Questo articolo è stato pubblicato originariamente
su «American Scientist», vol. 91 n. 1, gennaio-febbraio 2003,
con il titolo *On the Threshold*.

Ai limiti del calcolo

di Brian Hayes

È ancora un mistero perché esistano problemi computazionali intrinsecamente difficili; ma una nuova classe di algoritmi promette di allontanare un poco la frontiera dell'intrattabilità

La notte scorsa ho telefonato al supporto tecnico dell'universo per segnalare un baco. Mi hanno messo in attesa per un'eternità, ma alla fine ho potuto presentare il mio reclamo: alcune cose in questo mondo richiedono un tempo di calcolo semplicemente troppo lungo, nei casi peggiori addirittura esponenziale. «Non è un baco, è una caratteristica voluta» è stata l'inevitabile risposta. «Impedisce che l'universo si esaurisca troppo rapidamente. E poi le elaborazioni NP-complete sono un'opzione non supportata, il che invalida la sua garanzia. E dove sta scritto che una cosa qualsiasi deve essere calcolabile in modo efficiente? Si consideri fortunato che 1 + 1 è un calcolo fattibile in tempo polinomiale.»

Forse il supporto tecnico cosmico ha ragione: risposte rapide e semplici a problemi computazionali non sono qualcosa che abbiamo il diritto di aspettarci in questo mondo. Eppure, lascia perplessi il fatto che alcuni calcoli siano tanto più complessi di altri. L'esempio classico è quello della moltiplicazione e della scomposizione in fattori. Se vengono dati due numeri primi, è semplice moltiplicarli, ottenendo come prodotto un numero più grande di essi. Ma sembra che cercare di eseguire a rovescio questo procedimento - prendere il prodotto e ritrovare i due fattori sconosciuti - sia molto più difficile. Abbiamo algoritmi veloci per moltiplicare ma non per scomporre in fattori. Come mai? Anche se l'ufficio reclami rimane interdetto di fronte a queste domande, di recente c'è stato qualche progresso nel comprendere da dove deriva la difficoltà in almeno una famiglia di problemi computazionali, ossia quelli noti come problemi di «soddisfacimento di vincoli».

IL PROBLEMA DEL COMMESSO VIAGGIATORE è uno dei più celebri tra i problemi non calcolabili. Una sua versione piuttosto diffusa è la seguente: è possibile stimare il percorso più breve per un commesso viaggiatore che deve fare visita ai suoi clienti in tutte le città indicate sulla mappa? A prima vista, sembra facile, ma all'aumentare del numero dei nodi, ovvero delle città, il problema diventa esponenzialmente più difficile, mettendo nei guai anche i più potenti computer.

Questa nuova linea di ricerca non spiega perché alcuni di questi problemi siano difficili e altri facili, ma traccia molto dettagliatamente il confine tra i due casi. Inoltre, avere una mappa migliore del paesaggio dei procedimenti per risolvere problemi ha permesso di scoprire un nuovo algoritmo che fa indietreggiare un po' la frontiera di ciò che è intrattabile. L'algoritmo, chiamato «propagazione del sondaggio», potrebbe avere importanti applicazioni pratiche.

Dove sono i problemi difficili

Il nuovo algoritmo intreccia fili provenienti da almeno tre discipline: matematica, informatica e fisica. Il tema che lega il tutto è la presenza di transizioni improvvise da un tipo di comportamento a un altro.

Il filo matematico comincia negli anni sessanta con lo studio dei grafi aleatori, iniziato da Paul Erdós e Alfred Rényi. Un grafo è una struttura matematica astratta: un insieme di vertici e lati, disegnato in genere come uno schema di punti (i vertici) e linee che li uniscono (i lati, o spigoli). Per disegnare un grafo aleatorio si inizia distribuendo n vertici sul foglio e si considerano poi tutte le possibili coppie di vertici, scegliendo casualmente per ogni coppia, con probabilità p, se tracciare o no un lato che connetta i due punti. Quando p è vicino a 0, gli spigoli sono pochi e il grafo è composto di molti piccoli pezzi, o componenti connesse, separati gli uni dagli altri. Via via che p cresce, il grafo comincia a essere dominato da una singola componente connessa «gigante», che comprende la maggior parte dei vertici. L'esistenza di questa componente gigante non è certo una sorpresa, ma il modo in cui essa si sviluppa non è ovvio. La componente non evolve gradualmente al crescere di p, bensì emerge all'improvviso quando viene superata una certa soglia. La soglia è definita in termini di un parametro che chiamerò α: il numero di lati diviso per il numero di vertici. La componente gigante nasce quando α è circa 1/2.

In campo informatico un simile fenomeno di soglia attirò molta attenzione nei primi anni novanta. In questo caso la soglia determina la probabilità che certi problemi computazionali abbiano soluzione. Uno di questi problemi deriva proprio dalla teoria dei grafi: è il problema della k-colorazione, che chiede di dipingere ogni vertice di un grafo con uno di k colori, con la regola che due vertici uniti da un lato non possono avere lo stesso colore. Trovare una colorazione corretta diventa sempre più difficile al crescere di α perché ci sono più lati che impongono vincoli su ogni vertice. Di nuovo, la soglia è netta: al di sotto di un certo valore del rapporto α quasi tutti i grafi sono k-colorabili, mentre al di sopra di questa soglia non lo è quasi nessuno. Inoltre la soglia non solo influisce sull'esistenza di soluzioni ma anche sulla difficoltà di trovarne. Lo sforzo computazionale necessario per decidere se un grafo è k-colorabile ha un picco significativo vicino al valore critico di α. (Un importante articolo su questo effetto fu appropriatamente intitolato *Where the Really Hard Problems Are*, ossia: «Dove sono i problemi *veramente* difficili».)

Anche i fisici sanno qualcosa dei problemi di soglia; li chiamano transizioni di fase. Ma i cambiamenti di stato osservati nei grafi aleatori e nei problemi di soddisfacimento di vincoli sono veramente analoghi ad eventi fisici come il congelamento dell'acqua e la comparsa della magnetizzazione nel ferro? O la somiglianza è una semplice coincidenza? Per qualche tempo l'argomento fu controverso, ma ora è chiaro che i fenomeni di soglia nei grafi e in altre strutture matematiche sono autentiche transizioni di fase. Gli strumenti e le tecniche della fisica statistica sono adattissimi a studiarli. In particolare, il problema della k-colorazione ha una corrispondenza esatta con un modello di un sistema magnetico nella fisica dello stato solido. L'algoritmo di propagazione del sondaggio si basa su idee sviluppate in origine per descrivere simili modelli fisici.

Dove non sono i problemi difficili

La propagazione del sondaggio è in realtà una famiglia di algoritmi, che possono essere applicati in molti diversi ambiti. Finora il metodo è stato sperimentato su due problemi specifici. Il primo di questi è la soddisfacibilità booleana o SAT, in cui lo scopo è risolvere una formula complessa della logica simbolica, assegnando valori di vero o falso a tutte le variabili in modo tale che l'intera formula assuma un valore di vero. Il secondo problema è la k-colorazione. In questo articolo adotterò come esempio principale la k-colorazione, concentrandomi sulla 3-colorazione, in cui abbiamo a disposizione solo tre colori.

La 3-colorazione è un problema complesso, ma non impossibile. La domanda «Questo grafo è 3-colorabile?» ha sempre risposta, almeno in linea di principio. Visto che a ogni vertice può essere assegnato uno qualunque dei tre colori e che ci sono n vertici, ci devono essere esattamente 3^n modi di colorare il grafo. Per decidere se uno specifico grafo è 3-colorabile, basta prendere in esame, una per una, tutte le possibilità. Se si trova un'assegnazione di colori che soddisfa il vincolo, cioè in cui nessun lato congiunge vertici dello stesso colore, allora la risposta alla domanda è sì. Se si esauriscono tutte le possibilità senza trovare una colorazione appropriata, si può essere certi che non esiste.

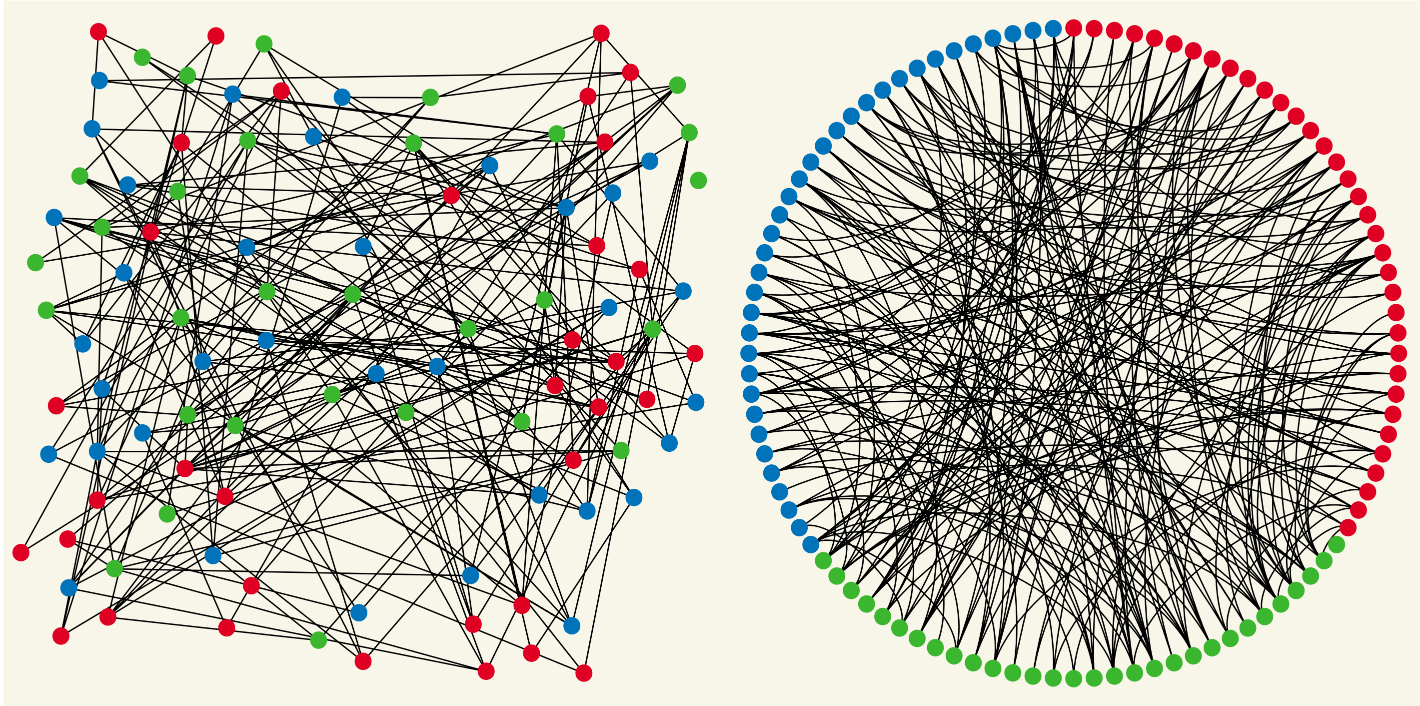
Questo algoritmo è semplice e sicuro. Ma anche inutile, perché enumerare 3^n colorazioni è al di là di ciò che si può fare in pratica per qualsiasi n maggiore di 15 o 20. Procedure più sofisticate possono continuare a garantire che la ricerca sia esatta ed esaustiva pur riducendo il numero di operazioni a meno di $1{,}5^n$. È un miglioramento significativo, ma si tratta sempre di una funzione esponenziale e non fa altro che innalzare il limite a $n = 50$ o giù di lì. Per grafi grandi, con migliaia di vertici, tutti questi metodi basati sulla forza bruta non offrono speranze.

D'altro canto, se si potesse in qualche modo sbirciare la soluzione di un problema di 3-colorazione su molti vertici, se ne potrebbe controllare la correttezza facendo assai meno fatica. Tutto quello che si dovrebbe fare sarebbe scorrere l'elenco dei lati e verificare che i vertici alle estremità di ciascun lato abbiano colori diversi. Il numero di lati in un grafo non può essere maggiore di n^2, che è una funzione polinomiale anziché esponenziale e quindi cresce molto più lentamente.

I problemi con risposte che sono difficili da trovare ma facili da verificare sono caratteristici della classe chiamata NP (che sta per *nondeterministic polynomial*: polinomiale non deterministico). La 3-colorazione è membro della prim'ora della classe NP e appartiene anche al più elitario gruppo di problemi descritti come NP-completi; la stessa cosa è vera per la soddisfacibilità. A meno di un miracolo, non si avranno mai algoritmi in tempo polinomiale per i problemi NP-completi.

Avendo ora appurato le credenziali della 3-colorazione come

Alcune cose, in questo mondo, richiedono un tempo di calcolo semplicemente troppo lungo



IN SINTESI

- Lo studio dei problemi computazionali di soddisfacimento di vincoli, intrinsecamente difficili, ha consentito la messa a punto di una nuova classe di algoritmi, detti di «propagazione del sondaggio», che in taluni casi facilitano l'ottenimento di soluzioni.
- Questi algoritmi si basano sull'analogia fra l'esistenza di valori di soglia oltre i quali la difficoltà dei problemi aumenta enormemente e le transizioni di fase nei sistemi fisici.
- Pur con le loro limitazioni, gli algoritmi di propagazione del sondaggio hanno applicazioni nella gestione di orari, nell'ingegneria dei circuiti e nell'ottimizzazione di programmi.

LA COLORAZIONE DEI GRAFI È UNO DEI PROBLEMI COMPUTAZIONALI che hanno una soglia piuttosto netta oltre la quale diventano molto più difficili da risolvere. Un grafo è un insieme di vertici e spigoli rappresentati con punti e linee; il problema della colorazione consiste nell'assegnare a ogni vertice un colore - in questo caso scelto fra rosso, verde e blu - in modo tale che nessuno spigolo connetta due vertici dello stesso colore. Quando il rapporto tra il numero di spigoli e quello di vertici è minore di circa 2,35, quasi tutti i grafi ottenuti aleatoriamente possono essere colorati; al di sopra di questa soglia, sono pochissimi i grafi che ammettono una soluzione. Il grafo mostrato nella figura ha 100 vertici e 218 spigoli ed è quindi poco al di sotto della soglia. Un nuovo algoritmo, chiamato «propagazione del sondaggio», ha permesso di colorare grafi aventi fino a un milione di vertici. I due diagrammi qui sopra sono raffigurazioni dello stesso grafo; la disposizione circolare dei vertici, a destra, rende più facile verificare che nessuno spigolo connette vertici dello stesso colore.

problema ufficialmente difficile, possiamo rivelare che la maggior parte dei problemi di 3-colorazione su grafi aleatori sono in realtà piuttosto semplici. Dato un grafo tipico, si hanno buone probabilità di trovare rapidamente una 3-colorazione o di dimostrare che non esiste. Questa curiosa situazione non è veramente paradossale. La classificazione della 3-colorazione come problema NP-completo si basa sull'analisi del peggiore dei casi. Verrebbe smentita solo da un algoritmo che offrisse la garanzia di dare la risposta corretta e di concludere l'elaborazione in un tempo polinomiale per ogni possibile grafo. Nessuno ha scoperto un tale algoritmo. Ma ci sono molti algoritmi che hanno, nella maggior parte dei casi, un tempo di elaborazione rapido, a patto di accettare un occasionale fallimento.

Una strategia diffusa per algoritmi che colorano grafi è il *backtracking* (letteralmente: «tornare sui propri passi»). Assomiglia al modo in cui la maggior parte delle persone affronterebbe il problema se dovesse cercare di colorare il grafo a mano. Si inizia assegnando un colore arbitrario a un vertice arbitrario; poi si passa ai vertici vicini, assegnando loro colori che non causino un conflitto. Proseguendo così si può arrivare a un vertice per cui non c'è un colore lecito; a questo punto si torna sui propri passi, annullando alcune scelte precedenti, e si riprova.

Per mostrare che un grafo non può essere 3-colorato occorre un altro tipo di algoritmo. L'approccio fondamentale consiste nel cercare un sottoinsieme di vertici che, anche se fosse isolato dal resto del grafo, non potrebbe essere 3-colorato. Per esempio, una *clique* (cricca) costituita da quattro vertici ognuno dei quali sia collegato con ognuno degli altri ha questa proprietà. Se si trova anche solo una di queste strutture, la questione è risolta per l'intero grafo.

Algoritmi come questi sono molto diversi dai metodi basati sulla forza bruta e sulla ricerca esaustiva. La semplice enumerazione di tutte le 3^n colorazioni può essere inammissibilmente lenta, ma almeno è prevedibile; il tempo di elaborazione è lo stesso per tutti i grafi con lo stesso numero di vertici. Ciò non è vero per il backtracking e per altri algoritmi inesatti o incompleti; le loro prestazioni variano notevolmente a seconda della natura del grafo. In particolare, questi algoritmi sono sensibili al valore di α - il rapporto tra il numero di lati e il numero di vertici - che è di nuovo il parametro che controlla la transizione tra fasi colorabili e non colorabili. Molto al di sotto del valore critico di α, dove gli spigoli sono radi, c'è un numero tale di modi di colorare il grafo che qualsiasi strategia ragionevole ha buone probabilità di trovarne uno. All'estremo opposto, molto al di sopra della soglia, i grafi sono densamente interconnessi, ed è facile trovare un sottografo che renda impossibile la 3-colorazione. La regione problematica è situata tra questi estremi, vicino alla soglia. In questa zona intermedia possono esserci pochissime colorazioni o nessuna. Distinguere tra queste due situazioni può rendere necessario controllare quasi ogni possibile assegnazione di colori.

Dove sono le soluzioni

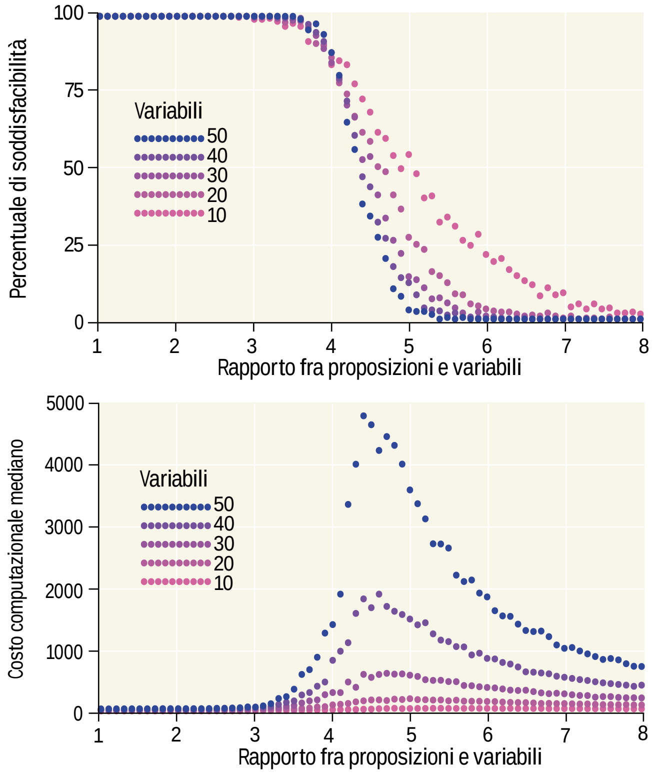
Il valore critico di α è circa 2,35. In altre parole, se un grafo aleatorio con *n* vertici ha meno di 2,35*n* lati, può essere quasi sicuramente 3-colorato; se ne ha di più, una 3-colorazione è improbabile. Inoltre si sa che la transizione tra questi due regimi è netta; è una vera discontinuità, un salto improvviso anziché un passaggio graduale. Per esprimere più formalmente quest'idea, l'ampiezza della regione di transizione tende a zero quando *n* tende all'infinito.

La nettezza della transizione di fase può essere considerata una notizia incoraggiante. Se gli algoritmi per decidere la colorabilità si impantanano solo nella regione di transizione, e se essa è tanto ristretta da essere quasi trascurabile, allora la probabilità di incontrare un grafo difficile da classificare è proporzionalmente piccola. Ma sembra che l'universo abbia un altro baco (o caratteristica voluta). In primo luogo, la nettezza della transizione è garantita solo per grafi infinitamente grandi; se *n* è finito, gli angoli della curva di transizione sono arrotondati. E c'è un altro fattore di disturbo, che è stato riconosciuto solo di recente: non ha a che fare con la struttura del grafo in sé, ma con la struttura dell'insieme di tutte le soluzioni al problema della colorabilità.

Sebbene la fase non colorabile non inizi fino ad α ≈ 2,35, gli esperimenti hanno mostrato che gli algoritmi cominciano a rallentare un po' prima, a valori di α attorno a 2,2. La discrepanza può sembrare poco importante, ma è troppo grande per poter essere spiegata solo dalla «sfocatura» della transizione di fase per valori finiti di *n*. Sta accadendo qualcos'altro.

Per capire la causa, giova visualizzare tutte le possibili 3-colorazioni di un grafo distese su una superficie. L'altezza della superficie in ogni punto rappresenta il numero di conflitti nella colorazione corrispondente. Così le colorazioni perfette (quelle senza conflitti) si trovano tutte al livello del mare, mentre le colorazioni peggiori creano picchi o altipiani ad alta quota. Naturalmente la topografia di questo paesaggio dipende dal particolare grafo che stiamo esaminando. Si consideri come evolve la superficie via via che α cresce gradualmente. Per bassi valori di α ci sono ampi bacini e vallate, che rappresentano i molti modi di colorare perfettamente il grafo. Per alti valori di α il paesaggio è alpino, e anche il punto più basso è molto al di sopra del livello del mare, denotando una completa assenza di colorazioni perfette. Il valore di transizione α ≈ 2,35 segna il momento in cui scompaiono le ultime aree estese che si trovano al livello del mare.

Che cosa avviene in questo «spazio delle soluzioni» per α ≈ 2,2? Si è scoperto che questo è il momento in cui un'ampia distesa di terreno a livello zero si frammenta in piccoli bacini isolati. Al di sotto di 2,2, quasi tutte le colorazioni perfette formano un'unica gigantesca regione connessa. Esse sono connesse nel senso che si può trasformare una soluzione in un'altra facendo relativamente pochi cambiamenti, e senza introdurre troppi conflitti in alcuno degli stati intermedi. Al di sopra di 2,2, ogni bacino rappresenta un insieme isolato di soluzioni. Le colorazioni che si trovano in bacini separati sono sostanzialmente diverse e per trasformarne una in un'altra si dovrebbe scalare una catena montuosa formata da colorazioni che presentano un alto numero di conflitti. È improbabile che gli algoritmi che funzionano conducendo una ricerca locale riescano a valicare queste catene montuose, e quindi rimangono confinati a lungo nel primo bacino in cui capitano. Al crescere di α sopra 2,2, il numero di colorazioni perfette all'interno di ogni bacino decresce fino a zero e quindi gli algoritmi possono non riuscire a trovare una soluzione anche se esistono ancora molte colorazioni valide in altre parti della superficie delle soluzioni.

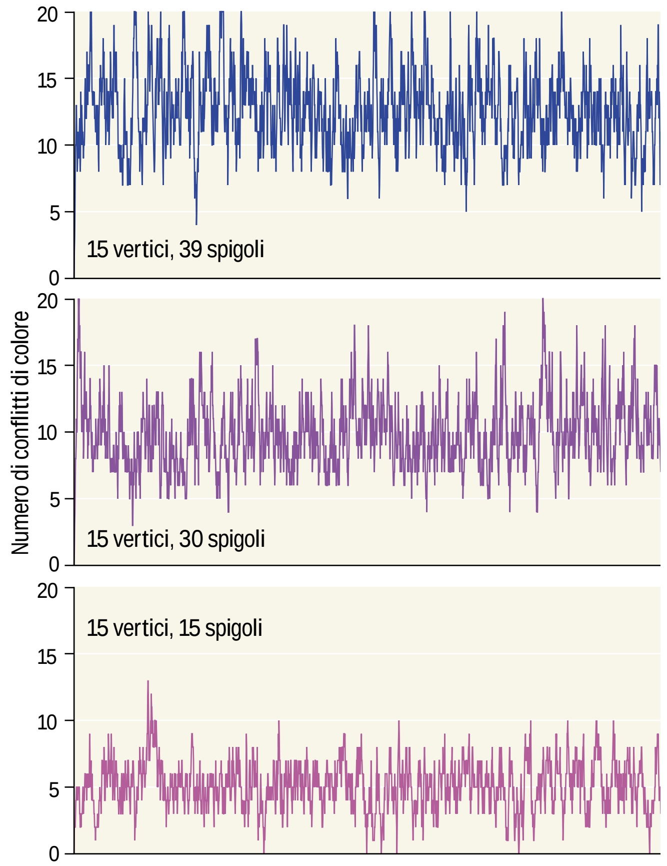


LA TRANSIZIONE TRA FASE RISOLUBILE E FASE NON RISOLUBILE per il problema noto come soddisfacibilità (*in alto*), in cui il compito è di assegnare i valori vero o falso alle variabili di una formula logica fatta di molte proposizioni. All'aumentare del rapporto tra il numero di proposizioni e quello di variabili, il problema passa dall'essere quasi sempre risolubile al non esserlo quasi mai. La transizione diventa più netta per problemi più grandi. Lo sforzo computazionale richiesto per decidere se una formula è soddisfacibile o no (*in basso*) ha un picco vicino alla transizione di fase. I problemi molto al di sotto della soglia sono facili da risolvere, e per quelli molto al di sopra è facile dimostrare che non vi è soluzione; i più complessi da analizzare sono quelli intermedi.

Questa visione delle soluzioni sparse su un paesaggio ondulato è uno strumento concettuale familiare in molte aree della fisica. Spesso il paesaggio è descritto come una superficie di energia, e si assume che i sistemi fisici tendano verso stati di energia minima. Questa analogia può essere spinta oltre, stabilendo una precisa corrispondenza tra le *k*-colorazioni dei grafi e un modello dei materiali magnetici.

Dove sono gli spin?

Esiste una tale varietà di modelli del magnetismo da lasciare a bocca aperta. Le componenti di base sono vettori che rappresentano spin atomici. Solitamente gli spin sono disposti in un reticolo regolare, come in un solido cristallino, e i vettori sono vincolati a puntare solo in poche direzioni possibili. In un modello di materiale ferromagnetico, gli spin vicini hanno accoppiamenti positivi, il che significa che l'energia del sistema è minore quando gli spin sono allineati parallelamente. Un materiale antiferromagnetico ha accoppiamenti negativi, favorendo così spin che puntano in direzioni differenti. Il problema della 3-colorazione di un grafo può essere visto come un modello di un materiale antiferromagnetico in cui ogni spin ha tre possibili direzioni, corrispondenti ai tre colori. È *anti*ferromagnetico perché lo stato preferito è uno in cui i colori o gli spin differiscono.



COMPIERE PASSEGGIATE ALEATORIE ATTRAVERSO LO SPAZIO delle colorazioni di grafi fa capire perché sia difficile trovare una colorazione perfetta. Ogni passeggiata comincia con una colorazione perfetta e modifica via via il colore di un vertice scelto a caso. I grafici registrano il numero di conflitti di colore (spigoli che connettono vertici dello stesso colore) nel corso di 2000 passi. Sebbene si sappia che tutti e tre i grafi sono colorabili, la passeggiata aleatoria scopre colorazioni perfette solo per il grafo più piccolo (*grafico inferiore*).

La maggior parte dei modelli di sistemi di spin si concentra sugli effetti delle fluttuazioni termiche e sugli imperativi contrapposti di minimizzare l'energia e massimizzare l'entropia. Da questo punto di vista il modello della colorazione di grafi è più semplice della maggior parte degli altri, poiché la condizione che interessa è a temperatura zero, quando l'entropia può essere trascurata. D'altro canto, il modello è più complicato sotto un altro aspetto: gli spin sono immersi in un grafo con interconnessioni casuali, più simile a un vetro che al reticolo geometricamente regolare di un cristallo. Avendo tradotto il problema della colorazione nel linguaggio della fisica degli spin, la meta è di identificare lo stato fondamentale, cioè la configurazione di spin di energia minima. Se l'energia dello stato fondamentale è zero, allora esiste almeno una colorazione perfetta. Se l'energia degli spin non può essere ridotta a zero, allora il grafo corrispondente non è 3-colorabile. L'energia minima indica quanti conflitti ineliminabili esistono nel grafo colorato.

Naturalmente riformulare il problema in un nuovo lessico non fa sparire la difficoltà di fondo. Nella colorazione dei grafi, quando si risolve un conflitto cambiando il colore di un vertice, si può creare un nuovo conflitto in un altro punto del grafo. Analogamente, nel sistema di spin, quando si abbassa l'energia di una coppia di spin appaiati, la si può innalzare per un'altra coppia. I fisici chiamano questo effetto «frustrazione».

Le interazioni tra spin adiacenti possono essere viste come una sorta di trasmissione di messaggi in cui ogni spin dice ai suoi vicini che cosa devono fare (o, visto che l'accoppiamento è antiferromagnetico, che cosa non devono fare). Ritraducendo ciò nel linguaggio della colorazione di grafi, un vertice verde dirama un segnale ai suoi vicini dicendo «Non siate verdi». I vicini rimandano i propri messaggi: «Non siate rossi», «Non siate blu». Il problema è che ogni lato trasporta messaggi in entrambe le direzioni, e alcuni di essi possono essere contraddittori. E circuiti chiusi di retroazione possono impedire che il sistema trovi uno stato stazionario in cui arrestarsi.

Un rimedio per questo tipo di frustrazione è noto nella fisica della materia condensata come metodo della cavità. Esso prescrive la seguente successione di azioni: prima si sceglie un singolo spin e lo si rimuove temporaneamente dal sistema (creando così una «cavità»). Poi, tra i vicini che circondano la cavità si sceglie un nodo da considerare come output e si considerano gli altri come input. Si sommano i segnali che arrivano lungo i lati di input e si passa il risultato all'output. L'effetto è di spezzare i circuiti chiusi e forzare comunicazioni unidirezionali. Infine, si ripete l'intera procedura con un altro spin e si continua finché il sistema converge a qualche stato stazionario.

Il metodo della cavità fu applicato per la prima volta a problemi di soddisfacimento di vincoli da Marc Mézard dell'Università di Parigi Sud, Giorgio Parisi dell'Università di Roma «La Sapienza» e Riccardo Zecchina dell'Abdus Salam International Centre for Theoretical Physics di Trieste. All'inizio era uno strumento per calcolare le proprietà medie di insiemi statistici di sistemi a molti spin. Circa un anno fa, Mézard e Zecchina si sono resi conto che può anche essere adattato per affrontare singoli problemi. Ma serviva una modifica essenziale. Invece di messaggi semplici come «Non siate verdi», l'informazione trasmessa da nodo a nodo consiste in intere distribuzioni di probabilità, che danno una valutazione numerica a ogni possibile stato dello spin o colore del vertice.

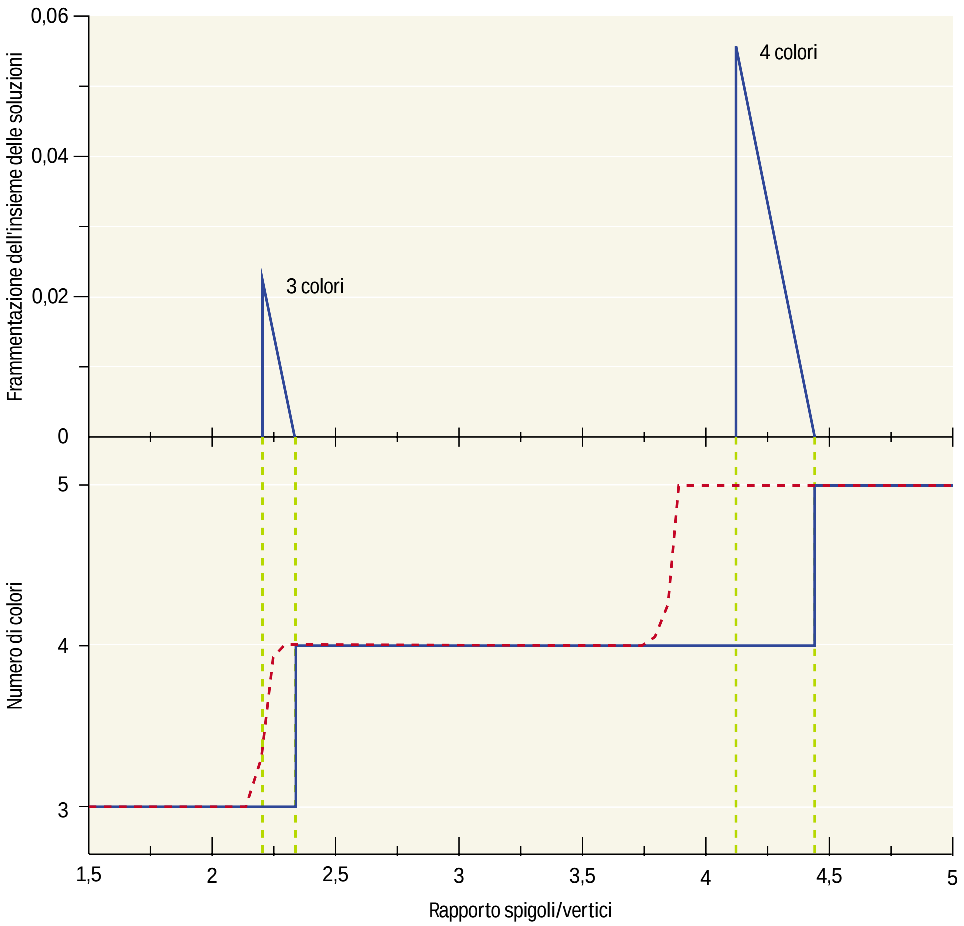
Mézard e Zecchina hanno chiamato l'algoritmo «propagazione del sondaggio». Hanno tratto il termine di «propagazione» da un altro algoritmo che ha pure ispirato il loro lavoro: una tecnica chiamata «propagazione delle opinioni», che viene usata in certi codici di correzione di errori. «Sondaggio» è pensato nel senso dei sondaggi d'opinione: si sondano i siti che circondano una cavità per apprendere quali consigli hanno da offrire ai loro vicini.

Dove sono i bachi

Nel corso dell'ultimo anno il concetto di propagazione del sondaggio è stato ulteriormente raffinato e incorporato in una serie di programmi per computer da Mézard e Zecchina con un gruppo di collaboratori, tra i quali Alfredo Braunstein, Silvio Franz, Michele Leone, Andrea Montanari, Roberto Mulet, Andrea Pagnani, Federico Ricci-Tersenghi e Martin Weigt.

Per risolvere un problema di 3-colorazione su un grafo con *n* vertici, l'algoritmo dapprima trova il vertice che è più fortemente vincolato ad avere un certo colore e fissa di conseguenza il colore di quel vertice. L'algoritmo viene poi invocato ricorsivamente per il rimanente grafo su $n - 1$ vertici, in modo da fissare il colore di un altro vertice. Ovviamente questo procedimento deve terminare dopo non più di *n* ripetizioni. In pratica, spesso si arresta prima, quando tutti i segnali che si propagano attraverso il grafo sono diventati messaggi di indifferenza, che non pongono vincoli ai nodi vicini. A questo punto la propagazione del sondaggio non ha più nulla da offrire, ma il grafo rimanente è stato ridotto a un caso risolubile banalmente con altri metodi.

IL PROBLEMA DELLA COLORAZIONE DEI GRAFI si fa più duro via via che essi sono più densamente connessi, ma l'aumentare della difficoltà non è regolare. Presenta, anzi, salti apparentemente improvvisi. Il numero minimo di colori necessario (*curva blu nel grafico in basso*) sale da 3 a 4 quando il rapporto tra spigoli e vertici raggiunge all'incirca il valore 2,35. Gli algoritmi che cercano una soluzione a tre colori cominciano ad avere problemi già per valori un po' più bassi (*curva in rosso*) della soglia. La ragione del prematuro fallimento degli algoritmi è che, anche se esistono soluzioni, diventano improvvisamente più difficili da trovare. Al di fuori di questa regione problematica, quasi tutte le soluzioni sono strettamente legate fra loro: un piccolo cambiamento a una colorazione corretta (come scambiare i colori di due vertici) porta facilmente a un'altra soluzione. Nella regione di transizione, invece, l'insieme delle soluzioni si spezza in «isole» separate. Il grafico in alto mostra che questa frammentazione dell'insieme delle soluzioni insorge in modo improvviso. E lo stesso fenomeno è ancora più pronunciato nel passaggio da 4 a 5 colori.



L'AUTORE

BRIAN HAYES è *senior writer* di «American Scientist». Il suo indirizzo e-mail è: bhayes@amsci.org

«Questo articolo ha avuto la sua genesi durante una permanenza di 10 settimane presso l'Abdus Salam International Centre for Theoretical Physics di Trieste, dove ho avuto utili discussioni con Riccardo Zecchina, Muli Safra, Roberto Mulet, Marc Mézard, Stephan Mertens, Alfredo Braunstein, Johannes Berg e altri.»

BIBLIOGRAFIA

MÉZARD MARC, PARISI GIORGIO e VIRASORO MIGUEL ANGEL (a cura), *Spin Glass Theory and Beyond*, World Scientific, Philadelphia, 1987.
CHEESEMAN PETER, KANEFSKY BOB e TAYLOR WILLIAM M., *Where the* Really *Hard Problems Are*, in *Proceedings of the International Joint Conference on Artificial Intelligence*, vol. 1, 1991.
HAYES BRIAN, *Computing Science: Can't Get no Satisfaction*, in «American Scientist», 85, pp. 108-112, 1997.
MÉZARD MARC, PARISI GIORGIO e ZECCHINA RICCARDO, *Analytic and Algorithmic Solution of Random Satisfiability Problem*, in «Science», 297, pp. 812-815, 2002.
MÉZARD MARC e ZECCHINA RICCARDO, *The Random K- Satisfiability Problem: From an Analytic Solution to an Efficient Algorithm*, in «Physical Review E», 66, 2002.
MÉZARD MARC e PARISI GIORGIO, *The Cavity Method at Zero Temperature*, in «Journal of Statistical Physics» (in stampa).

Come accade per altri algoritmi per problemi NP-completi, la propagazione del sondaggio non è coperta da garanzia, e talvolta fallisce. Il procedimento per decidere quale sarà il successivo vertice da fissare non è infallibile e, quando si compie una scelta sbagliata, potrebbe non esserci in seguito un'opportunità di rimediare. (L'aggiunta di qualche forma di backtracking o di ripartenza casuale può alleviare questo problema.) Nella sua forma attuale, l'algoritmo è anche strettamente unilaterale: solitamente è in grado di colorare un grafo colorabile, ma non può dimostrare che un grafo è non colorabile. Ciononostante, l'algoritmo ha già conseguito successi rilevanti, in particolare nella regione di difficile soluzione vicino alla transizione di fase. La versione per la soddisfacibilità ha risolto problemi con 8 milioni di variabili. Il programma per la colorazione di grafi può trattare grafi di un milione di vertici. Entrambi questi numeri superano di due ordini di grandezza ciò che è di routine per altri metodi.

La colorazione di grafi e la soddisfacibilità non sono solo problemi giocattolo per teorici. Sono al centro di vari problemi pratici nella gestione di orari, nell'ingegneria dei circuiti e nell'ottimizzazione di programmi. Disporre di un algoritmo in grado di risolvere casi molto più grandi potrebbe aprire la porta a ulteriori applicazioni.

Paradossalmente, mentre la propagazione del sondaggio funziona bene con problemi enormi, talora si inceppa in situazioni molto più piccole, come grafi aleatori con poche centinaia di vertici. Ciò non costituisce un'impellente preoccupazione pratica, dato che altri metodi funzionano bene in questo intervallo, ma è fastidioso, e fa temere che gli stessi problemi possano emergere per grafi più grandi non aleatori. La causa di questi fallimenti non è ancora chiara. Può essere dovuta a un'abbondanza di circuiti chiusi densamente nidificati e altre strutture entro i grafi. Oppure può essere semplicemente un altro baco dell'universo.