero più grande di tutti quelli visti fino a quel momento (se si trova... altrimenti vuol dire che la pallina migliore era fra le prime r, e purtroppo è andata!). Vogliamo ora scegliere r per avere una ragionevole speranza di vittoria.

1. Dimostrare che con questa strategia, se le palline totali sono n e le palline che guardiamo per farci un'idea sono r, la probabilità di vittoria è

$$\frac{r}{n}\left(\frac{1}{r}+\frac{1}{r+1}+\cdots+\frac{1}{n-1}\right)=\frac{r}{n}(H_{n-1}-H_{r-1}),$$

   dove il simbolo H_n indica la somma $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}$.

   Indicazione. Come si calcola questa probabilità? Immaginiamo di continuare a estrarre fino alla fine (anche dopo aver deciso di fermarci, tanto per vedere come sarebbe andata) e dividiamo in casi a seconda di quando estraiamo la pallina con il numero più alto. Se è la prima che estraiamo non c'è niente da fare. Nemmeno se è la seconda, o la terza, o la r-esima. Se è la i-esima con $i > r$ può ancora andarci bene: serve però che non ci siamo già fermati! Quindi serve che la migliore fra le prime $i-1$ palline sia uscita fra le prime r, altrimenti ci saremmo accontentati...

2. Mostrare che per ogni $n \geq 1$ si ha $H_{2n} - H_n \geq \frac{1}{2}$.

3. Supponendo (tanto per fissare le idee) che $n = 2k+1$ sia dispari, dimostrare che possiamo scegliere r in modo che la nostra probabilità di vittoria sia almeno $1/4$.

Commento. Questo è un problema famoso, spesso citato come *il problema delle segretarie* o anche, vista una diffusa formulazione più romanzesca del quesito, *il problema della dote del sultano*. La scelta di r suggerita qui sopra, per quanto consenta di avere una probabilità di vittoria non trascurabile, non è la migliore: la strategia ottimale richiede infatti di guardare le prime $r \approx \frac{n}{e}$ palline, e per n grande garantisce come probabilità di vittoria uno strabiliante 36% (qui e è la costante di Eulero, o numero di Nepero, $e = 2.71828...$, e 36% è un'approssimazione di $1/e = 0.3678...$). Questo problema rientra nell'ambito di quella che viene chiamata *teoria*

dell'arresto ottimo, uno strumento importante non solo per considerazioni astratte ma anche nelle applicazioni, ivi compreso – per esempio – lo studio matematico delle transazioni finanziarie.

Inoltre, con tecniche più raffinate si può dimostrare che quando n è grande la differenza $H_{2n} - H_n$ è approssimativamente uguale a $\log(2n) + \gamma - (\log n + \gamma) = \log 2 \approx 0.693...$, quindi (quando n è grande) la 'vera' probabilità di vittoria garantita dalla strategia appena descritta è circa $\frac{1}{2} \log 2 \approx 0.3465...$, ovvero quasi il 35%!


LE SCOPERTE · LE INVENZIONI
MARIO LIVIO
LA SEZIONE AUREA
Giorgio Tomaso Bagni
Storia della Matematica
GODFREY H. HARDY
MARCUS DU SAUTOY
L'ENIGMA DEI NUMERI PRIMI
L'ipotesi di Riemann, il più grande mistero della matematica
CHE COS'È LA MATEMATICA?

Alcuni consigli: libri, pagine web e altri media

Raccogliamo ora una breve lista di libri, pagine web e film che possono essere uno spunto per ulteriori approfondimenti. Alcuni contengono delle vere e proprie pagine di matematica, altri invece sono biografie di celebri matematici o trattano di argomenti "più leggeri".

- C. B. Boyer, *Storia della Matematica*, Mondadori.
- 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. du Sautoy, *L'enigma dei numeri primi*, BUR: storia, problemi ed applicazioni sulla ricerca dei numeri primi con una notevole enfasi sull'ipotesi di Riemann.
- 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.
- M. Livio, *La sezione aurea*, BUR: Un percorso storico su uno dei numeri che ha maggiormente affascinato l'intelletto umano.

- G. Lolli, *Tavoli, sedie, boccali di birra. David Hilbert e la matematica del Novecento*, Raffaello Cortina Editore: Hilbert è stato protagonista di una straordinaria impresa intellettuale, che ha messo a nostra disposizione nuovi strumenti per indagare la realtà che ci circonda come la precisazione dei linguaggi, delle tecniche e dei problemi della logica matematica.

- 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.

- S. Singh, *Codici e segreti. La storia affascinante dei messaggi cifrati dall'Antico Egitto a Internet*, BUR: dal Cifrario di Cesare ai moderni metodi di Crittografia, scopriamo come la matematica permetta di proteggere la nostra privacy.

- 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 André 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.

Negli ultimi anni sono stati prodotti molti film a tema matematico. Eccone alcuni, dai classici alle perle poco note.

- D. Aronofsky, π – *Il teorema del delirio*, 1998.

- M. Brown, *L'uomo che vide l'infinito*, 2015.

- R. Howard, *A beautiful mind*, 2001.

- M. Martone, *Morte di un matematico napoletano*, 1992.

- M. Tyldum, *The imitation game*, 2014.

- G. Van Sant, *Will Hunting – Genio ribelle*, 1997.

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

- Sito di Maddmaths! Matematica, Divulgazione, Didattica:
  `http://maddmaths.simai.eu/`
- 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:
  `https://poisson.phc.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`