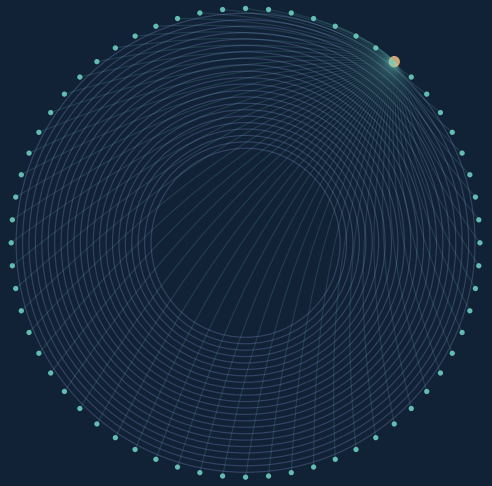
CALCOLO QUANTISTICO

L’algoritmo di Grover

Ricerca e interferenza
Amplificazione di ampiezza



Una lettura con esempi numerici e richiami alle definizioni

Il percorso di lettura

I capitoli seguono l'ordine dei prerequisiti. I richiami a fondo pagina riportano le definizioni che servono al ragionamento corrente.

01 Il problema della ricerca
Candidati, verificatori e interrogazioni

02 Stati e probabilità
La preparazione dei registri e le ampiezze

03 Oracoli reversibili
Il valore della funzione e la marcatura di fase

04 Quattro candidati
Un esempio completo dell'iterazione

05 Otto candidati
Il calcolo delle ampiezze dopo ogni passaggio

06 Riflessioni e sottospazi
Proiezioni, ortogonalità e conservazione della norma

Il percorso di lettura

I capitoli seguono l'ordine dei prerequisiti. I richiami a fondo pagina riportano le definizioni che servono al ragionamento corrente.

07 La rotazione di Grover
Il piano degli stati e la probabilità di successo

08 Circuiti e ausiliari
La costruzione dell'oracolo e del diffusore

09 Più soluzioni
La probabilità complessiva e i singoli esiti

10 Il numero di soluzioni ignoto
Conteggi casuali, verifiche e nuovi tentativi

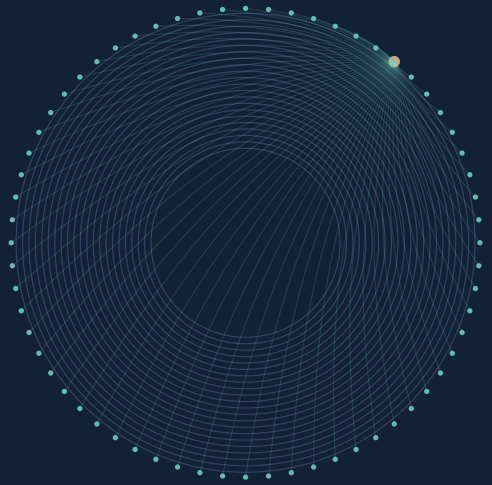
11 Amplificazione di ampiezza
La ricerca con una preparazione non uniforme

12 Costo e uso dell'algoritmo
Risorse, limiti ed esercizi svolti

CAPITOLO 01

Il problema della ricerca

Candidati, verificatori e interrogazioni



L’algoritmo di Grover

La ricerca di un candidato valido

Consideriamo un insieme finito di candidati e una procedura che, ricevuto un candidato, risponde se soddisfa una certa condizione. Il problema consiste nel trovare almeno un candidato valido. La procedura di verifica può essere disponibile anche quando nessuno ci ha ancora fornito una soluzione. È la situazione ordinaria di molti problemi di ricerca: possiamo controllare una proposta, ma dobbiamo ancora stabilire quale proposta sottoporre al controllo.

Useremo inizialmente otto candidati, numerati da 0 a 7, e una sola soluzione. Per seguire i calcoli mostreremo quale sia la soluzione degli esempi. L'algoritmo, invece, dovrà individuarla attraverso le operazioni ammesse sul verificatore. Questa distinzione fra ciò che conosce il lettore e ciò che riceve l'algoritmo sarà importante anche quando costruiremo piccoli circuiti dimostrativi.

Richiami Un candidato è un possibile risultato; una soluzione è un candidato che supera la verifica.

Periodicità e ricerca senza struttura

Negli algoritmi di Simon e Shor, la funzione possiede una regolarità promessa dal problema. Simon cerca uno spostamento binario che lascia invariati i valori della funzione; Shor ricava un periodo delle potenze modulari. Queste relazioni collegano fra loro ingressi diversi e rendono possibile una ricostruzione della struttura nascosta.

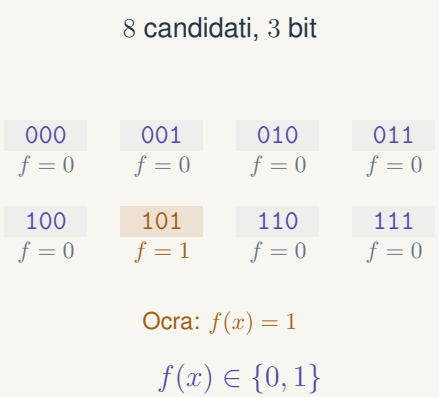
La ricerca di Grover parte da un'ipotesi meno ricca: sappiamo riconoscere i candidati validi, ma non assumiamo una periodicità, un ordinamento utile o una relazione fra candidati vicini. Il termine «senza struttura» riguarda le informazioni offerte all'algoritmo, non l'aspetto della tabella con cui rappresentiamo gli ingressi. Se un problema concreto possiede altre proprietà utilizzabili, anche gli algoritmi classici possono sfruttarle. Il confronto corretto deve allora tenerne conto.

Richiami Il verificatore stabilisce se un singolo candidato soddisfa la condizione richiesta.

Il predicato che riconosce le soluzioni

Rappresentiamo il verificatore mediante una funzione f che restituisce un bit. Scriviamo $f(x) = 1$ quando il candidato x è valido e $f(x) = 0$ quando non lo è. Una funzione con questo tipo di risultato si chiama predicato: esprime una proprietà attraverso i due valori vero e falso.

Nell'esempio illustrato il solo candidato valido è 5. Il disegno mostra l'intera funzione per permetterci di controllare i passaggi successivi. Un algoritmo di ricerca non riceve gratuitamente questa tabella completa: può interrogare la procedura che calcola f. Se ricevesse già tutti i valori esplicitamente, il costo necessario per leggerli diventerebbe parte del problema.



Le etichette identificano gli otto candidati. Il colore distingue il solo valore per cui il verificatore restituisce 1; la disposizione sul foglio non fornisce informazioni alla ricerca.

Richiami Una soluzione soddisfa $f(x) = 1$; il valore 0 indica che il candidato non supera il controllo.

Le stringhe che identificano i candidati

Per identificare otto candidati bastano tre bit. In una stringa scritta da sinistra a destra, i bit hanno peso 4, 2 e 1. La stringa 101 rappresenta quindi $1 \cdot 4 + 0 \cdot 2 + 1 \cdot 1 = 5$. Useremo indifferentemente il numero 5 e la sua scrittura binaria, precisando quando la lunghezza del registro è rilevante.

Con n bit si possono formare $N = 2^n$ stringhe distinte, perché ogni posizione ammette due scelte. Due bit danno quattro candidati, tre bit ne danno otto e dieci bit ne danno 1024. Il numero di bit e il numero di candidati descrivono dimensioni diverse dello stesso problema: una crescita modesta del primo può produrre una crescita notevole del secondo.

La stringa 101

x_2	x_1	x_0
1	0	1
$2^2 = 4$	$2^1 = 2$	$2^0 = 1$

$$4 + 0 + 1 = 5$$

Il bit più significativo è a sinistra.

$$N = 2^n$$

Il bit a sinistra è il più significativo: pesa quattro unità. Negli esempi su tre bit manterremo sempre l'ordine $x_2 x_1 x_0$.

Richiami N indica il numero di candidati; n indica il numero di bit necessario a identificarli.

Il numero dei candidati validi

Indichiamo con M il numero di candidati per i quali il predicato vale 1. La richiesta di trovare una soluzione non impone, in generale, che ne esista una sola. Se otto candidati contengono due soluzioni, un risultato che coincide con una qualsiasi delle due soddisfa il problema.

Distingueremo tre informazioni: conoscere il numero totale N, conoscere eventualmente il numero M delle soluzioni e conoscere la loro posizione. Sapere che esiste una sola soluzione non significa sapere quale sia. Allo stesso modo, sapere che le soluzioni sono due può aiutare a scegliere la durata dell'algoritmo senza rivelare alcuna delle due. Per iniziare studieremo il caso $M = 1$, poi estenderemo il ragionamento a più soluzioni e a un numero di soluzioni ignoto.

$$0 \leq M \leq N$$

Richiami M conta i candidati con $f(x) = 1$; N conta tutti i candidati, validi e non validi.

Le informazioni disponibili all'algoritmo

Il modello della ricerca specifica quali operazioni può eseguire l'algoritmo. Nel caso classico più semplice, esso sceglie un candidato x, chiama il verificatore e riceve il bit $f(x)$. La chiamata prende il nome di interrogazione. Il numero delle interrogazioni misura quante volte l'algoritmo usa la procedura che riconosce le soluzioni.

Una risposta negativa esclude il candidato appena provato, ma nel problema senza struttura non segnala quale altro candidato sia preferibile. Per esempio, sapere che $f(010) = 0$ non rende 011 più promettente di 110. L'ordine numerico delle etichette serve a identificarle e non costituisce un ordinamento per validità. Un vantaggio nella ricerca deve dunque derivare da qualcosa di diverso da una scelta classica sempre migliore del candidato successivo.

Richiami Un'interrogazione classica fornisce il valore $f(x)$ per il candidato scelto.

Una verifica su otto possibilità

Supponiamo che la sola soluzione possa trovarsi con uguale probabilità in ciascuna delle otto posizioni. Se controlliamo un candidato fissato in anticipo, la probabilità che sia quello valido è $1/8$. Lo stesso vale se scegliamo il candidato uniformemente a caso: nessuna posizione ha un privilegio.

Questo valore descrive l'incertezza sulla soluzione o la casualità della scelta, non una proprietà probabilistica del verificatore. Una volta fissati f e x, il risultato del controllo è determinato. La probabilità diventa utile quando confrontiamo algoritmi che devono operare prima di conoscere la posizione della soluzione. Nel modello senza struttura possiamo anche considerare il caso peggiore; per i primi esempi, la distribuzione uniforme rende immediato il conteggio.

$$P(\text{successo con una verifica}) = \frac{1}{8}$$

Richiami Una scelta uniforme assegna la stessa probabilità a ogni candidato.

Le verifiche senza ripetizioni

Dopo un controllo negativo conviene scegliere un candidato diverso: ripetere una verifica deterministica non aggiunge informazioni. Se controlliamo quattro posizioni distinte su otto, avremo successo esattamente quando la soluzione appartiene a quelle quattro posizioni. Con la distribuzione uniforme considerata, la probabilità è quindi $4/8 = 1/2$.

Più in generale, q verifiche distinte coprono q delle N posizioni. La probabilità di avere effettivamente interrogato la soluzione è q/N, finché $0 \le q \le N$. Qui contiamo un successo verificato. Con la promessa di unicità, dopo $N - 1$ risposte negative si può anche dedurre la soluzione rimasta senza interrogarla: quel caso richiede di distinguere l'inferenza finale dal controllo diretto.

$$P(\text{soluzione fra le posizioni provate}) = \frac{q}{N}$$

Richiami q è il numero di candidati distinti verificati; la formula assume una sola soluzione distribuita uniformemente.

La probabilità dopo una risposta negativa

Con otto candidati e una sola soluzione, un primo controllo negativo lascia sette posizioni possibili. Se scegliamo uniformemente fra queste, la probabilità condizionata di riuscire al secondo controllo è $1/7$. «Condizionata» significa che stiamo usando l'informazione appena ricevuta: il primo candidato è stato escluso.

La probabilità di fallire entrambi i controlli si calcola seguendo i due eventi: prima falliamo con probabilità $7/8$, poi con probabilità $6/7$. Il prodotto vale $6/8$, perciò la probabilità complessiva di successo entro due controlli è $1 - 6/8 = 1/4$. Otteniamo lo stesso risultato del conteggio di due posizioni distinte su otto. La moltiplicazione segue le probabilità aggiornate, non due tentativi indipendenti con probabilità immutata.

$$P(\text{due fallimenti}) = \frac{7}{8} \cdot \frac{6}{7} = \frac{6}{8}$$

$$P(\text{almeno un successo}) = 1 - \frac{6}{8} = \frac{1}{4}$$

Richiami La probabilità condizionata incorpora un'informazione già osservata; gli eventi successivi possono quindi avere probabilità diverse.

I tentativi indipendenti con ripetizioni

Un'altra procedura sceglie ogni volta uniformemente fra tutti i candidati, senza ricordare quelli già provati. In questo caso possiamo controllare due volte la stessa posizione. Per una sola soluzione su otto, ogni tentativo indipendente fallisce con probabilità $7/8$. Dopo due tentativi, la probabilità di fallire entrambi è $(7/8)^2 = 49/64$.

La probabilità di almeno un successo è dunque $15/64$, leggermente inferiore al valore $1/4 = 16/64$ ottenuto evitando le ripetizioni. Per M soluzioni su N candidati, la probabilità di successo di un singolo tentativo uniforme è M/N. L'indipendenza permette allora di moltiplicare q volte la probabilità di fallimento. Useremo questa formula anche per discutere esperimenti che ripartono ogni volta dallo stesso stato iniziale.

$$P(\text{successo entro } q \text{ tentativi}) = 1 - \left(1 - \frac{M}{N}\right)^q$$

Richiami Tentativi indipendenti ripetono la stessa distribuzione senza usare l'esito precedente per cambiare i candidati possibili.

Due soluzioni nel confronto classico

Consideriamo otto candidati, due dei quali validi. Un controllo uniforme riesce con probabilità $2/8 = 1/4$. Se evitiamo di ripetere un candidato e il primo controllo fallisce, rimangono sette candidati: due validi e cinque non validi. Il secondo controllo fallisce quindi con probabilità condizionata $5/7$.

La probabilità di due fallimenti consecutivi è $(6/8)(5/7) = 15/28$. Di conseguenza, la probabilità di almeno un successo entro due controlli è $13/28$, circa il $46{,}4\%$. Con tentativi indipendenti che ammettono ripetizioni avremmo invece $1 - (3/4)^2 = 7/16$, circa il $43{,}8\%$. Avere più soluzioni facilita già la ricerca classica; anche il confronto quantistico dovrà includere questo cambiamento.

$$P(\text{successo senza ripetizioni}) = 1 - \frac{6}{8}\cdot\frac{5}{7} = \frac{13}{28}$$

Richiami Con M soluzioni, la probabilità iniziale di successo di una scelta uniforme è M/N.

Il numero medio di controlli

Fissiamo un ordine in cui esaminare gli otto candidati senza ripetizioni e fermiamoci appena il verificatore restituisce 1. Se la soluzione è uniforme, può occupare ciascuna delle otto posizioni dell'ordine con la stessa probabilità. I possibili numeri di controlli sono allora 1, 2, fino a 8.

Il valore medio si ottiene sommando questi otto numeri e dividendo per otto: servono $36/8 = 4{,}5$ controlli. Un singolo esperimento non esegue quattro controlli e mezzo; la frazione descrive la media di molte istanze o di molti ordinamenti casuali. Per N candidati, la stessa regola dà $(N+1)/2$, se richiediamo il controllo positivo anche dell'ultima posizione. Il caso peggiore rimane N verifiche.

$$\frac{1+2+\cdots+8}{8} = \frac{36}{8} = 4{,}5$$

Richiami Il costo medio pesa i possibili costi con le loro probabilità; non coincide necessariamente con un costo realizzato in una singola esecuzione.

Interrogazioni e lavoro del verificatore

Una singola interrogazione può nascondere molto lavoro. Un verificatore può confrontare pochi bit oppure eseguire un calcolo aritmetico complesso. Contare le interrogazioni permette di separare la strategia di ricerca dal contenuto della verifica, ma non fornisce ancora il tempo totale dell'esecuzione.

Per esempio, ridurre cento chiamate a dieci è utile solo dopo avere chiarito il costo delle chiamate nelle due realizzazioni confrontate. Nel circuito quantistico occorrerà rendere la verifica coerente e reversibile, con porte e possibili registri ausiliari. Il modello delle interrogazioni tratta queste operazioni come accessi a una stessa funzione astratta. Più avanti distingueremo questa misura teorica dal numero di porte e dalle risorse necessarie a realizzare realmente il circuito.

Richiami Il costo in interrogazioni conta gli usi del verificatore, non tutte le operazioni interne necessarie a calcolarlo.

Il verificatore e l’accesso ai dati

La parola «ricerca» può suggerire una tabella già memorizzata in un computer. Il predicato può però essere anche un calcolo che stabilisce se una stringa soddisfa un insieme di vincoli. In questo secondo caso, i candidati sono descritti dalle loro etichette e la verifica calcola la proprietà richiesta senza leggere un archivio completo.

Se invece la verifica dipende da dati esterni, bisogna specificare come il circuito vi accede. La possibilità matematica di interrogare un oracolo quantistico non costruisce automaticamente una memoria capace di fornire quei dati in sovrapposizione. Per questa ragione useremo piccoli predicati e circuiti espliciti: rendono controllabile l’azione del verificatore, mentre mantengono separata la questione del vantaggio su una particolare applicazione.

Richiami Il predicato f descrive la condizione di validità; la sua realizzazione determina come ottenere concretamente il bit $f(x)$.

La probabilità da modificare

Con una sola soluzione su N candidati, una scelta uniforme riesce con probabilità $1/N$. Grover prepara inizialmente una distribuzione di questo tipo, poi applica operazioni che modificano le ampiezze associate ai candidati. Alla fine misura il registro e controlla il candidato ottenuto.

La misura produce una singola stringa. Non fornisce una lista di tutte le verifiche compiute in sovrapposizione, né permette di leggere arbitrariamente ogni componente dello stato. Il progetto dell'algoritmo consiste quindi nel rendere probabile un esito valido prima della misura. Per comprendere questo meccanismo occorre distinguere con precisione le ampiezze, che le operazioni trasformano, dalle probabilità, che determinano gli esiti osservabili. Cominceremo da tale distinzione prima di usare il verificatore in un circuito quantistico.

Richiami La ricerca ha successo quando il candidato finale x soddisfa $f(x) = 1$.

Un confronto su sedici candidati

Supponiamo che una sola soluzione sia distribuita uniformemente fra sedici candidati. Vogliamo confrontare quattro controlli distinti con quattro tentativi indipendenti che possono ripetere lo stesso candidato. Nel primo caso, le posizioni coperte sono quattro su sedici, quindi la probabilità di successo verificato è $1/4$.

Nel secondo caso ogni tentativo fallisce con probabilità $15/16$. I quattro fallimenti hanno probabilità $(15/16)^4 = 50625/65536$, perciò il successo ha probabilità $14911/65536$, circa il $22{,}8\%$. La differenza non dipende dalla rapidità del verificatore: nasce dal fatto che alcuni tentativi possono ripetere controlli già effettuati. Questi due modelli torneranno utili quando parleremo di riavvii dell'algoritmo quantistico e della verifica dei risultati.

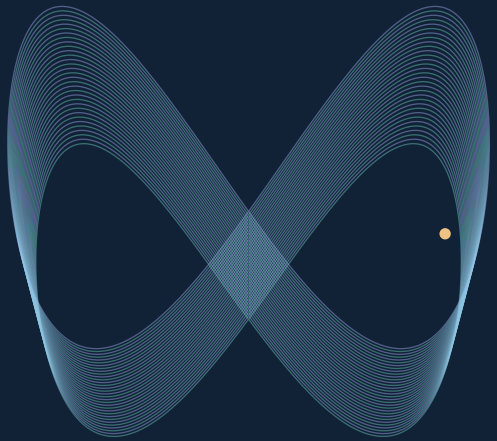
$$\frac{4}{16} = 25\%, \qquad 1 - \left(\frac{15}{16}\right)^4 \approx 22{,}8\%$$

Richiami Con una soluzione, q posizioni distinte danno probabilità q/N; q tentativi indipendenti danno $1 - (1 - 1/N)^q$.

CAPITOLO 02

Stati e probabilità

La preparazione dei registri e le ampiezze



L’algoritmo di Grover

Le etichette degli stati di un registro

Un bit classico può contenere 0 oppure 1. Per descrivere i due stati corrispondenti di un qubit usiamo i simboli $|0\rangle$ e $|1\rangle$, letti «ket zero» e «ket uno». Le parentesi distinguono uno stato del sistema dal numero che lo identifica. La notazione sarà particolarmente utile quando uno stato coinvolgerà più etichette contemporaneamente.

Per due qubit, gli stati associati alle quattro stringhe classiche sono $|00\rangle$, $|01\rangle$, $|10\rangle$ e $|11\rangle$. Fissiamo questo ordine per elencare le coordinate. Anche $|3\rangle$ può indicare $|11\rangle$, purché sia chiaro che il registro contiene due qubit. Il numero di etichette rimane $N = 2^n$ per un registro di n qubit; quello che cambia rispetto a un registro classico è il tipo di stato ammesso.

Richiami Un registro di n bit ha 2^n stringhe possibili; $|x\rangle$ indica lo stato associato alla stringa x.

Gli stati di base come vettori

Possiamo rappresentare gli stati di un qubit mediante colonne di due numeri. Lo stato $|0\rangle$ ha coordinate $(1, 0)$ e lo stato $|1\rangle$ ha coordinate $(0, 1)$. Questi vettori costituiscono una base: ogni vettore a due componenti si può esprimere come una loro combinazione.

La stessa costruzione vale per due qubit, ma richiede quattro coordinate. Nell'ordine $00, 01, 10, 11$, il vettore di $|10\rangle$ è $(0, 0, 1, 0)$. La coordinata uguale a 1 indica l'etichetta presente. Per un registro di tre qubit occorrono otto coordinate. Non disegniamo dunque un piano fisico per ogni qubit: stiamo organizzando matematicamente le possibili etichette dell'intero registro, che formeranno la base computazionale.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Richiami La base computazionale comprende uno stato $|x\rangle$ per ogni stringa binaria del registro.

Le ampiezze di una sovrapposizione

Uno stato quantistico può essere una combinazione lineare degli stati di base. Per un qubit scriviamo $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$. I coefficienti a_0 e a_1 si chiamano ampiezze. «Combinazione lineare» significa che moltiplichiamo ciascun vettore per il suo coefficiente e sommiamo le coordinate corrispondenti.

Uno stato con più coefficienti non nulli è una sovrapposizione rispetto alla base scelta. Per esempio, $(|0\rangle + |1\rangle)/\sqrt{2}$ ha due ampiezze uguali, entrambe $1/\sqrt{2}$. La scrittura non afferma che la misura mostrerà contemporaneamente 0 e 1. Specifica invece lo stato su cui agiranno le operazioni e da cui ricaveremo le probabilità dei singoli esiti. Per interpretare queste probabilità dobbiamo anche imporre una condizione sulla grandezza delle ampiezze.

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle \quad \longleftrightarrow \quad \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

Richiami Le ampiezze sono i coefficienti dello stato nella base computazionale; non sono direttamente probabilità.

La normalizzazione delle ampiezze

Per ora prendiamo ampiezze reali. La probabilità di misurare un'etichetta è il quadrato della sua ampiezza. Poiché gli esiti possibili devono avere probabilità totale 1, le ampiezze soddisfano $a_0^2 + a_1^2 = 1$. Questa condizione si chiama normalizzazione. Anche un'ampiezza negativa contribuisce con un quadrato positivo.

Se le due probabilità devono essere uguali, ciascuna vale $1/2$. Possiamo allora scegliere ampiezze positive uguali a $1/\sqrt{2}$, perché $(1/\sqrt{2})^2 = 1/2$. La radice quadrata compare proprio per passare dalla probabilità all'ampiezza. Due ampiezze uguali a $1/2$ darebbero invece probabilità totale $1/4 + 1/4 = 1/2$: quel vettore, da solo, non rappresenterebbe uno stato normalizzato di un qubit.

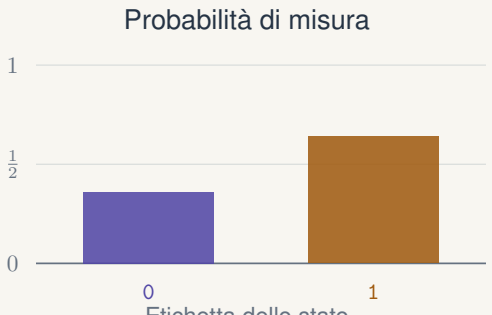
$$\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = 1$$

Richiami Per ampiezze reali, $P(x) = a_x^2$ e la normalizzazione richiede $\sum_x a_x^2 = 1$.

Due esiti con probabilità diverse

Consideriamo lo stato $|\psi\rangle = (3/5)|0\rangle + (4/5)|1\rangle$. Il primo esito ha probabilità $9/25$, il secondo $16/25$. La somma vale 1, dunque lo stato è normalizzato. Il fatto che le ampiezze sommino a $7/5$ non è un problema: la condizione riguarda la somma dei quadrati, non la somma delle ampiezze.

Ripetendo molte volte la stessa preparazione e la stessa misura, le frequenze tendono a riflettere queste probabilità. In un gruppo di cento esperimenti ci aspettiamo in media 36 zeri e 64 uni, senza che tali conteggi siano garantiti per ogni gruppo. Una singola esecuzione restituisce soltanto uno dei due risultati. Il grafico riporta probabilità teoriche, non conteggi osservati.



$$\frac{9}{25} + \frac{16}{25} = 1$$

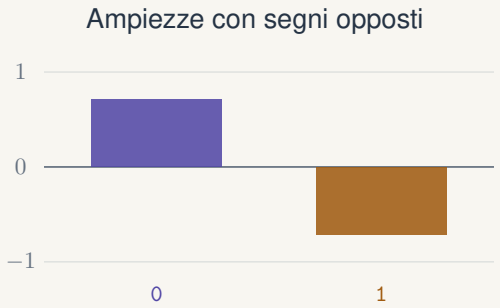
Le due barre rappresentano $P(0) = 0{,}36$ e $P(1) = 0{,}64$. Il colore distingue un esito per facilitarne la lettura, senza modificare la misura.

Richiami Le probabilità sono quadrati delle ampiezze reali; la loro somma deve essere 1.

Il segno di un’ampiezza

I due stati $(|0\rangle + |1\rangle)/\sqrt{2}$ e $(|0\rangle - |1\rangle)/\sqrt{2}$ producono entrambi probabilità $1/2$ per 0 e $1/2$ per 1 quando misurati subito nella base computazionale. Il quadrato elimina infatti il segno. Questa coincidenza non rende però identici i due stati.

Un’operazione successiva può sommare contributi provenienti da componenti diverse. I segni determinano allora se i contributi si aggiungono o si cancellano. La differenza fra i due stati riguarda il segno di una componente rispetto all’altra, quindi una relazione interna allo stato. Mostreremo con la porta Hadamard come tale differenza possa trasformarsi in esiti di misura opposti. Grover sfrutterà proprio una modifica inizialmente invisibile alle probabilità immediate.



$$\left(\frac{1}{\sqrt{2}}\right)^2 = \left(-\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$$

Le barre hanno la stessa lunghezza e segni opposti. I quadrati delle due ampiezze valgono entrambi $1/2$, ma i contributi possono interferire diversamente.

Richiami Un’ampiezza può essere negativa; la probabilità corrispondente è sempre non negativa.

Le ampiezze complesse

In generale, le ampiezze possono essere numeri complessi. Scriviamo un numero complesso come $a + ib$, dove a e b sono reali e $i^2 = -1$. Il suo modulo quadro è $|a + ib|^2 = a^2 + b^2$. È questa quantità, sempre non negativa, che fornisce la probabilità associata a un'ampiezza complessa.

Per esempio, l'ampiezza $(1 + i)/2$ ha modulo quadro $1/4 + 1/4 = 1/2$. Anche $(1 - i)/2$ ha modulo quadro $1/2$. Lo stato che assegna questi due coefficienti rispettivamente a 0 e 1 è quindi normalizzato. Le prime costruzioni di Grover useranno soltanto ampiezze reali, positive o negative; introdurre ora il modulo quadro ci permette però di formulare correttamente le regole generali senza confondere un caso particolare con una limitazione della teoria.

$$P(x) = |a_x|^2, \qquad \sum_x |a_x|^2 = 1$$

Richiami Per $z = a + ib$, il modulo quadro è $|z|^2 = a^2 + b^2$; per un numero reale coincide con il quadrato ordinario.

L’azione lineare di una porta

Una porta quantistica ideale agisce linearmente sullo stato. Se conosciamo la sua azione su ciascun vettore della base, possiamo quindi ricavare l’azione su ogni sovrapposizione: applichiamo la porta ai singoli vettori, manteniamo i loro coefficienti e sommiamo i risultati.

La linearità è una regola di calcolo, non una procedura per leggere separatamente tutte le componenti. Dopo aver sommato i contributi dobbiamo ancora calcolare le probabilità e misurare un solo esito. Le porte reversibili ammesse devono inoltre preservare la normalizzazione degli stati. La loro descrizione completa usa operatori unitari, che riprenderemo quando studieremo le riflessioni. Per i primi esempi verificheremo direttamente gli effetti sulle ampiezze e sulla somma dei moduli quadrati.

$$U(a|0\rangle + b|1\rangle) = aU|0\rangle + bU|1\rangle$$

Richiami Una combinazione lineare moltiplica i vettori per coefficienti e somma le coordinate corrispondenti.

La fase comune a tutto lo stato

Moltiplichiamo tutte le ampiezze di uno stato per lo stesso numero complesso di modulo 1. Per esempio, possiamo moltiplicarle tutte per -1 oppure tutte per i. Le probabilità non cambiano, perché il modulo quadro del fattore comune è 1. Chiamiamo questa modifica una fase globale.

La sua irrilevanza non riguarda soltanto una misura immediata. Per un'operazione lineare U, il fattore comune rimane comune: $U(i|\psi\rangle) = iU|\psi\rangle$. Anche dopo l'operazione, ogni probabilità coincide con quella ottenuta partendo da $|\psi\rangle$. Questa affermazione vale quando il fattore moltiplica l'intero stato del sistema considerato. Un segno applicato soltanto ad alcune componenti modifica invece una fase relativa e può influire sull'interferenza.

$$|-a_x|^2 = |a_x|^2, \qquad |ia_x|^2 = |a_x|^2$$

Richiami Una fase globale moltiplica tutte le ampiezze per lo stesso fattore di modulo 1.

La porta Hadamard sui due stati di base

La porta Hadamard, indicata con H, trasforma $|0\rangle$ nella sovrapposizione con due ampiezze positive uguali e trasforma $|1\rangle$ nella sovrapposizione che ha il secondo coefficiente negativo. Le formule riportate definiscono completamente la sua azione su un qubit, grazie alla linearità.

In entrambi i casi la misura immediata dà 0 e 1 con probabilità uguali. Tuttavia il segno diverso conserva la distinzione fra i due ingressi: un'altra applicazione di H permetterà di recuperarli. Una trasformazione che mandasse entrambi gli ingressi nello stesso stato perderebbe invece informazione e non potrebbe avere un'inversa. La coppia di uscite di Hadamard contiene dunque una differenza essenziale, benché non appaia nelle probabilità immediate della base computazionale.

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

Richiami La linearità determina l'azione di una porta su tutte le sovrapposizioni a partire dall'azione sugli stati di base.

Hadamard come somma e differenza

Applichiamo la linearità allo stato $a|0\rangle + b|1\rangle$. Il termine $a|0\rangle$ contribuisce con $a/\sqrt{2}$ a entrambe le uscite. Il termine $b|1\rangle$ contribuisce con $b/\sqrt{2}$ all'uscita 0 e con $-b/\sqrt{2}$ all'uscita 1. Per ogni uscita sommiamo i contributi che arrivano alla stessa etichetta.

Otteniamo così le nuove ampiezze $(a+b)/\sqrt{2}$ e $(a-b)/\sqrt{2}$. La prima contiene una somma, la seconda una differenza: un segno relativo negli ingressi può favorire una delle due uscite e cancellare l'altra. Questa combinazione dei contributi è ciò che chiamiamo interferenza. Il calcolo avviene sulle ampiezze; elevare al quadrato ciascun contributo prima di sommarli darebbe, in generale, probabilità sbagliate.

$$H \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} a+b \\ a-b \end{pmatrix}$$

Richiami I contributi alla stessa ampiezza si sommano prima di calcolare il modulo quadro.

La cancellazione dell'uscita uno

Prepariamo $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ e applichiamo Hadamard. L'ampiezza dell'uscita 0 è $(1/\sqrt{2} + 1/\sqrt{2})/\sqrt{2} = 1$. Quella dell'uscita 1 è $(1/\sqrt{2} - 1/\sqrt{2})/\sqrt{2} = 0$. La misura finale restituisce quindi 0 con certezza.

Il simbolo $|+\rangle$ è soltanto un nome breve per la sovrapposizione appena scritta. Non indica un terzo esito della misura nella base computazionale. Prima della porta, quella misura avrebbe avuto due esiti equiprobabili; dopo la porta, l'interferenza ha concentrato tutta la probabilità su uno solo. Poiché $|+\rangle = H|0\rangle$, il calcolo dimostra anche che due Hadamard consecutive riportano l'ingresso $|0\rangle$ al suo stato iniziale.

$$H|+\rangle = |0\rangle$$

Richiami $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$; Hadamard produce le ampiezze $(a+b)/\sqrt{2}$ e $(a-b)/\sqrt{2}$.

La cancellazione dell'uscita zero

Consideriamo ora $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. L'ampiezza dell'uscita 0 dopo Hadamard è $(1/\sqrt{2} - 1/\sqrt{2})/\sqrt{2} = 0$. Quella dell'uscita 1 è $(1/\sqrt{2} + 1/\sqrt{2})/\sqrt{2} = 1$. La misura finale restituisce dunque 1 con certezza.

Gli stati $|+\rangle$ e $|-\rangle$ avevano le stesse probabilità prima della porta, ma danno esiti opposti dopo la stessa operazione. Il segno relativo contiene perciò un'informazione che le probabilità immediate non mostravano. Inoltre, poiché $|-\rangle = H|1\rangle$, due Hadamard riportano anche $|1\rangle$ al suo stato iniziale. Per linearità il risultato vale su ogni sovrapposizione: $H^2 = I$, dove I indica l'operazione identità, che lascia lo stato invariato.

$$H|-\rangle = |1\rangle, \qquad H^2 = I$$

Richiami $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$; una fase relativa può modificare l'esito di operazioni successive.

Due qubit preparati separatamente

Per descrivere due qubit preparati separatamente usiamo il prodotto tensoriale dei loro stati. Nella base computazionale, $|0\rangle \otimes |1\rangle = |01\rangle$: il simbolo $\otimes$ combina due sistemi e concatena le loro etichette. Non rappresenta una somma né una scelta fra i due qubit.

Se entrambi sono in sovrapposizione, sviluppiamo il prodotto distribuendo ogni termine del primo fattore su ogni termine del secondo. Il coefficiente di una coppia è il prodotto dei due coefficienti. Per esempio, $|+\rangle \otimes |0\rangle = (|00\rangle + |10\rangle)/\sqrt{2}$. Soltanto il primo bit varia, mentre il secondo rimane 0. Ometteremo spesso il simbolo di prodotto e scriveremo $|+\rangle|0\rangle$; l'ordine dei fattori continuerà a identificare l'ordine dei qubit.

$$|+\rangle|0\rangle = \frac{|00\rangle + |10\rangle}{\sqrt{2}}$$

Richiami Nel prodotto di stati preparati separatamente, le etichette si concatenano e le ampiezze si moltiplicano.

La preparazione uniforme di quattro etichette

Partiamo da due qubit in $|00\rangle$ e applichiamo Hadamard a ciascuno. Ogni qubit diventa $|+\rangle$, quindi lo stato complessivo è il prodotto $|+\rangle|+\rangle$. Sviluppando i due fattori otteniamo quattro termini, uno per ogni stringa di due bit.

Il coefficiente di ciascun termine è $(1/\sqrt{2})(1/\sqrt{2}) = 1/2$. Ogni etichetta ha pertanto probabilità $1/4$, e le quattro probabilità sommano a 1. La notazione $H \otimes H$ indica che la stessa porta agisce sui due qubit, non che una singola Hadamard abbia improvvisamente quattro ingressi. Questa preparazione sarà lo stato iniziale dell'esempio più piccolo di Grover, nel quale cercheremo una soluzione fra le quattro etichette.

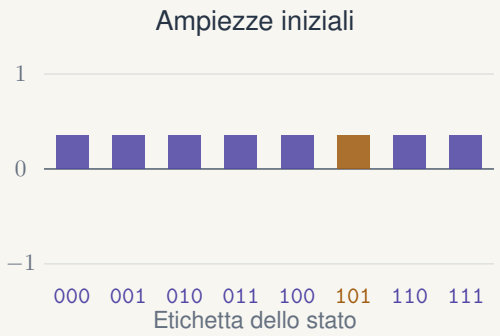
$$(H \otimes H)|00\rangle = \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2}$$

Richiami Una probabilità uniforme su quattro etichette vale $1/4$ per etichetta, quindi un'ampiezza positiva uniforme vale $1/2$.

La preparazione uniforme di otto etichette

Aggiungendo un terzo qubit inizialmente in $|0\rangle$ e applicandogli Hadamard, ciascuno dei quattro termini precedenti si divide in due termini. Si ottengono le otto stringhe da 000 a 111, tutte con ampiezza $(1/\sqrt{2})^3 = 1/\sqrt{8}$.

Il simbolo $\sum$ nella formula richiede di sommare i termini delle otto etichette. Ogni probabilità vale $1/8$. Le barre del grafico hanno tutte la stessa altezza, anche quella del candidato 5 che useremo come soluzione. Il colore ne evidenzia soltanto la posizione per il lettore: la preparazione uniforme non conosce il risultato del verificatore e non tratta quella posizione in modo speciale. Il vantaggio dovrà apparire nelle operazioni successive.



$$|s\rangle = \frac{1}{\sqrt{8}} \sum_{x=0}^{7} |x\rangle$$

Tutte le ampiezze iniziali valgono $1/\sqrt{8}$. La barra evidenziata corrisponde a 101, ma ha la stessa altezza delle altre sette.

Richiami Il simbolo $\sum_{x=0}^{7}$ richiede di sommare un termine per ciascuna delle otto etichette.

La sovrapposizione uniforme su un registro

Per n qubit inizialmente in $|0^n\rangle$, dove 0^n indica una stringa di n zeri, applichiamo Hadamard a ogni qubit. Scriviamo questa operazione come $H^{\otimes n}$. La distribuzione dei prodotti genera tutte le $N = 2^n$ stringhe e assegna a ciascuna ampiezza $1/\sqrt{N}$.

Indichiamo lo stato risultante con $|s\rangle$, abbreviazione che useremo lungo tutto il calcolo. La sua normalizzazione segue direttamente: ci sono N termini, ognuno con modulo quadro $1/N$, quindi la somma è 1. Preparare questo stato richiede n porte Hadamard. Questo conteggio non significa aver ottenuto N risultati leggibili: la misura del registro produrrà comunque una sola etichetta.

$$|s\rangle = H^{\otimes n}|0^n\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$$

$$\sum_{x=0}^{N-1} \frac{1}{N} = 1$$

Richiami $N = 2^n$; $H^{\otimes n}$ applica Hadamard separatamente a ciascuno degli n qubit.

La misura prima dell'amplificazione

Se misuriamo subito lo stato uniforme $|s\rangle$, ogni candidato compare con probabilità $1/N$. Se le soluzioni sono M, la probabilità complessiva di ottenere una soluzione si calcola sommando le probabilità delle M etichette valide: il risultato è M/N.

Non sommiamo le loro ampiezze prima di elevarle al quadrato, perché si tratta di esiti distinti della misura. L'interferenza riguardava invece contributi diversi alla stessa ampiezza d'uscita. Con due soluzioni su otto etichette, le probabilità $1/8$ e $1/8$ danno successo totale $1/4$. Questa è la stessa probabilità di una scelta classica uniforme. La sola preparazione della sovrapposizione, seguita immediatamente dalla misura, non realizza ancora l'amplificazione cercata.

$$P(\text{soluzione nella misura di } |s\rangle) = \frac{M}{N}$$

Richiami Per esiti distinti si sommano le probabilità; per contributi alla stessa ampiezza si sommano prima le ampiezze.

Ripetere una preparazione e ripetere una misura

La misura nella base computazionale restituisce un'etichetta e lascia il registro nello stato di base corrispondente. Se misuriamo subito di nuovo lo stesso registro, senza modificarlo, otteniamo ancora quell'etichetta. Le due letture non sono campioni indipendenti della distribuzione precedente alla prima misura.

Per ottenere nuovi campioni dobbiamo ripetere la preparazione e il circuito, poi misurare nuovamente. Per esempio, quattro misure consecutive dello stesso registro appena collassato in $|101\rangle$ danno quattro volte 101. Quattro esecuzioni indipendenti della preparazione uniforme su tre qubit possono invece dare quattro etichette diverse oppure alcune ripetizioni. Quando parleremo di riavvio di Grover, intenderemo una nuova esecuzione a partire dallo stato iniziale, non una nuova lettura del risultato già prodotto.

Richiami Una nuova esecuzione riprepara lo stato iniziale; una seconda misura dello stesso stato di base ne ripete l'etichetta.

Il controllo di uno stato a quattro componenti

Consideriamo le quattro ampiezze $1/2$, $1/2$, $1/2$ e $-1/2$, nell'ordine $00, 01, 10, 11$. Per verificare la normalizzazione calcoliamo quattro moduli quadrati uguali a $1/4$: la somma è 1. La somma delle ampiezze è invece 1, ma questa uguaglianza numerica è accidentale e non costituisce il criterio di normalizzazione.

La probabilità della stringa 11 è ancora $1/4$, nonostante la sua ampiezza negativa. Se moltiplichiamo tutte e quattro le ampiezze per -1, otteniamo lo stesso stato fisico a meno di una fase globale. Se cambiamo soltanto l'ultima da $-1/2$ a $1/2$, modifichiamo invece una fase relativa. Quest'ultimo confronto anticipa esattamente il tipo di marcatura che applicherà l'oracolo: prima cambierà un segno, mentre un'operazione successiva ne renderà osservabili gli effetti.

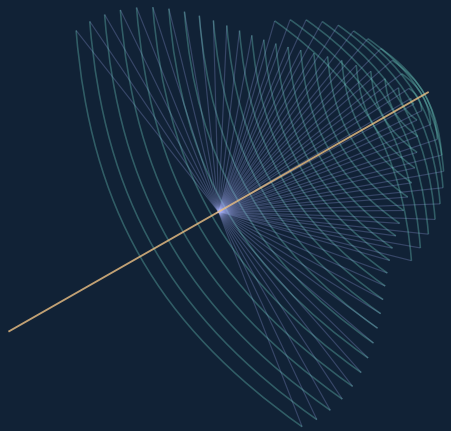
$$\left|\frac{1}{2}\right|^2 + \left|\frac{1}{2}\right|^2 + \left|\frac{1}{2}\right|^2 + \left|-\frac{1}{2}\right|^2 = 1$$

Richiami La normalizzazione usa i moduli quadrati; una fase globale cambia tutte le ampiezze con lo stesso fattore.

CAPITOLO 03

Oracoli reversibili

Il valore della funzione
e la marcatura di fase



L’algoritmo di Grover

La verifica come operazione reversibile

Una funzione booleana ordinaria può mandare molti ingressi nello stesso risultato. Con quattro candidati e una sola soluzione, tre ingressi diversi producono 0. Se sostituissimo lo stato dell'ingresso soltanto con il risultato $f(x)$, perderemmo la possibilità di ricostruire quale candidato era stato presentato.

Le porte quantistiche ideali che useremo sono reversibili: la loro azione completa deve ammettere un'inversa. Per rendere reversibile la verifica, conserveremo quindi l'ingresso x e affiancheremo un qubit in cui combinare il risultato con un valore precedente. Il qubit aggiuntivo si chiama ausiliario, o ancilla. Questa costruzione non impedisce di calcolare una funzione con molte risposte uguali: richiede di registrare l'informazione in un sistema abbastanza ampio, senza cancellare l'ingresso.

Richiami Un predicato f restituisce 0 oppure 1; un'operazione reversibile permette di ricostruire lo stato precedente.

Lo XOR come aggiornamento reversibile

Per combinare il risultato con l'ausiliario useremo lo XOR, indicato da $\oplus$. Su due bit, lo XOR vale 1 se i bit sono diversi e 0 se sono uguali. Si può anche leggere come somma modulo 2: dopo 1 si ritorna a 0.

Aggiungere 0 tramite XOR lascia il bit invariato; aggiungere 1 lo scambia, da 0 a 1 oppure da 1 a 0. Ripetendo lo stesso aggiornamento si recupera sempre il bit iniziale. Per esempio, $(0 \oplus 1) \oplus 1 = 1 \oplus 1 = 0$. Questa proprietà permette di scrivere un risultato su un ausiliario senza introdurre un'operazione irreversibile, purché il valore da combinare resti ricostruibile dall'ingresso conservato.

b	c	$b \oplus c$
0	0	0
0	1	1
1	0	1
1	1	0

$$(b \oplus c) \oplus c = b$$

La tabella comprende tutte le quattro coppie di bit. Fissato c, l'aggiornamento di b può essere annullato applicando di nuovo lo XOR con lo stesso valore.

Richiami Lo XOR $b \oplus c$ vale 1 quando i due bit differiscono e 0 quando coincidono.

Il registro dei candidati e il qubit ausiliario

Il sistema comprende ora due parti: un registro di n qubit che identifica x e un qubit ausiliario che contiene un bit b. Uno stato di base si scrive $|x\rangle|b\rangle$. Il prodotto indica i due registri, nello stesso ordine in cui li leggeremo nei circuiti.

Per tre qubit di ingresso, $|101\rangle|0\rangle$ contiene quattro qubit complessivi. Il candidato è ancora 5, mentre l'ultimo qubit è un supporto del calcolo e non un quarto bit della sua etichetta. Questa separazione evita di confondere il numero di candidati con il numero totale di stati del dispositivo: il registro di ingresso ammette otto candidati anche quando il circuito usa ulteriori qubit di lavoro.

$$|x\rangle|b\rangle = |x\rangle \otimes |b\rangle$$

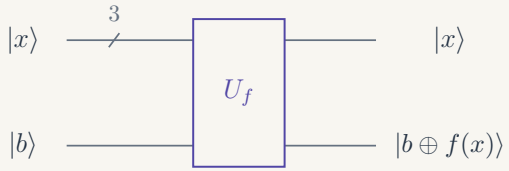
Richiami Il prodotto tensoriale combina registri diversi; le etichette specificano lo stato di ciascun registro.

La definizione dell'oracolo XOR

Chiamiamo U_f l'operazione che conserva il candidato e aggiorna il qubit ausiliario mediante XOR con il risultato della verifica. Su uno stato di base agisce quindi come $|x\rangle|b\rangle \mapsto |x\rangle|b \oplus f(x)\rangle$. Il simbolo U_f indica una trasformazione del sistema quantistico, mentre f indica il predicato classico.

Se l'ausiliario parte da 0, il nuovo valore è $f(x)$. Se parte da 1, il nuovo valore è il complemento di $f(x)$. Entrambi i casi fanno parte della stessa definizione e sono necessari per descrivere correttamente l'azione su una sovrapposizione dell'ausiliario. Il circuito a blocchi lascia ancora astratto il calcolo interno di f; in seguito costruiremo un verificatore concreto con porte elementari.

Oracolo con registro di risposta



$$b \longmapsto b \oplus f(x)$$

$$U_f|x\rangle|b\rangle = |x\rangle|b \oplus f(x)\rangle$$

Il blocco riceve il candidato e l'ausiliario. Conserva l'ingresso x e combina il risultato con il qubit inferiore attraverso lo XOR.

Richiami x identifica il candidato; b è il bit ausiliario; $f(x) = 1$ segnala una soluzione.

La verifica con l'ausiliario inizialmente a zero

Prendiamo quattro candidati e il predicato che vale 1 soltanto su 11. Se l'ausiliario parte da 0, i candidati 00, 01 e 10 lo lasciano a 0, mentre 11 lo porta a 1. L'ingresso rimane sempre leggibile nel primo registro.

Le uscite complete sono tutte diverse, benché tre valori del predicato coincidano. Per esempio, $|00\rangle|0\rangle$ e $|01\rangle|0\rangle$ hanno lo stesso ausiliario, ma registri di ingresso differenti. La possibilità di distinguere queste uscite è ciò che mancava alla trasformazione irreversibile che avrebbe conservato soltanto $f(x)$. Il risultato della funzione occupa qui un posto aggiuntivo anziché sostituire l'etichetta del candidato.

x	$f(x)$	**Ausiliario finale**
00	0	$\|0\rangle$
01	0	$\|0\rangle$
10	0	$\|0\rangle$
11	1	$\|1\rangle$

La tabella riguarda l'ausiliario inizialmente in $|0\rangle$. Per ogni riga il registro del candidato mantiene la stessa stringa mostrata nella prima colonna.

Richiami Per $b = 0$, lo XOR produce $0 \oplus f(x) = f(x)$ e il primo registro conserva x.

L’annullamento della verifica

Per lo stesso predicato, se il candidato è 11 e l’ausiliario parte da 1, l’oracolo produce $1 \oplus 1 = 0$. Esso scambia dunque $|11\rangle|0\rangle$ e $|11\rangle|1\rangle$. Per gli altri tre candidati, $f(x) = 0$ e il qubit ausiliario rimane invariato.

Applicando due volte l’oracolo recuperiamo ogni stato di base iniziale: il candidato non cambia e l’ausiliario riceve due volte lo stesso XOR. La linearità estende il risultato a tutte le sovrapposizioni. Scriviamo quindi $U_f^2 = I$, ossia U_f è la propria inversa. La descrizione come permutazione degli stati di base mostra anche che nessuna probabilità totale viene persa: si riorganizzano le etichette delle ampiezze senza modificarne i moduli.

$$b \oplus f(x) \oplus f(x) = b, \qquad U_f^2 = I$$

Richiami I è l’identità, che lascia invariato lo stato; un doppio XOR con lo stesso bit annulla il primo.

La verifica di una sovrapposizione uniforme

Prepariamo i quattro candidati con ampiezza $1/2$ e l'ausiliario in $|0\rangle$. La linearità richiede di applicare U_f a ciascun termine. I primi tre termini conservano l'ausiliario a 0, mentre il termine associato a 11 lo porta a 1.

Lo stato finale contiene una correlazione fra i due registri: osservare il valore dell'ausiliario fornisce informazione sul candidato. Non possiamo più descrivere il risultato come il prodotto di uno stato del primo registro e di uno stato dell'ausiliario. Una correlazione quantistica di questo tipo si chiama entanglement. Le quattro ampiezze del sistema completo hanno ancora modulo $1/2$, perciò la sola applicazione dell'oracolo non ha aumentato la probabilità di trovare 11 nella misura del registro di ingresso.

$$\frac{|00\rangle|0\rangle + |01\rangle|0\rangle + |10\rangle|0\rangle + |11\rangle|1\rangle}{2}$$

Richiami La linearità applica l'oracolo a ogni termine mantenendone il coefficiente; lo stato completo comprende entrambi i registri.

La misura del risultato della verifica

Nello stato appena costruito, l'ausiliario vale 1 con probabilità $1/4$, perché un solo termine su quattro lo contiene. Se misuriamo 1, il registro del candidato si trova in $|11\rangle$ e abbiamo trovato la soluzione. Se misuriamo 0, evento di probabilità $3/4$, rimangono soltanto i tre candidati non validi.

In quest'ultimo caso, le tre ampiezze uguali devono essere rinormalizzate. Ciascuna diventa $1/\sqrt{3}$, così i tre moduli quadrati sommano a 1. Il successo immediato ha però probabilità $1/4$, proprio come una scelta uniforme fra quattro candidati. Misurare il bit della verifica non ha creato un'amplificazione. Cercheremo invece di usare coerentemente il risultato del predicato per cambiare i segni delle ampiezze prima della misura finale.

$$|\psi \mid b = 0\rangle = \frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}$$

Richiami La rinormalizzazione rende uguale a 1 la somma delle probabilità degli esiti rimasti dopo una condizione osservata.

La porta che scambia zero e uno

La porta X scambia i due stati di base di un qubit: $X|0\rangle = |1\rangle$ e $X|1\rangle = |0\rangle$. È la versione quantistica della negazione di un bit, estesa per linearità a tutte le sovrapposizioni. Sullo stato $a|0\rangle + b|1\rangle$ produce quindi $b|0\rangle + a|1\rangle$.

L'oracolo XOR può essere letto usando questa porta. Per ogni candidato x, esso applica X all'ausiliario se $f(x) = 1$ e non lo modifica se $f(x) = 0$. Il primo registro decide quale delle due azioni corrisponde al termine considerato. Questa lettura rende utile scegliere uno stato dell'ausiliario il cui comportamento sotto X sia particolarmente semplice: useremo gli stati $|+\rangle$ e $|-\rangle$ già ottenuti con Hadamard.

$$X \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a \end{pmatrix}$$

Richiami $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ e $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$.

Il segno prodotto da X sullo stato meno

Possiamo preparare $|-\rangle$ partendo da $|0\rangle$: prima X produce $|1\rangle$, poi Hadamard produce $|-\rangle$. Applichiamo ora X allo stato così preparato. Il termine $|0\rangle$ diventa $|1\rangle$ e il termine $-|1\rangle$ diventa $-|0\rangle$. Riordinando la somma, otteniamo $(-|0\rangle + |1\rangle)/\sqrt{2}$.

Questo risultato è uguale a $-|-\rangle$: basta raccogliere il fattore -1 davanti alla parentesi. L'ausiliario torna dunque allo stesso stato a meno di un segno. Se osservassimo soltanto questo qubit, il segno globale non avrebbe conseguenze. Quando però X viene applicata soltanto per alcuni candidati di una sovrapposizione, il segno distingue quei termini dagli altri. Sarà quindi una fase relativa dello stato complessivo, capace di influire sulle operazioni successive.

$$X|-\rangle = \frac{|1\rangle - |0\rangle}{\sqrt{2}} = -|-\rangle$$

Richiami Un segno comune all'intero stato è globale; un segno applicato soltanto ad alcune componenti è relativo.

Lo stato più lascia il predicato senza effetto

Per confronto, applichiamo X a $|+\rangle$. Lo scambio dei due termini non cambia la loro somma: $X|+\rangle = (|1\rangle + |0\rangle)/\sqrt{2} = |+\rangle$. La porta non produce neppure un segno negativo.

Se usassimo questo stato per l'ausiliario dell'oracolo, ogni candidato lo lascerebbe in $|+\rangle$: sia il caso $f(x) = 0$, che non applica X, sia il caso $f(x) = 1$, che la applica. L'intero stato del registro di ingresso rimarrebbe allora immutato. La scelta dell'ausiliario è quindi parte della costruzione dell'algoritmo. Un qubit aggiuntivo non trasferisce automaticamente il risultato della verifica alle ampiezze in una forma utile; lo stato $|-\rangle$ realizza precisamente la differenza di segno di cui abbiamo bisogno.

$$U_f|x\rangle|+\rangle = |x\rangle|+\rangle$$

Richiami $X|+\rangle = |+\rangle$, mentre $X|-\rangle = -|-\rangle$.

L’ausiliario meno per un candidato non valido

Supponiamo che $f(x) = 0$ e prepariamo l’ausiliario in $|-\rangle$. Nella definizione di U_f, lo XOR con 0 lascia invariati entrambi i valori di base dell’ausiliario. Applicando la linearità, ogni termine della differenza $|0\rangle - |1\rangle$ resta quindi identico.

Il risultato è $|x\rangle|-\rangle$, senza alcuna modifica dell’ampiezza del candidato. Nel caso su quattro stringhe, questo comportamento riguarda 00, 01 e 10. Abbiamo esaminato separatamente un candidato non valido per rendere esplicito che il segno non viene cambiato per tutti gli ingressi: la trasformazione dipende dal valore del predicato e conserverà così una distinzione fra le componenti della sovrapposizione.

$$f(x) = 0 : \quad U_f|x\rangle|-\rangle = |x\rangle|-\rangle$$

Richiami Lo XOR con 0 lascia invariato un bit; l’oracolo applica l’identità all’ausiliario quando $f(x) = 0$.

L’ausiliario meno per un candidato valido

Se $f(x) = 1$, lo XOR scambia invece i due stati di base dell’ausiliario. La differenza $|0\rangle - |1\rangle$ diventa $|1\rangle - |0\rangle$, che è l’opposto della differenza iniziale. Otteniamo dunque $-|x\rangle|-\rangle$.

Il fattore -1 può essere scritto davanti al prodotto completo oppure davanti al primo fattore: $-|x\rangle|-\rangle = (-|x\rangle)|-\rangle$. Questa uguaglianza algebrica permette di descrivere l’azione come un cambio di segno della componente del candidato, mentre l’ausiliario rimane nello stesso stato $|-\rangle$. Riunendo i casi valido e non valido, il fattore è $(-1)^{f(x)}$: vale 1 quando $f(x) = 0$ e -1 quando $f(x) = 1$.

$$U_f|x\rangle|-\rangle = (-1)^{f(x)}|x\rangle|-\rangle$$

Richiami $(-1)^0 = 1$ e $(-1)^1 = -1$; l’ausiliario $|-\rangle$ trasforma lo scambio in un fattore di segno.

Il fattore ausiliario comune alla sovrapposizione

Applichiamo ora il risultato a una sovrapposizione $\sum_x a_x|x\rangle$ del registro dei candidati. La linearità assegna a ogni termine il fattore $(-1)^{f(x)}$. Tutti i termini conservano lo stesso stato ausiliario $|-\rangle$, che possiamo quindi raccogliere come fattore comune del prodotto.

Il sistema finale è nuovamente un prodotto fra il registro dei candidati e l'ausiliario. A differenza della preparazione dell'ausiliario in $|0\rangle$, non abbiamo lasciato un valore distinguibile che identifichi le componenti valide. L'informazione del predicato compare ora nei segni delle ampiezze del primo registro. Questo fenomeno viene spesso chiamato trasferimento di fase, o phase kickback; la formula mostra esplicitamente quale fase cambia e perché l'ausiliario può essere separato dalla descrizione successiva.

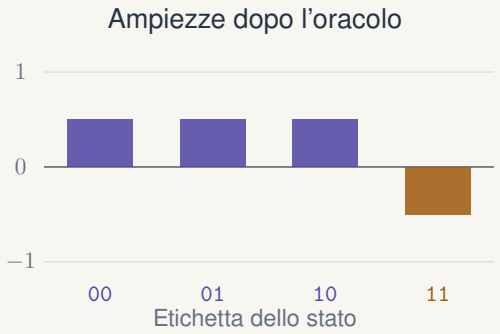
$$U_f\left(\sum_x a_x|x\rangle\right)|-\rangle = \left(\sum_x (-1)^{f(x)} a_x|x\rangle\right)|-\rangle$$

Richiami La linearità trasforma ogni termine della somma; un fattore ausiliario identico in tutti i termini può essere raccolto fuori dalla somma.

L’oracolo che cambia il segno delle soluzioni

Indichiamo con O_f l’azione ottenuta sul solo registro dei candidati: ogni stato di base $|x\rangle$ riceve il fattore $(-1)^{f(x)}$. Chiamiamo O_f oracolo di fase per distinguerlo dall’oracolo XOR U_f, che nella sua definizione include anche il qubit ausiliario.

Nel caso su quattro candidati, con unica soluzione 11, le ampiezze uniformi $(1/2, 1/2, 1/2, 1/2)$ diventano $(1/2, 1/2, 1/2, -1/2)$. Solo l’ultima cambia segno. Questa marcatura non elimina le altre componenti e non porta ancora tutta la probabilità sulla soluzione. Prepara una differenza relativa che un’ulteriore operazione potrà trasformare in una diversa distribuzione degli esiti.



$$O_f|x\rangle = (-1)^{f(x)}|x\rangle$$

La barra di 11 passa sotto lo zero perché la sua ampiezza cambia segno. Le altezze in valore assoluto rimangono tutte uguali a $1/2$.

Richiami U_f aggiorna un ausiliario con XOR; O_f indica il cambio di fase indotto sul registro dei candidati.

Le probabilità subito dopo la marcatura

Dopo la marcatura di fase dell'esempio, i moduli quadrati delle quattro ampiezze sono ancora tutti $1/4$. Se misuriamo subito il registro, troviamo la soluzione 11 con la stessa probabilità iniziale. La marcatura ha modificato lo stato, ma la base di misura usata non rivela direttamente la differenza di segno.

Il fenomeno coincide con quello già osservato per $|+\rangle$ e $|-\rangle$: avevano probabilità immediate uguali, ma Hadamard le trasformava in risultati differenti. Non possiamo però applicare una qualunque trasformazione e aspettarci automaticamente un vantaggio. Occorre scegliere un'operazione che sfrutti la particolare disposizione dei segni generata dall'oracolo, senza conoscere in anticipo quale etichetta sia stata marcata. Nel capitolo successivo la costruiremo esplicitamente su quattro candidati.

$$\left|(-1)^{f(x)}a_x\right|^2 = |a_x|^2$$

Richiami Il segno di un'ampiezza non modifica il suo modulo quadro, ma può influire sull'interferenza successiva.

Due marcature consecutive si annullano

Applichiamo due volte l'oracolo di fase senza inserire altre operazioni. Un candidato non valido riceve due fattori 1 e rimane invariato. Una soluzione riceve due fattori -1, il cui prodotto è $+1$. Tutte le ampiezze tornano quindi ai valori iniziali.

La relazione $O_f^2 = I$ mostra che non basta ripetere la marcatura per aumentare la probabilità di successo. Nel caso uniforme su quattro etichette, una marcatura lascia le probabilità a $1/4$ e due marcature ripristinano perfino lo stato iniziale. Grover alternerà perciò la marcatura con una seconda trasformazione, comune a tutti i candidati. La sequenza avrà un effetto diverso dalla ripetizione isolata dell'oracolo, perché la seconda operazione combinerà le ampiezze prima della marcatura successiva.

$$O_f^2|x\rangle = (-1)^{2f(x)}|x\rangle = |x\rangle$$

Richiami I lascia lo stato invariato; $(-1)^2 = 1$ annulla il cambio di segno applicato due volte.

La marcatura di uno stato non uniforme

Prendiamo un qubit nello stato $(3/5)|0\rangle + (4/5)|1\rangle$ e un predicato che vale 1 soltanto su 1. Con un ausiliario in $|-\rangle$, l'oracolo XOR lascia la prima componente invariata e cambia il segno della seconda. Il registro diventa quindi $(3/5)|0\rangle - (4/5)|1\rangle$.

Le probabilità immediate rimangono $9/25$ e $16/25$. Se applichiamo poi Hadamard, le nuove ampiezze sono $-1/(5\sqrt{2})$ per 0 e $7/(5\sqrt{2})$ per 1. Le probabilità diventano $1/50$ e $49/50$, che sommano a 1. Questo esercizio separa i due momenti: il predicato modifica una fase relativa, poi una porta che combina le componenti rende osservabile quella differenza. L'esempio illustra il meccanismo, senza sostituire la costruzione generale di Grover.

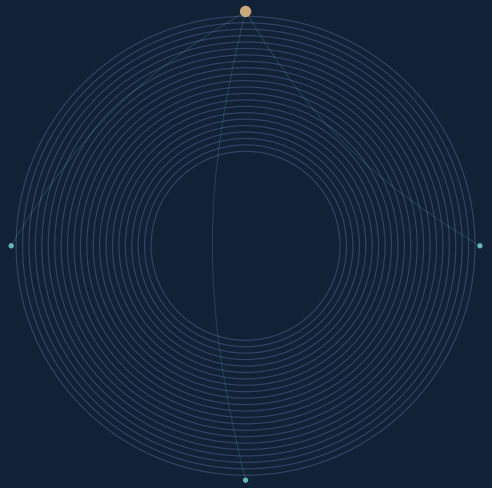
$$H\left(\frac{3}{5}|0\rangle - \frac{4}{5}|1\rangle\right) = \frac{-|0\rangle + 7|1\rangle}{5\sqrt{2}}$$

Richiami Hadamard trasforma le ampiezze (a, b) in $((a+b)/\sqrt{2}, (a-b)/\sqrt{2})$.

CAPITOLO 04

Quattro candidati

Un esempio completo dell’iterazione

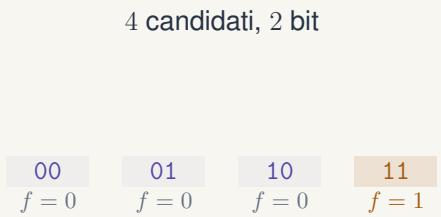


L’algoritmo di Grover

Una soluzione fra quattro candidati

Consideriamo un registro di due qubit. Le quattro etichette possibili sono 00, 01, 10 e 11. Il verificatore dell'esempio restituisce 1 soltanto per 11, che chiameremo bersaglio. L'algoritmo può interrogare il verificatore, mentre la nostra analisi conosce la risposta per poter controllare ogni passaggio.

Questa scelta permette di scrivere tutte le ampiezze, senza abbreviazioni. Partiremo da quattro candidati equiprobabili e seguiremo un'iterazione completa: prima l'oracolo cambia il segno dell'ampiezza del bersaglio, poi una seconda operazione modifica le ampiezze in modo che la misura restituisca 11. La certezza che otterremo riguarda questo piccolo caso ideale e non ogni dimensione del problema.



Il colore distingue il candidato accettato dal verificatore. Le posizioni nel disegno servono a leggere le etichette, non forniscono informazioni all'algoritmo.

Richiami $N = 2^n$ conta le etichette di un registro di n qubit. Qui $n = 2$ e $N = 4$. Una soluzione soddisfa $f(x) = 1$.

Le quattro ampiezze iniziali

Prepariamo lo stato uniforme applicando una porta Hadamard a ciascun qubit inizialmente in $|0\rangle$. Ogni etichetta riceve ampiezza $1/2$. La probabilità associata a ciascuna è quindi $(1/2)^2 = 1/4$, e le quattro probabilità sommano a 1.

Indicheremo questo stato con $|s\rangle$. Quando scriviamo un vettore di quattro coordinate, manterremo sempre l'ordine $00, 01, 10, 11$. La quarta coordinata riguarda dunque il bersaglio. Fissare l'ordine evita di confondere l'etichetta binaria di un candidato con la posizione della sua ampiezza nella lista. In questa preparazione nessuna etichetta è favorita rispetto alle altre.

$$|s\rangle = \tfrac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

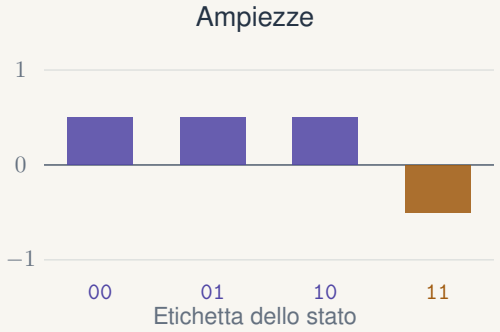
$$(a_{00}, a_{01}, a_{10}, a_{11}) = (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2})$$

Richiami La probabilità di misurare x è $|a_x|^2$. Uno stato normalizzato soddisfa $\sum_x |a_x|^2 = 1$.

Il segno del bersaglio dopo l'oracolo

L'oracolo di fase lascia invariate le componenti per cui $f(x) = 0$ e moltiplica per -1 quelle per cui $f(x) = 1$. Nel nostro esempio cambia soltanto l'ultima ampiezza, da $1/2$ a $-1/2$. Le altre tre restano positive.

Una misura eseguita ora darebbe ancora ciascuna etichetta con probabilità $1/4$, perché il quadrato di $-1/2$ è uguale a quello di $1/2$. L'informazione introdotta dall'oracolo riguarda il segno relativo fra le componenti. Occorre conservarla fino all'operazione successiva, che farà dipendere le nuove ampiezze da quelle dell'intero registro. La sola marcatura, seguita subito da una misura, non favorisce il bersaglio.



Le barre rappresentano ampiezze con segno. La barra del bersaglio scende sotto lo zero, ma il suo modulo resta uguale a quello delle altre.

Richiami $O_f|x\rangle = (-1)^{f(x)}|x\rangle$. Cambiare il segno di una sola componente introduce una fase relativa.

La media di una lista di ampiezze

La media aritmetica si ottiene sommando i valori di una lista e dividendo per il loro numero. Per esempio, la media di $2, 4, 6, 8$ è $20/4 = 5$. Il procedimento ammette anche valori negativi: la media di $2, 2, 2, -2$ è $4/4 = 1$. I segni partecipano alla somma.

Applichiamo la stessa definizione alle ampiezze. Se il registro ha N etichette e le loro ampiezze sono a_x, scriviamo $\bar{a}$ per la media. In questo capitolo tutte le ampiezze sono reali, perciò possiamo rappresentarle su una retta. La media non è una probabilità e non richiede di misurare le singole componenti: servirà a descrivere algebricamente una trasformazione del registro.

$$\bar{a} = \frac{1}{N} \sum_{x=0}^{N-1} a_x$$

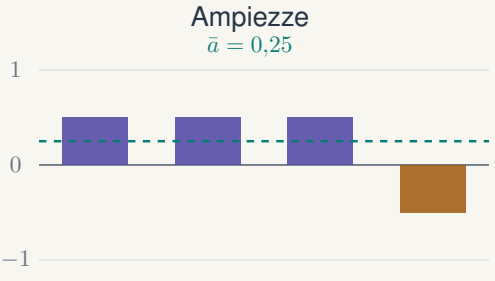
$$\frac{2 + 2 + 2 - 2}{4} = 1$$

Richiami a_x è l'ampiezza della componente $|x\rangle$, mentre N è il numero delle componenti. La somma delle probabilità vale 1; la somma delle ampiezze può avere altri valori.

La media dopo la marcatura

Dopo l'oracolo, le quattro ampiezze sono $1/2, 1/2, 1/2, -1/2$. La loro somma vale 1, perché le prime due componenti contribuiscono 1 e le ultime due si cancellano. Dividendo per quattro otteniamo la media $\bar{a} = 1/4$.

La media iniziale era $1/2$, dato che tutte le componenti avevano quel valore. La marcatura l'ha quindi abbassata senza modificare alcuna probabilità di misura. Rispetto alla nuova media, le tre componenti positive si trovano $1/4$ più in alto, mentre quella del bersaglio si trova $3/4$ più in basso. Queste distanze determineranno l'effetto dell'operazione seguente.



$$\bar{a} = \frac{\frac{1}{2} + \frac{1}{2} + \frac{1}{2} - \frac{1}{2}}{4} = \frac{1}{4}$$

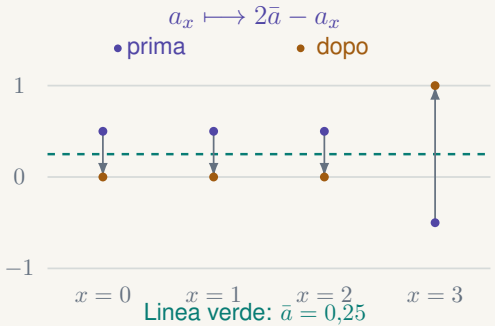
La linea verde mostra la media delle quattro ampiezze dopo l'oracolo. Il bersaglio e le altre componenti si trovano a distanze diverse da questa linea.

Richiami La media $\bar{a}$ si ricalcola sulle ampiezze dello stato corrente. Un segno negativo conta nella media, anche se non cambia il modulo quadro.

La riflessione rispetto alla media

Per riflettere un numero a rispetto al punto $\bar{a}$, conserviamo la distanza da quel punto e la portiamo sul lato opposto. Il risultato è $\bar{a} + (\bar{a} - a) = 2\bar{a} - a$. Con media $1/4$, il valore $1/2$ diventa zero, mentre il valore $-1/2$ diventa 1.

La seconda operazione di Grover applica questa regola a tutte le ampiezze ed è chiamata diffusore. La stessa formula agisce sul bersaglio e sugli altri candidati. Non riceve come dato l'etichetta cercata: la differenza fra le componenti è già stata prodotta dall'oracolo. Nel capitolo sulle riflessioni dimostreremo che questa trasformazione è lineare, reversibile e conserva la normalizzazione.



Per ogni componente, il valore prima e quello dopo la riflessione hanno la stessa distanza dalla media. Il valore negativo del bersaglio raggiunge 1.

Richiami La riflessione del valore a_x rispetto a $\bar{a}$ dà $a'_x = 2\bar{a} - a_x$. Nel caso corrente $\bar{a} = 1/4$.

Il calcolo di tutte le nuove ampiezze

Usiamo la media appena trovata, senza modificarla durante il calcolo delle singole componenti. Per ciascuno dei tre candidati non accettati dal verificatore otteniamo $2(1/4) - 1/2 = 0$. Per il bersaglio otteniamo invece $2(1/4) - (-1/2) = 1$. Il doppio segno negativo produce una somma.

Le quattro sostituzioni descrivono un'unica operazione sullo stato. Non sono un ciclo che aggiorna una coordinata e ricalcola la media prima di passare alla successiva. Il vettore risultante ha soltanto la quarta componente non nulla, perciò coincide con lo stato di base $|11\rangle$. La normalizzazione è preservata: tre probabilità sono zero e la quarta vale 1.

$$a'_{00} = a'_{01} = a'_{10} = 2 \cdot \tfrac{1}{4} - \tfrac{1}{2} = 0$$

$$a'_{11} = 2 \cdot \tfrac{1}{4} + \tfrac{1}{2} = 1$$

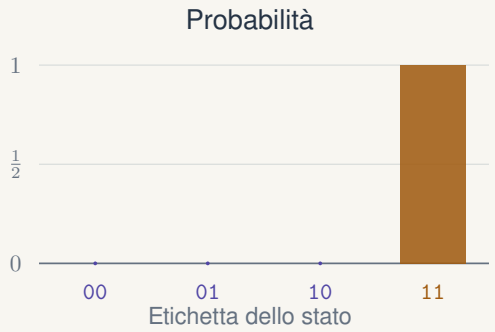
$$(a'_{00}, a'_{01}, a'_{10}, a'_{11}) = (0, 0, 0, 1)$$

Richiami La media del diffusore si riferisce al vettore in ingresso all'operazione. $|11\rangle$ ha ampiezza 1 per l'etichetta 11 e zero per le altre.

La misura dopo un’iterazione

La misura dello stato $|11\rangle$ restituisce 11 con probabilità 1. Una chiamata all’oracolo di fase, seguita dal diffusore, è bastata a trasformare la distribuzione uniforme in una distribuzione concentrata sul bersaglio. Un successivo controllo classico del predicato conferma che $f(11) = 1$.

Il grafico riporta probabilità teoriche del circuito ideale. Non rappresenta una serie di esecuzioni su un dispositivo reale, sul quale errori di preparazione, di porta o di misura potrebbero produrre risultati diversi. Qui la certezza segue direttamente dalle ampiezze calcolate. Nei problemi più grandi dovremo invece scegliere il numero di iterazioni in modo da ottenere una probabilità elevata, generalmente inferiore a 1.



Le barre sono ora probabilità, tutte non negative. Il bersaglio ha probabilità 1, mentre gli altri tre candidati hanno probabilità zero.

Richiami La misura legge un’etichetta con probabilità uguale al modulo quadro della sua ampiezza. La verifica classica accetta il risultato se $f(x) = 1$.

Un diffusore uguale per ogni bersaglio

Indichiamo con D il diffusore e con O_f l'oracolo. La formula del diffusore usa soltanto il numero N delle componenti e le loro ampiezze. Se il verificatore marcasse 01 anziché 11, cambierebbe la posizione dell'ampiezza negativa, ma la media dopo l'oracolo resterebbe $1/4$. La stessa operazione D porterebbe allora l'ampiezza di 01 a 1.

Questa separazione è essenziale per interpretare l'esempio. Le informazioni sul problema entrano attraverso f e il suo oracolo. Il diffusore non contiene una seconda copia della soluzione. La scrittura mediante la media è una descrizione matematica del suo effetto, non una richiesta di leggere le ampiezze durante l'esecuzione. Più avanti costruiremo il circuito che realizza questa trasformazione.

$$a'_x = \frac{2}{N} \sum_{y=0}^{N-1} a_y - a_x$$

Richiami O_f dipende dal predicato da verificare. D applica la stessa regola a ogni componente del registro.

L’ordine delle due operazioni

Una iterazione di Grover comprende prima l’oracolo e poi il diffusore. Scriviamo il loro prodotto come $G = DO_f$. Gli operatori agiscono sul vettore alla loro destra, quindi $G|s\rangle = D(O_f|s\rangle)$: nella formula il fattore più a destra si applica per primo.

Fra queste due operazioni non inseriamo una misura del registro dei candidati. La trasformazione deve conservare le relazioni fra le componenti che l’oracolo ha appena modificato. La misura è prevista dopo il numero scelto di iterazioni. Scriviamo G^k per indicare k applicazioni consecutive dello stesso blocco. Nel caso su quattro candidati basta una iterazione, ma questa notazione descrive anche due o più applicazioni.

$$G = DO_f, \qquad G|s\rangle = D(O_f|s\rangle) = |11\rangle$$

$$G^2|s\rangle = DO_f DO_f|s\rangle$$

Richiami Nel prodotto $AB|\psi\rangle$, l’operatore B agisce prima di A. G^k indica k iterazioni consecutive.

Il risultato con le operazioni invertite

Possiamo verificare l'importanza dell'ordine invertendo le due operazioni. Nello stato uniforme tutte le ampiezze valgono $1/2$, e anche la loro media vale $1/2$. Il diffusore lascia quindi ogni componente uguale a se stessa, perché $2(1/2) - 1/2 = 1/2$.

Se applichiamo l'oracolo dopo questa operazione, otteniamo soltanto il cambio di segno del bersaglio. Le probabilità restano tutte uguali a $1/4$. Perciò $O_f D|s\rangle$ e $DO_f|s\rangle$ producono distribuzioni di misura diverse. Il diffusore può convertire la marcatura di fase in una variazione delle probabilità solo quando riceve lo stato che porta già quella marcatura.

$$D|s\rangle = |s\rangle$$

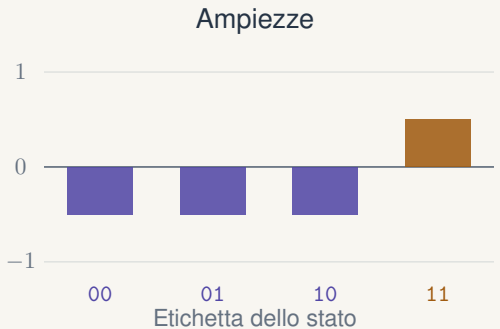
$$O_f D|s\rangle = \tfrac{1}{2}(|00\rangle + |01\rangle + |10\rangle - |11\rangle)$$

Richiami La riflessione di un valore uguale alla media lascia quel valore invariato. In generale, cambiare l'ordine di due operatori cambia il risultato.

Una seconda iterazione dopo il successo

Non misuriamo e applichiamo una seconda iterazione allo stato $|11\rangle$. L'oracolo cambia l'ampiezza del bersaglio da 1 a -1. Le altre tre sono zero, perciò la nuova media è $-1/4$. Il diffusore porta ciascuna ampiezza nulla a $-1/2$ e quella del bersaglio a $1/2$.

Ora tutte le probabilità valgono nuovamente $1/4$. L'informazione non è scomparsa dallo stato, che conserva segni relativi diversi, ma la misura non restituisce più il bersaglio con certezza. Questo calcolo mostra già che il numero di iterazioni fa parte dell'algoritmo. Aggiungere operazioni dopo aver raggiunto una probabilità elevata può peggiorare il risultato.



$$\bar{a} = -\tfrac{1}{4}, \qquad 2\bar{a} - 0 = -\tfrac{1}{2}$$

$$2\bar{a} - (-1) = \tfrac{1}{2}$$

Dopo due iterazioni le ampiezze hanno ancora moduli uguali. I tre candidati non marcati hanno segno negativo, mentre il bersaglio ha segno positivo.

Richiami Una nuova iterazione applica di nuovo O_f e poi D. La probabilità del bersaglio è il quadrato della sua ampiezza reale.

Il ritorno a uno stato uniforme

Alla terza iterazione partiamo dalle ampiezze $-1/2, -1/2, -1/2, 1/2$. L'oracolo cambia il segno dell'ultima, rendendole tutte uguali a $-1/2$. Il diffusore le lascia invariate, dato che coincidono con la loro media. Lo stato è dunque $-|s\rangle$.

Il segno meno moltiplica adesso l'intero vettore e costituisce una fase globale. Le probabilità coincidono con quelle della preparazione iniziale. Per linearità, un'ulteriore iterazione porta $-|s\rangle$ in $-|11\rangle$, che dà ancora successo certo. Indichiamo con P_k la probabilità di successo dopo k iterazioni. I valori alternano probabilità elevate e basse. Conviene fermarsi al primo successo certo, evitando le iterazioni successive.

$$G^3|s\rangle = -|s\rangle, \qquad G^4|s\rangle = -|11\rangle$$

$$P_0 = \tfrac{1}{4},\ P_1 = 1,\ P_2 = \tfrac{1}{4},\ P_3 = \tfrac{1}{4}$$

Richiami Una fase globale comune a tutte le ampiezze non cambia le probabilità. P_k indica la probabilità di successo dopo k iterazioni.

Esercizio su un bersaglio diverso

Consideriamo ancora quattro candidati, ma il verificatore accetta soltanto 01. Prepariamo lo stato uniforme e manteniamo l'ordine delle coordinate $00, 01, 10, 11$. Scrivere il vettore dopo l'oracolo, calcolarne la media e applicare il diffusore a ciascuna componente.

Il calcolo deve mostrare esplicitamente quale coordinata cambia segno e quale raggiunge ampiezza 1. Conviene controllare separatamente la somma delle ampiezze e la somma delle probabilità, perché non rappresentano la stessa quantità. Infine, determinare la probabilità di ottenere 01 se si misura subito dopo l'oracolo e quella ottenuta dopo l'iterazione completa. La pagina successiva svolge tutti i passaggi.

$$(a_{00}, a_{01}, a_{10}, a_{11}) = (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2})$$

Richiami O_f cambia segno soltanto alle componenti con $f(x) = 1$. Il diffusore usa $a'_x = 2\bar{a} - a_x$, con $\bar{a}$ calcolata prima degli aggiornamenti.

Soluzione dell'esercizio sul bersaglio 01

Il bersaglio occupa la seconda coordinata. Dopo l'oracolo il vettore è $(1/2, -1/2, 1/2, 1/2)$ e la somma delle ampiezze è 1. La media vale dunque $1/4$. La riflessione porta le tre ampiezze positive a zero e quella negativa a 1.

Prima del diffusore, la probabilità di misurare 01 è $|-1/2|^2 = 1/4$. Dopo il diffusore è $|1|^2 = 1$. Il risultato finale è lo stato $|01\rangle$. Il passaggio dal bersaglio 11 al bersaglio 01 ha soltanto permutato le coordinate coinvolte. Non abbiamo cambiato la regola del diffusore e non abbiamo modificato il numero di iterazioni.

$$(\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}) \xrightarrow{D} (0, 1, 0, 0)$$

$$P(01 \text{ dopo } O_f) = \tfrac{1}{4}, \qquad P(01 \text{ dopo } G) = 1$$

Richiami L'etichetta della soluzione determina quale componente viene marcata. La probabilità finale si calcola dalle ampiezze dopo tutte le operazioni previste.

La media delle probabilità non produce il diffusore

Se dopo l'oracolo prendessimo le quattro probabilità anziché le ampiezze, avremmo ancora la lista $1/4, 1/4, 1/4, 1/4$. La sua media è $1/4$. Riflettere questi quattro numeri rispetto alla loro media lascerebbe la lista invariata, senza aumentare la probabilità del bersaglio.

Il calcolo di Grover deve quindi conservare il segno delle ampiezze fino alla loro trasformazione. Il modulo quadro si applica quando vogliamo ricavare le probabilità di misura. Anticiparlo cancellerebbe la distinzione fra $1/2$ e $-1/2$ dalla quale dipende l'esempio. Nel caso su otto candidati questa differenza sarà ancora più evidente, perché alcune ampiezze non marcate diventeranno negative mentre le probabilità resteranno sempre non negative.

$$|\tfrac{1}{2}|^2 = |-\tfrac{1}{2}|^2 = \tfrac{1}{4}$$

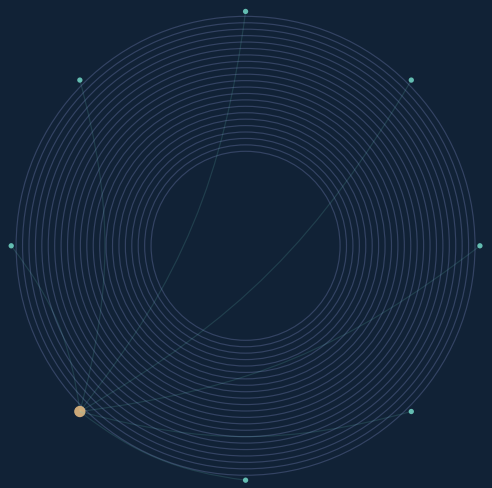
$\bar{a} = \frac{1}{4}$ dopo la marcatura del caso a quattro candidati

Richiami Le operazioni quantistiche trasformano ampiezze. Le probabilità si ottengono mediante il modulo quadro dopo aver calcolato lo stato.

CAPITOLO 05

Otto candidati

Il calcolo delle ampiezze dopo ogni passaggio



L’algoritmo di Grover

La ricerca del candidato 101

Passiamo a tre qubit, che rappresentano otto etichette da 000 a 111. Scegliamo come bersaglio 101, corrispondente all'intero 5 perché $1 \cdot 4 + 0 \cdot 2 + 1 \cdot 1 = 5$. Il verificatore accetta soltanto questa stringa. Per analizzare il circuito manteniamo l'ordine numerico delle etichette.

Il procedimento resta quello appena studiato: preparazione uniforme, oracolo e diffusore. Cambiano però il numero delle ampiezze che contribuiscono alla media e il risultato di ogni iterazione. Seguiremo tre iterazioni senza misure intermedie. I calcoli mostreranno che due iterazioni danno una probabilità maggiore di una, mentre la terza la riduce. La spiegazione geometrica di questa oscillazione verrà dopo i calcoli.

8 candidati, 3 bit

000 $f=0$	001 $f=0$	010 $f=0$	011 $f=0$
100 $f=0$	101 $f=1$	110 $f=0$	111 $f=0$

Ocra: $f(x) = 1$

La stringa accettata è 101. Le sette altre etichette hanno lo stesso ruolo nel calcolo, poiché l'oracolo non modifica il loro segno.

Richiami Con $n = 3$ qubit si hanno $N = 2^3 = 8$ candidati. Un'iterazione è $G = DO_f$: prima la marcatura, poi il diffusore.

La normalizzazione su otto componenti

Lo stato uniforme deve assegnare la stessa probabilità a ciascuno degli otto candidati. Ogni probabilità vale dunque $1/8$. Scegliendo ampiezze reali positive, ciascuna ampiezza è la radice quadrata di $1/8$, cioè $c = 1/\sqrt{8} = 1/(2\sqrt{2})$. Useremo la lettera c per abbreviare i calcoli.

Il valore numerico di c è circa $0{,}353553$. Le otto ampiezze non sommano a 1, ma a $8c = \sqrt{8}$. Sono i loro quadrati a sommare a 1. Questa distinzione conta nel diffusore, la cui media usa la somma delle ampiezze. Prima di interrogare l'oracolo, la probabilità di ottenere il bersaglio è soltanto $1/8$, pari al $12{,}5\%$.

$$|s\rangle = c\sum_{x=0}^{7}|x\rangle, \qquad c = \frac{1}{\sqrt{8}}$$

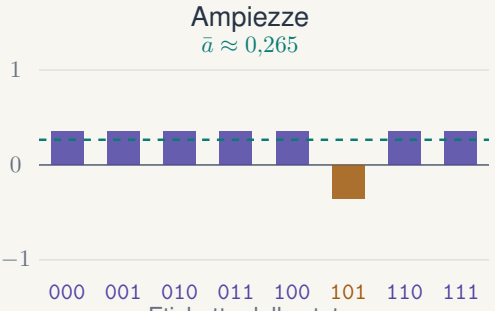
$$\sum_{x=0}^{7}|c|^2 = 8\cdot\frac{1}{8} = 1$$

Richiami Lo stato uniforme ha ampiezza $1/\sqrt{N}$ per ciascuna etichetta. La probabilità iniziale del singolo bersaglio è $1/N$.

La media alla prima marcatura

La prima applicazione dell'oracolo rende negativa soltanto l'ampiezza di 101. Abbiamo quindi sette componenti uguali a c e una uguale a $-c$. La somma è $7c - c = 6c$, e dividendo per otto otteniamo $\bar{a} = 3c/4$.

La media è positiva, ma inferiore all'ampiezza iniziale. Le sette componenti non marcate si trovano leggermente sopra di essa, mentre il bersaglio si trova sotto lo zero. Il diffusore userà questa media per tutte le componenti. Come nel caso precedente, le probabilità immediatamente dopo l'oracolo restano uguali a $1/8$: il cambiamento osservabile comparirà soltanto dopo la trasformazione delle ampiezze.



$$\bar{a} = \frac{7c - c}{8} = \frac{3c}{4}$$

La linea verde è la media delle otto ampiezze dopo la marcatura. La sola barra negativa identifica la componente del bersaglio nell'analisi dell'esempio.

Richiami $c = 1/\sqrt{8}$ è l'ampiezza uniforme iniziale. Il segno del bersaglio cambia a ogni applicazione di O_f.

Le ampiezze dopo la prima iterazione

Per il bersaglio, la riflessione rispetto alla media dà $2(3c/4) - (-c) = 5c/2$. Per ciascuno degli altri candidati dà invece $2(3c/4) - c = c/2$. L'ampiezza marcata aumenta, mentre tutte le altre diminuiscono mantenendo lo stesso valore fra loro.

Sostituendo $c = 1/(2\sqrt{2})$, otteniamo $5/(4\sqrt{2})$ per il bersaglio e $1/(4\sqrt{2})$ per ciascun altro candidato. Il valore maggiore è circa $0{,}883883$, quindi rimane inferiore a 1. Non abbiamo ancora uno stato di base come nel caso su quattro etichette. Per conoscere la probabilità di successo dobbiamo elevare al quadrato queste ampiezze.

$$a_{101} = \frac{5c}{2} = \frac{5}{4\sqrt{2}}$$

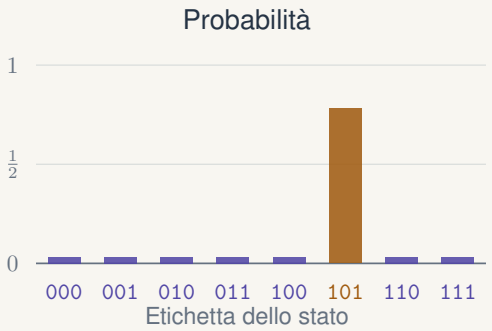
$$a_x = \frac{c}{2} = \frac{1}{4\sqrt{2}} \quad (x \neq 101)$$

Richiami Il diffusore trasforma ogni ampiezza mediante $a'_x = 2\bar{a} - a_x$. La media usata in questa iterazione è $\bar{a} = 3c/4$.

La probabilità dopo la prima iterazione

Il quadrato dell'ampiezza del bersaglio vale $25/32$, ossia $0{,}78125$. Ciascuno degli altri sette candidati ha probabilità $1/32$. La somma totale è $(25 + 7)/32 = 1$, come richiede la normalizzazione.

La probabilità di successo è passata dal $12{,}5\%$ iniziale al $78{,}125\%$. Tuttavia una misura può ancora restituire una stringa non accettata, con probabilità complessiva $7/32$. Il grafico separa il successo del singolo bersaglio dai sette esiti alternativi. In una singola esecuzione non si leggono tutte queste probabilità: si ottiene una sola etichetta, e il verificatore permette di stabilire se è una soluzione.



$$P_1 = \left(\frac{5}{4\sqrt{2}}\right)^2 = \frac{25}{32}$$

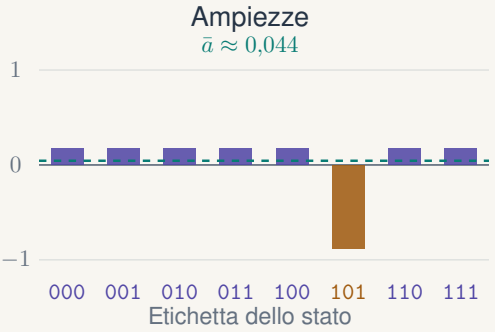
Il grafico mostra le probabilità teoriche dopo una iterazione. Il bersaglio concentra $25/32$ della probabilità, mentre i sette esiti rimanenti condividono $7/32$.

Richiami P_k è la probabilità del bersaglio dopo k iterazioni. Le probabilità di esiti distinti si sommano.

La media alla seconda marcatura

Proseguiamo senza misurare. Prima della seconda iterazione, il bersaglio ha ampiezza $5c/2$ e ciascuno degli altri candidati ha ampiezza $c/2$. L'oracolo rende negativa l'ampiezza del bersaglio, mentre lascia invariate le altre sette.

La somma diventa $-5c/2 + 7c/2 = c$. La nuova media è dunque $c/8$, molto più piccola della media $3c/4$ usata alla prima iterazione. Occorre ricalcolarla, perché lo stato in ingresso è cambiato. Riutilizzare la vecchia media darebbe una trasformazione diversa dal diffusore. I valori dopo questa marcatura restano normalizzati e hanno le stesse probabilità presenti prima dell'oracolo.



$$\bar{a} = \frac{-5c/2 + 7c/2}{8} = \frac{c}{8}$$

Alla seconda marcatura il bersaglio ha una grande ampiezza negativa. La media resta positiva, ma è vicina allo zero rispetto ai valori delle componenti.

Richiami Dopo la prima iterazione, il bersaglio ha ampiezza $5c/2$ e gli altri candidati $c/2$. $c = 1/\sqrt{8}$.

L’origine delle ampiezze negative

Con media $c/8$, il diffusore porta l’ampiezza del bersaglio a $2(c/8) + 5c/2 = 11c/4$. Per ciascuno degli altri candidati otteniamo invece $2(c/8) - c/2 = -c/4$. Le componenti non marcate attraversano quindi lo zero e diventano negative.

Il segno negativo è compatibile con uno stato valido. Le probabilità corrispondenti saranno positive, perché si calcolano elevando al quadrato il modulo delle ampiezze. Il segno conserva però un significato per le iterazioni successive: entra nella somma con cui calcoliamo la nuova media. Eliminandolo, potremmo riprodurre la distribuzione di questa misura ma non l’evoluzione successiva del circuito.

$$a_{101} = \frac{11c}{4} = \frac{11}{8\sqrt{2}}$$

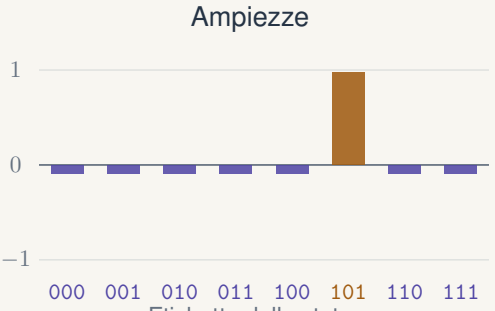
$$a_x = -\frac{c}{4} = -\frac{1}{8\sqrt{2}} \quad (x \neq 101)$$

Richiami Un’ampiezza negativa non è una probabilità negativa. La media corrente è $c/8$, con $c = 1/\sqrt{8}$.

La probabilità dopo due iterazioni

Il bersaglio ha ora probabilità $121/128$, cioè $0{,}9453125$. Ciascuno degli altri sette candidati ha probabilità $1/128$. Il controllo della normalizzazione dà $(121 + 7)/128 = 1$. La probabilità complessiva di fallimento è quindi $7/128$.

Le barre delle ampiezze mostrano una componente positiva vicina a 1 e sette piccole componenti negative. Se le trasformiamo in probabilità, tutte le barre diventano non negative, ma il bersaglio rimane dominante. Due iterazioni hanno migliorato il risultato rispetto a una. Non possiamo però dedurre che un'ulteriore iterazione lo migliori ancora: la successiva marcatura agirà su un vettore con segni diversi da quello iniziale.



$$P_2 = \frac{121}{128} = 0{,}9453125$$

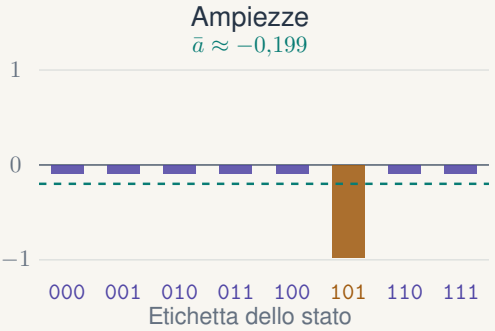
Le piccole barre negative contribuiscono alla normalizzazione attraverso i loro quadrati. Il loro segno conta, invece, quando si applica un'altra iterazione.

Richiami Il bersaglio ha ampiezza $11/(8\sqrt{2})$; gli altri candidati hanno ampiezza $-1/(8\sqrt{2})$. La probabilità di fallimento è $1 - P_2$.

Una media negativa alla terza marcatura

La terza applicazione dell'oracolo cambia l'ampiezza del bersaglio da $11c/4$ a $-11c/4$. Le altre sette erano già uguali a $-c/4$ e non cambiano. Tutte le ampiezze sono ora negative.

La loro somma è $-(11+7)c/4 = -9c/2$. Dividendo per otto otteniamo $\bar{a} = -9c/16$. La media può essere negativa perché riguarda ampiezze, non probabilità. Rispetto a questa media, il bersaglio è molto più in basso, mentre gli altri valori sono più vicini allo zero. La riflessione successiva aumenterà il modulo delle componenti non marcate, sottraendo probabilità al successo.



$$\bar{a} = \frac{-11c/4 - 7c/4}{8} = -\frac{9c}{16}$$

Tutte le ampiezze sono sotto lo zero dopo la terza marcatura. La linea della media è anch'essa negativa e serve da riferimento per la riflessione.

Richiami $c = 1/\sqrt{8}$. Dopo due iterazioni, le sette ampiezze non marcate valgono $-c/4$. L'oracolo cambia soltanto il segno della componente accettata.

Le ampiezze dopo la terza iterazione

La riflessione del bersaglio rispetto alla media $-9c/16$ produce $-9c/8 + 11c/4 = 13c/8$. Per ciascuno degli altri candidati produce invece $-9c/8 + c/4 = -7c/8$. Il bersaglio torna positivo, ma la sua ampiezza è più piccola di quella raggiunta dopo due iterazioni.

Anche il confronto fra i valori assoluti è istruttivo. Le componenti non marcate avevano modulo $c/4$ e ora hanno modulo $7c/8$, quindi contribuiscono molto di più alla probabilità totale. La trasformazione continua a conservare la norma del vettore. Non aggiunge probabilità al sistema: la distribuisce diversamente fra il bersaglio e gli esiti alternativi.

$$a_{101} = \frac{13c}{8} = \frac{13}{16\sqrt{2}}$$

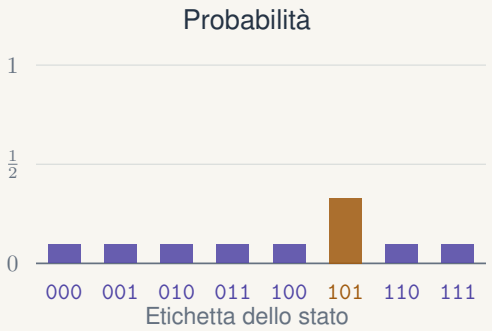
$$a_x = -\frac{7c}{8} = -\frac{7}{16\sqrt{2}} \quad (x \neq 101)$$

Richiami $c = 1/\sqrt{8}$ e la media dopo la terza marcatura è $-9c/16$. Il diffusore conserva la somma dei moduli quadri.

Il calo della probabilità di successo

Dopo tre iterazioni, il bersaglio ha probabilità $169/512$, circa $0{,}330078$. Ciascuno dei sette candidati non marcati ha probabilità $49/512$. La somma è $(169 + 7 \cdot 49)/512 = 1$, perciò il calo del successo corrisponde a un aumento della probabilità complessiva degli altri esiti.

Il valore ottenuto è ancora superiore a quello iniziale, ma molto inferiore al $121/128$ raggiunto dopo due iterazioni. In questo esempio la terza iterazione è quindi svantaggiosa se vogliamo misurare subito dopo. La sequenza ideale è reversibile e non ha un meccanismo che trattenga automaticamente lo stato quando la probabilità del bersaglio diventa grande. L'arresto deve essere deciso dall'algoritmo.



$$P_3 = \frac{169}{512} \approx 0{,}330078$$

Il bersaglio rimane il singolo esito più probabile, ma i sette esiti alternativi sommano ormai a una probabilità maggiore della sua.

Richiami $P_2 = 121/128$ era la probabilità dopo due iterazioni. Una maggiore quantità di iterazioni non implica una maggiore probabilità finale.

Il confronto fra le prime iterazioni

Raccogliamo le probabilità ottenute con zero, una, due e tre iterazioni. Il numero zero indica la misura immediata dello stato uniforme. Gli altri valori derivano dai vettori calcolati, non da una stima statistica. Fra queste quattro scelte, due iterazioni offrono la probabilità maggiore.

La tabella consente anche di distinguere due decisioni diverse: quante iterazioni eseguire in un singolo tentativo e quanti tentativi ripetere se la verifica finale fallisce. La prima decisione controlla la distribuzione prima della misura. La seconda riparte da una nuova preparazione dopo aver ottenuto un risultato classico. In seguito ricaveremo una formula che permette di scegliere il primo arresto favorevole senza enumerare tutti i vettori.

k	P_k	**Percentuale**
0	$1/8$	$12{,}5\%$
1	$25/32$	$78{,}125\%$
2	$121/128$	$94{,}53125\%$
3	$169/512$	$33{,}0078\%$

Le percentuali esprimono le stesse probabilità delle frazioni. L'ultima percentuale è arrotondata, mentre le frazioni riportate sono esatte.

Richiami Il conteggio k riguarda le iterazioni $G = DO_f$ prima della misura. Le probabilità qui si riferiscono a $N = 8$ e a una sola soluzione.

L’oscillazione nelle iterazioni successive

Continuando ad applicare lo stesso oracolo e lo stesso diffusore, la probabilità del bersaglio torna ad aumentare e poi diminuisce ancora. Il grafico mostra i valori ottenuti proseguendo il medesimo calcolo. Le iterazioni effettivamente eseguibili sono indicate dai punti con indice intero.

Non è necessario attendere un picco successivo quando un primo picco fornisce già un’elevata probabilità. In questo caso il primo arresto conveniente è dopo due iterazioni. La forma oscillante verrà spiegata come effetto di una rotazione in un piano dello spazio degli stati. Per ora il grafico riassume un fatto già visibile nei conti: la crescita del successo ha una durata limitata e richiede una scelta dell’arresto.



I punti rappresentano conteggi interi di iterazioni. L’andamento della probabilità mostra che continuare dopo un picco può ridurre sensibilmente la probabilità di successo.

Richiami k conta le applicazioni del blocco DO_f. La probabilità è compresa fra 0 e 1, anche quando alcune ampiezze sono negative.

Lo stato prodotto da una misura intermedia

Consideriamo una modifica del procedimento: subito dopo la prima marcatura misuriamo il registro, otteniamo un'etichetta x e poi applichiamo il diffusore. La misura lascia lo stato di base $|x\rangle$, a meno di una fase globale. Le sue ampiezze sono una uguale a 1 e sette uguali a zero.

La media di queste ampiezze è $1/8$. Il diffusore porta la componente x a $2/8 - 1 = -3/4$, mentre porta ciascuna delle altre a $1/4$. Se la misura intermedia aveva trovato il bersaglio, il successo finale ha probabilità $9/16$. Se aveva trovato un candidato diverso, il successo finale ha probabilità $1/16$. Scriviamo $P(A \mid B)$ per indicare la probabilità di A sotto la condizione B. Il diffusore riceve stati diversi da quello coerente usato in Grover.

$$a'_x = -\frac{3}{4}, \qquad a'_y = \frac{1}{4} \quad (y \neq x)$$

$$P(101 \mid x = 101) = \frac{9}{16}, \quad P(101 \mid x \neq 101) = \frac{1}{16}$$

Richiami Dopo una misura con esito x, il registro è nello stato $|x\rangle$. La scrittura $P(A \mid B)$ indica la probabilità di A sotto la condizione B.

La probabilità con la misura intermedia

Alla prima marcatura tutte le etichette hanno ancora probabilità $1/8$. La misura intermedia trova quindi il bersaglio con probabilità $1/8$ e un altro candidato con probabilità $7/8$. Per ottenere la probabilità finale dobbiamo pesare i due casi con queste probabilità.

Il risultato è $(1/8)(9/16) + (7/8)(1/16) = 1/8$. Con la misura intermedia non otteniamo l'amplificazione a $25/32$ della prima iterazione coerente. Questo confronto riguarda proprio la procedura descritta: misurare e applicare lo stesso diffusore, senza un'altra strategia condizionata all'esito. Illustra perché il progetto dell'algoritmo conserva le relazioni di fase fino al momento previsto per la misura finale.

$$P_{\text{finale}} = \frac{1}{8} \cdot \frac{9}{16} + \frac{7}{8} \cdot \frac{1}{16} = \frac{1}{8}$$

$$P_{\text{Grover},1} = \frac{25}{32}$$

Richiami Per combinare casi alternativi, si sommano le probabilità dei casi moltiplicate per le probabilità condizionate. Le due probabilità condizionate sono $9/16$ e $1/16$.

La verifica del risultato e un nuovo tentativo

Scegliamo di misurare dopo due iterazioni. La probabilità di ottenere il bersaglio è $121/128$, ma la verifica resta necessaria: applichiamo classicamente f all'etichetta osservata. Se il risultato è 0, prepariamo un nuovo registro uniforme e ripetiamo le due iterazioni.

Nel modello ideale, tentativi preparati da capo sono indipendenti. Ciascuno fallisce con probabilità $7/128$. Due tentativi falliscono entrambi con probabilità $(7/128)^2 = 49/16384$, quindi la probabilità di almeno un successo è $16335/16384$, circa $99{,}70\%$. Non stiamo aggiungendo altre iterazioni al registro già misurato. Il riavvio riproduce la stessa distribuzione favorevole da una nuova preparazione.

$$P(\text{due fallimenti}) = \left(\frac{7}{128}\right)^2 = \frac{49}{16384}$$

$$P(\text{almeno un successo}) = \frac{16335}{16384}$$

Richiami La verifica accetta soltanto etichette con $f(x) = 1$. Per eventi indipendenti, la probabilità che avvengano entrambi è il prodotto delle probabilità.

Esercizio su sedici candidati

Un registro di quattro qubit rappresenta sedici etichette. Supponiamo che una sola sia accettata dal verificatore e prepariamo lo stato uniforme, con ampiezza $1/4$ su ciascuna etichetta. Eseguire una sola iterazione di Grover, indicando con a_t l'ampiezza finale del bersaglio e con a_u quella di un candidato non marcato.

Calcolare prima la somma delle ampiezze dopo l'oracolo e dividerla per sedici. Applicare poi la regola $2\bar{a} - a_x$ al bersaglio e a uno qualsiasi degli altri quindici candidati. Determinare le probabilità corrispondenti e verificarne la somma. Infine, confrontare il successo di una iterazione con i risultati per quattro e otto candidati. Il confronto riguarda una sola iterazione, non la migliore scelta dell'arresto per ogni dimensione.

$$N = 16, \qquad a_x = \frac{1}{4} \quad \text{prima dell'oracolo}$$

Richiami Una sola ampiezza cambia segno quando esiste una sola soluzione. Dopo il diffusore, la normalizzazione richiede $|a_t|^2 + 15|a_u|^2 = 1$.

Soluzione dell'esercizio su sedici candidati

Dopo l'oracolo ci sono quindici ampiezze $1/4$ e una $-1/4$. La somma è $14/4 = 7/2$ e la media vale $7/32$. Il diffusore porta il bersaglio a $2(7/32) + 1/4 = 11/16$ e gli altri candidati a $2(7/32) - 1/4 = 3/16$.

La probabilità di successo è dunque $121/256$, circa $47{,}27\%$. Ciascun altro candidato ha probabilità $9/256$, e la somma è $(121 + 15 \cdot 9)/256 = 1$. Una iterazione migliora il valore iniziale $1/16$, ma non raggiunge i risultati ottenuti con una sola iterazione per quattro o otto candidati. All'aumentare dei candidati, a parità di una soluzione, occorre generalmente un maggior numero di iterazioni prima del primo picco favorevole.

$$\bar{a} = \frac{7}{32}, \quad a_t = \frac{11}{16}, \quad a_u = \frac{3}{16}$$

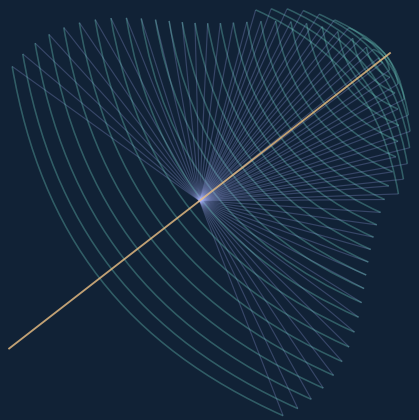
$$P_1 = \frac{121}{256}, \qquad \frac{121 + 15 \cdot 9}{256} = 1$$

Richiami a_t indica l'ampiezza del bersaglio e a_u quella di un candidato non marcato. Il calcolo riguarda un'unica iterazione, con $N = 16$ e una soluzione.

CAPITOLO 06

Riflessioni e sottospazi

Proiezioni, ortogonalità
e conservazione della norma



L’algoritmo di Grover

Le ampiezze come coordinate di un vettore

Gli esempi precedenti hanno trasformato liste di quattro o otto ampiezze. Una lista di questo tipo è un vettore, le cui coordinate si riferiscono a una base fissata. Per un qubit, le coordinate si riferiscono a $|0\rangle$ e $|1\rangle$; per due qubit, alle quattro etichette $00, 01, 10, 11$.

La descrizione geometrica userà operazioni su questi vettori. Nel piano reale, la lunghezza, chiamata norma, è la radice quadrata della somma dei quadrati delle coordinate. Il vettore $(3/5, 4/5)$ ha norma 1, perché $(3/5)^2 + (4/5)^2 = 1$. Se lo interpretiamo come stato di un qubit, le probabilità sono $9/25$ e $16/25$. Le coordinate sono le ampiezze, mentre i loro quadrati determinano le probabilità: la geometria riguarda anzitutto il vettore delle ampiezze.

$$|v\rangle = \frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle$$

$$\|v\| = \sqrt{\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = 1$$

Richiami La norma di un vettore è la radice quadrata della somma dei moduli quadri delle coordinate. Gli stati qui considerati sono rappresentati da vettori di norma 1.

Il prodotto scalare fra vettori reali

Per confrontare due vettori reali, moltiplichiamo le coordinate corrispondenti e sommiamo i prodotti. Nel piano, il prodotto scalare di $u = (u_0, u_1)$ e $v = (v_0, v_1)$ è $u_0v_0 + u_1v_1$. Il prodotto di un vettore con se stesso dà il quadrato della sua norma.

Consideriamo $u = (3/5, 4/5)$ e $v = (-4/5, 3/5)$. Entrambi hanno norma 1, ma il loro prodotto scalare è $-12/25 + 12/25 = 0$. Diciamo che sono ortogonali, cioè perpendicolari nel disegno geometrico. Due vettori normalizzati e ortogonali formano una base ortonormale del piano. Queste condizioni permettono di separare le componenti lungo direzioni indipendenti senza confondere i loro contributi alla norma.

$$\langle u, v \rangle = u_0v_0 + u_1v_1$$

$$\left\langle (\tfrac{3}{5}, \tfrac{4}{5}), (-\tfrac{4}{5}, \tfrac{3}{5}) \right\rangle = 0$$

Richiami Due vettori sono ortogonali quando il loro prodotto scalare è zero. La normalizzazione richiede prodotto scalare 1 con se stessi.

Il prodotto scalare con ampiezze complesse

Per ampiezze complesse, prima di moltiplicare le coordinate coniughiamo quelle del primo vettore. Il coniugato di $a + ib$, con a, b reali, è $a - ib$. Il simbolo i soddisfa $i^2 = -1$. Coniugare garantisce che il prodotto di un vettore con se stesso sia la somma dei moduli quadri, quindi un numero reale non negativo.

Prendiamo $u = (1, i)/\sqrt{2}$ e $v = (1, -i)/\sqrt{2}$. Il prodotto di u con se stesso è $(1 + (-i)i)/2 = 1$. Quello di u con v è $(1 + (-i)(-i))/2 = 0$. I due vettori sono dunque normalizzati e ortogonali, proprio come i due vettori reali della pagina precedente. Nei calcoli di Grover useremo ampiezze reali, ma gli operatori devono essere validi anche su stati complessi.

$$\langle u, v \rangle = \sum_x \overline{u_x}\, v_x$$

$$\langle u, u \rangle = 1, \qquad \langle u, v \rangle = 0$$

Richiami Il coniugato cambia il segno della parte immaginaria. $\overline{z}\, z = |z|^2$ per ogni numero complesso z.

Una direzione rappresentata da un vettore unitario

Una direzione non cambia se moltiplichiamo un vettore per un numero reale positivo. Il vettore $(3, 4)$ e il vettore $(3/5, 4/5)$ indicano la stessa direzione, ma hanno norme diverse: il primo ha norma 5, il secondo ha norma 1. Per normalizzare il primo, dividiamo ogni coordinata per la sua norma.

Useremo vettori di norma 1 per descrivere le direzioni rispetto alle quali proiettiamo o riflettiamo altri vettori. Questa scelta semplifica le formule: il prodotto scalare con un vettore unitario misura direttamente la componente lungo quella direzione. Se la direzione non fosse normalizzata, occorrerebbe un fattore correttivo. Per questo la normalizzazione compare sia negli stati sia nelle basi geometriche impiegate nell'analisi.

$$\|(3, 4)\| = \sqrt{3^2 + 4^2} = 5$$

$$u = \frac{(3, 4)}{5} = (\tfrac{3}{5}, \tfrac{4}{5}), \qquad \|u\| = 1$$

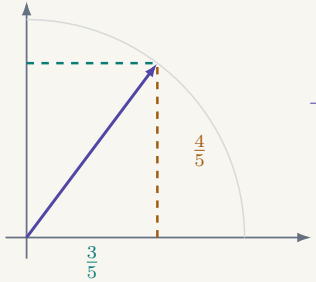
Richiami Per normalizzare un vettore non nullo, si divide il vettore per la sua norma. Un vettore unitario ha norma 1.

La componente lungo una direzione

Consideriamo il vettore unitario $v = (3/5, 4/5)$ e l'asse orizzontale, rappresentato dal vettore unitario $e_0 = (1, 0)$. La componente di v lungo l'asse è $3/5$, che coincide con il prodotto scalare $\langle e_0, v\rangle$. Il vettore proiettato sull'asse è quindi $(3/5)e_0 = (3/5, 0)$.

La parte restante è $(0, 4/5)$ ed è ortogonale all'asse. La somma delle due parti restituisce il vettore iniziale. La proiezione descrive dunque una componente di un vettore, non una probabilità di misura. Il quadrato della lunghezza della componente può diventare una probabilità quando la direzione rappresenta uno stato normalizzato della base di misura.

Proiezioni di un vettore unitario



$$\left(\tfrac{3}{5}\right)^2 + \left(\tfrac{4}{5}\right)^2 = 1$$

Il vettore ha norma 1. I segmenti orizzontale e verticale rappresentano le sue componenti, la cui somma vettoriale ricostruisce il vettore iniziale.

Richiami Il prodotto scalare con una direzione unitaria fornisce il coefficiente della proiezione. $v = (3/5, 0) + (0, 4/5)$.

La proiezione su una direzione obliqua

Il procedimento non dipende dall'orientamento degli assi. Scegliamo la direzione unitaria $u = (3/5, 4/5)$ e il vettore $v = (1, 0)$. Il prodotto scalare è $\langle u, v \rangle = 3/5$. Per ottenere il vettore parallelo a u, moltiplichiamo l'intero vettore u per questo coefficiente.

La proiezione vale quindi $(9/25, 12/25)$. Le sue coordinate non coincidono con quelle del vettore di partenza, perché la direzione su cui proiettiamo è obliqua. Il risultato resta però sulla retta generata da u: è esattamente $3/5$ volte quel vettore. Questa costruzione generale ci permetterà di riflettere lo stato rispetto alla direzione dello stato uniforme, che non coincide con una singola etichetta della base computazionale.

$$v_{\parallel} = u \langle u, v \rangle$$

$$v_{\parallel} = \frac{3}{5}(\frac{3}{5}, \frac{4}{5}) = (\frac{9}{25}, \frac{12}{25})$$

Richiami u deve avere norma 1 nella formula $v_{\parallel} = u\langle u, v \rangle$. Il simbolo $\parallel$ indica la componente parallela alla direzione scelta.

La parte perpendicolare alla proiezione

Sottraiamo dal vettore iniziale la sua proiezione. Nell'esempio $v = (1, 0)$ e $v_{\parallel} = (9/25, 12/25)$, quindi la parte restante è $v_{\perp} = (16/25, -12/25)$. La somma delle due parti restituisce esattamente $(1, 0)$.

Il prodotto scalare di $u = (3/5, 4/5)$ con questa parte vale $(48 - 48)/125 = 0$. Il simbolo $\perp$ ricorda proprio l'ortogonalità. Le lunghezze delle due componenti sono $3/5$ e $4/5$, e i loro quadrati sommano a 1. Abbiamo così separato il vettore in una parte lungo u e una parte perpendicolare, mantenendo esplicito il contributo di ciascuna alla norma.

$$v_{\perp} = v - v_{\parallel} = (\frac{16}{25}, -\frac{12}{25})$$

$$\langle u, v_{\perp} \rangle = 0, \qquad v = v_{\parallel} + v_{\perp}$$

$$\|v\|^2 = \|v_{\parallel}\|^2 + \|v_{\perp}\|^2 = \frac{9}{25} + \frac{16}{25}$$

Richiami La parte perpendicolare si ottiene sottraendo la proiezione dal vettore. Per componenti ortogonali si sommano i quadrati delle norme.

La notazione della proiezione

La notazione di Dirac distingue il vettore colonna $|u\rangle$ dal corrispondente vettore riga $\langle u|$, ottenuto trasponendo e coniugando le coordinate. Il prodotto $\langle u|v\rangle$ è uno scalare: moltiplica la riga associata a u per la colonna associata a v.

Il prodotto scritto nell'altro ordine, $|u\rangle\langle u|$, è invece un operatore. Quando agisce su $|v\rangle$, calcola prima il coefficiente $\langle u|v\rangle$ e poi lo moltiplica per $|u\rangle$. Per un vettore u normalizzato, questa è esattamente la proiezione appena costruita. L'ordine dei simboli distingue dunque un numero da una trasformazione che prende un vettore e restituisce un altro vettore.

$$P_u = |u\rangle\langle u|$$

$$P_u|v\rangle = |u\rangle\,\langle u|v\rangle = |v_\parallel\rangle$$

Richiami $\langle u|$ è la riga delle coordinate coniugate di $|u\rangle$. Il proiettore P_u restituisce la componente lungo u, se $\langle u|u\rangle = 1$.

La matrice di un proiettore

Possiamo verificare la notazione scrivendo una matrice esplicita. Per $u = (3/5, 4/5)$, il prodotto fra la colonna u e la riga u dà una matrice con elementi $u_i u_j$. Sulla diagonale compaiono $9/25$ e $16/25$, mentre gli elementi fuori diagonale valgono $12/25$.

Moltiplicando questa matrice per $v = (1, 0)$, otteniamo la sua prima colonna, cioè $(9/25, 12/25)$. È la stessa proiezione trovata mediante il prodotto scalare. Le due scritture descrivono una sola operazione: la formula con ket e bra mette in evidenza la direzione geometrica, mentre la matrice consente di calcolare direttamente ogni coordinata nella base scelta.

$$P_u = \frac{1}{25}\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}$$

$$P_u \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 9 \\ 12 \end{pmatrix}$$

Richiami Una matrice rappresenta un operatore rispetto a una base fissata. Il prodotto $|u\rangle\langle u|$ è una matrice, mentre $\langle u|u\rangle$ è un numero.

Proiettare due volte sulla stessa direzione

Una volta ottenuto un vettore parallelo a u, una seconda proiezione sulla stessa direzione lo lascia invariato. In simboli, $P_u^2 = P_u$. La proprietà segue anche dalla formula: fra i due prodotti $|u\rangle\langle u|$ compare il fattore $\langle u|u\rangle$, che vale 1 per normalizzazione.

La proiezione da sola non è reversibile e in generale non conserva la norma. Ogni vettore ortogonale a u viene mandato nel vettore zero. Qui è uno strumento algebrico per descrivere componenti, non una porta quantistica deterministica. Il diffusore userà il proiettore nell'espressione $2P_u - I$, dove I è l'identità. Questa operazione conserva anche la componente perpendicolare, cambiandole soltanto segno, e sarà reversibile.

$$P_u^2 = |u\rangle\langle u|u\rangle\langle u| = P_u$$

$$P_u|v_\perp\rangle = 0$$

Richiami Un proiettore su una direzione normalizzata soddisfa $P_u^2 = P_u$. I indica l'operatore identità, che lascia ogni vettore invariato.

La riflessione conserva una parte e inverte l'altra

Scomponiamo un vettore come $v = v_{\parallel} + v_{\perp}$ rispetto a una direzione unitaria u. La riflessione rispetto alla retta di u conserva la parte parallela e cambia segno a quella perpendicolare. Il risultato è quindi $v_{\parallel} - v_{\perp}$.

Poiché $v_{\perp} = v - v_{\parallel}$, lo stesso risultato si scrive $2v_{\parallel} - v$. Sostituendo $v_{\parallel} = P_u v$, otteniamo l'operatore $R_u = 2P_u - I$. La riflessione conserva la lunghezza: le parti rimangono ortogonali e cambiare un segno non modifica il quadrato della loro norma. Questa proprietà rende la riflessione adatta alla trasformazione di stati normalizzati.

$$R_u = 2P_u - I$$

$$R_u v = v_{\parallel} - v_{\perp}$$

$$\|R_u v\|^2 = \|v_{\parallel}\|^2 + \|v_{\perp}\|^2 = \|v\|^2$$

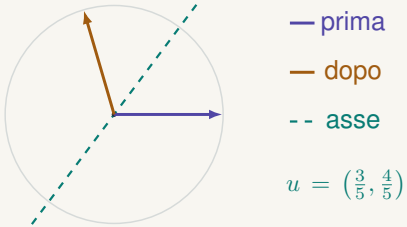
Richiami $P_u v$ è la componente di v parallela a u. La riflessione cambia il segno della componente perpendicolare, non la sua lunghezza.

Una riflessione calcolata nel piano

Riprendiamo la direzione $u = (3/5, 4/5)$ e il vettore $v = (1, 0)$. La proiezione vale $(9/25, 12/25)$, quindi la riflessione dà $2(9/25, 12/25) - (1, 0) = (-7/25, 24/25)$. Il risultato si trova dall'altro lato della retta di u.

Possiamo controllarne la norma senza usare gli angoli. La somma dei quadrati delle coordinate è $(49 + 576)/625 = 1$, uguale alla norma quadrata dell'ingresso. Il disegno rappresenta una riflessione ordinaria di vettori reali. Nel registro quantistico le coordinate sono ampiezze e lo spazio può avere molte dimensioni, ma la scomposizione in parti parallela e perpendicolare segue la stessa regola.

Riflessione rispetto a una retta



Il vettore iniziale è orizzontale. La retta obliqua indica la direzione di riflessione, rispetto alla quale i due vettori occupano posizioni simmetriche.

Richiami $R_u v = 2P_u v - v$. Il vettore proiettato è $(9/25, 12/25)$ e quello riflesso è $(-7/25, 24/25)$.

Due riflessioni uguali restituiscono l'ingresso

Applichiamo due volte la stessa riflessione. La parte parallela resta invariata in entrambi i passaggi, mentre la parte perpendicolare cambia segno due volte. Il risultato è il vettore iniziale. La riflessione è dunque la propria inversa.

La stessa conclusione emerge dal calcolo degli operatori. Poiché $P_u^2 = P_u$, il quadrato di $2P_u - I$ si riduce all'identità. Nell'esempio numerico, riflettere di nuovo $(-7/25, 24/25)$ rispetto alla retta di $u = (3/5, 4/5)$ restituisce $(1, 0)$. La reversibilità che mancava alla sola proiezione è stata recuperata conservando la parte perpendicolare anziché eliminarla.

$$R_u^2 = (2P_u - I)^2 = 4P_u^2 - 4P_u + I = I$$

$$R_u^{-1} = R_u$$

Richiami Il prodotto R_u^2 significa applicare due volte la stessa riflessione. Un'operazione è reversibile se ammette un'inversa che ricostruisce ogni ingresso.

L’operatore aggiunto

Per esprimere la conservazione del prodotto scalare usiamo l’operatore aggiunto, indicato dal simbolo $\dagger$. La matrice $A^\dagger$ si ottiene dalla matrice A scambiando righe e colonne e coniugando tutti gli elementi. Per matrici reali basta quindi la trasposizione.

Il piccolo esempio sotto mostra entrambe le operazioni: l’elemento i si sposta dall’angolo in alto a destra a quello in basso a sinistra e diventa $-i$. Questa matrice serve soltanto a illustrare la notazione. Per un vettore colonna, l’aggiunto produce proprio la riga coniugata che abbiamo chiamato bra. L’aggiunto di un prodotto inverte l’ordine dei fattori, proprietà che useremo per analizzare sequenze di porte.

$$A = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}, \qquad A^\dagger = \begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix}$$

$$(AB)^\dagger = B^\dagger A^\dagger$$

Richiami Coniugare cambia i in $-i$. Trasporre una matrice significa scambiare l’indice di riga con quello di colonna.

La condizione di unitarietà

Un operatore quadrato U è unitario quando $U^\dagger U = I$. In uno spazio di dimensione finita, il suo aggiunto è allora la sua inversa. Questa condizione garantisce che il prodotto scalare fra due vettori rimanga invariato dopo la trasformazione e, in particolare, che ogni norma sia conservata.

La porta Hadamard fornisce un esempio già incontrato. La sua matrice è reale e simmetrica, quindi $H^\dagger = H$. Il prodotto H^2 vale l'identità, perciò anche $H^\dagger H = I$. Le operazioni quantistiche ideali senza misura sono descritte da operatori unitari. Mostrare che il diffusore è unitario significa verificare che la regola sulle ampiezze può costituire un'operazione di questo tipo su qualunque ingresso, non soltanto sugli esempi calcolati.

$$U^\dagger U = I, \qquad \langle Uu | Uv \rangle = \langle u | v \rangle$$

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad H^\dagger H = I$$

Richiami † indica trasposizione e coniugazione. Un operatore unitario conserva prodotti scalari e normalizzazione.

La riflessione è un operatore unitario

Il proiettore $P_u = |u\rangle\langle u|$ coincide con il proprio aggiunto: $P_u^\dagger = P_u$. Nella matrice dell'esempio reale questo si vede dalla simmetria rispetto alla diagonale. Perciò anche la riflessione $R_u = 2P_u - I$ coincide con il proprio aggiunto.

Abbiamo già dimostrato che $R_u^2 = I$. Combinando le due proprietà otteniamo $R_u^\dagger R_u = I$, cioè l'unitarietà. La dimostrazione vale anche quando le coordinate sono complesse, purché il prodotto scalare includa la coniugazione e u sia normalizzato. La rappresentazione con frecce nel piano reale aiuta a intuire l'operazione, mentre il calcolo con proiettori ne stabilisce la validità su tutti gli stati del registro.

$$P_u^\dagger = P_u, \qquad R_u^\dagger = R_u$$

$$R_u^\dagger R_u = R_u^2 = I$$

Richiami $P_u^2 = P_u$ e $R_u = 2P_u - I$. La condizione di unitarietà è $R_u^\dagger R_u = I$.

La proiezione lungo lo stato uniforme

Scegliamo ora come direzione lo stato uniforme $|s\rangle = N^{-1/2}\sum_x |x\rangle$. Le sue N coordinate sono tutte uguali a $1/\sqrt{N}$. Per uno stato $|\psi\rangle = \sum_x a_x|x\rangle$, il prodotto scalare con $|s\rangle$ è quindi la somma delle ampiezze divisa per $\sqrt{N}$.

Moltiplichiamo questo numero per il vettore $|s\rangle$. Ogni coordinata del vettore proiettato riceve un altro fattore $1/\sqrt{N}$, perciò vale $(1/N)\sum_x a_x = \bar{a}$. La proiezione sulla direzione uniforme sostituisce dunque tutte le coordinate con la loro media. Questo collega la descrizione geometrica alla media usata nei calcoli precedenti, senza richiedere di leggere o misurare le ampiezze.

$$\langle s|\psi\rangle = \frac{1}{\sqrt{N}}\sum_x a_x = \sqrt{N}\,\bar{a}$$

$$P_s|\psi\rangle = |s\rangle\langle s|\psi\rangle = \sum_x \bar{a}|x\rangle$$

Richiami $\bar{a} = (1/N)\sum_x a_x$ è la media delle ampiezze. $P_s = |s\rangle\langle s|$ proietta sulla direzione dello stato uniforme.

Il diffusore come riflessione

La riflessione rispetto alla direzione dello stato uniforme è $D = 2P_s - I = 2|s\rangle\langle s| - I$. Poiché ogni coordinata di $P_s|\psi\rangle$ è $\bar{a}$, ogni coordinata di $D|\psi\rangle$ è $2\bar{a} - a_x$. Ritroviamo esattamente la regola applicata negli esempi.

Il diffusore conserva la media delle ampiezze: sommando le nuove coordinate otteniamo $2N\bar{a} - \sum_x a_x = N\bar{a}$. Le deviazioni dalla media cambiano invece segno. Negli esempi la media cambiava fra un'iterazione e la successiva a causa della nuova marcatura dell'oracolo. La riflessione del diffusore, da sola, conserva la propria media di ingresso ed è unitaria per la dimostrazione appena svolta.

$$D = 2|s\rangle\langle s| - I$$

$$D|\psi\rangle = \sum_x (2\bar{a} - a_x)|x\rangle$$

$$\frac{1}{N}\sum_x (2\bar{a} - a_x) = \bar{a}$$

Richiami Il diffusore riflette rispetto alla direzione $|s\rangle$. La formula $2\bar{a} - a_x$ vale anche per ampiezze complesse.

La matrice del diffusore su quattro candidati

Per $N = 4$, il vettore uniforme ha tutte le coordinate uguali a $1/2$. Il proiettore $|s\rangle\langle s|$ ha quindi tutti gli elementi uguali a $1/4$. Moltiplicando per 2 e sottraendo l'identità otteniamo una matrice con $-1/2$ sulla diagonale e $1/2$ fuori diagonale.

La riga associata a una componente sottrae metà della sua ampiezza e somma metà delle altre tre. È la stessa espressione $2\bar{a} - a_x$, perché $2\bar{a}$ contiene metà di tutte e quattro le ampiezze. La matrice è identica qualunque sia il bersaglio. Le sue colonne sono normalizzate e ortogonali, come deve accadere per una matrice unitaria.

$$D = \frac{1}{2}\begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$$

Richiami Nella base $00, 01, 10, 11$, lo stato uniforme è $(1, 1, 1, 1)/2$. L'identità ha elementi 1 sulla diagonale e zero altrove.

La scomposizione del primo esempio

Dopo l'oracolo del caso a quattro candidati, il vettore è $\psi = (1/2, 1/2, 1/2, -1/2)$. La sua media è $1/4$. La parte parallela allo stato uniforme è quindi $(1/4, 1/4, 1/4, 1/4)$, mentre la parte perpendicolare è $(1/4, 1/4, 1/4, -3/4)$.

Il diffusore conserva la prima parte e cambia segno alla seconda. Sottraendole otteniamo $(0, 0, 0, 1)$, cioè il bersaglio. Possiamo controllare l'ortogonalità osservando che la somma delle coordinate della seconda parte è zero: il suo prodotto scalare con il vettore uniforme è perciò nullo. La spiegazione geometrica ricostruisce così il calcolo numerico dell'inizio senza introdurre una diversa operazione.

$$\psi_{\parallel} = (\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4})$$

$$\psi_{\perp} = (\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, -\tfrac{3}{4})$$

$$D\psi = \psi_{\parallel} - \psi_{\perp} = (0, 0, 0, 1)$$

Richiami La componente parallela a $|s\rangle$ ha tutte le coordinate uguali alla media. Una componente ortogonale a $|s\rangle$ ha somma delle coordinate uguale a zero.

Le due azioni del diffusore

Sulla direzione uniforme, il diffusore agisce come l'identità: $D|s\rangle = |s\rangle$. Su ogni vettore ortogonale a $|s\rangle$ agisce invece cambiando il segno. Per esempio, nel caso a quattro coordinate, il vettore $w = (1, -1, 0, 0)/\sqrt{2}$ ha somma delle ampiezze zero e soddisfa $Dw = -w$.

Considerato da solo, il cambio di segno dell'intero vettore w è una fase globale. Consideriamo invece uno stato con ampiezze generiche α lungo $|s\rangle$ e β lungo $|w\rangle$. Il diffusore cambia il segno relativo delle due componenti e può così modificare le ampiezze della base computazionale. La distinzione fra fase globale e fase relativa riappare nella descrizione geometrica dell'interferenza.

$$D|s\rangle = |s\rangle, \qquad D|w\rangle = -|w\rangle$$

$$D(\alpha|s\rangle + \beta|w\rangle) = \alpha|s\rangle - \beta|w\rangle$$

Richiami $\langle s|w\rangle = 0$ significa che w è ortogonale allo stato uniforme. Una fase relativa riguarda il rapporto fra componenti diverse dello stesso stato.

La riflessione realizzata dall'oracolo

Con un solo bersaglio t, il proiettore $P_t = |t\rangle\langle t|$ seleziona la sua componente. L'oracolo di fase si scrive $O_f = I - 2P_t$: sottrae due volte la componente del bersaglio, trasformando a_t in $-a_t$, e lascia invariate tutte le altre.

Le combinazioni lineari delle direzioni non marcate formano un sottospazio, cioè un insieme di vettori chiuso rispetto a somma e moltiplicazione per scalari. L'oracolo lo conserva e cambia il segno della componente perpendicolare, diretta lungo $|t\rangle$. Per quattro candidati e bersaglio 11, le direzioni conservate sono $|00\rangle, |01\rangle, |10\rangle$. L'oracolo è una riflessione rispetto al loro sottospazio. Come il diffusore, coincide con il proprio inverso ed è unitario.

$$O_f = I - 2|t\rangle\langle t|$$

$$O_f|t\rangle = -|t\rangle, \qquad O_f|x\rangle = |x\rangle \quad (x \neq t)$$

Richiami Il proiettore $|t\rangle\langle t|$ conserva soltanto la componente del bersaglio. In questo caso c'è una sola soluzione, $f(t) = 1$.

La composizione delle due riflessioni

L'iterazione $G = DO_f$ compone due riflessioni rispetto a riferimenti diversi. Entrambe conservano la norma, quindi la composizione è unitaria. Tuttavia applicare due volte G non equivale in generale all'identità, perché la seconda iterazione ripete le operazioni nello stesso ordine anziché invertirlo.

L'inversa è $G^{-1} = O_f D$. Si verifica cancellando prima i due diffusori adiacenti e poi i due oracoli. La distinzione spiega perché un'iterazione può spostare lo stato e un'altra può spostarlo ancora, producendo l'oscillazione osservata. Nel caso con una soluzione, tutte le ampiezze non marcate restano uguali fra loro. L'evoluzione può quindi essere descritta con due componenti aggregate, che nel prossimo capitolo formeranno un piano.

$$G = DO_f, \qquad G^{-1} = O_f D$$

$$(O_f D)(DO_f) = O_f D^2 O_f = I$$

Richiami $D^2 = I$ e $O_f^2 = I$. Per invertire una sequenza di operazioni, si applicano le inverse in ordine opposto.

Esercizio sulla riflessione lungo una diagonale

Nel piano reale, consideriamo la direzione unitaria $u = (1,1)/\sqrt{2}$ e il vettore $v = (1,0)$. Calcolare il coefficiente $\langle u, v\rangle$, la proiezione $v_{\parallel}$ e la parte perpendicolare $v_{\perp}$. Applicare poi la riflessione $R_u = 2|u\rangle\langle u| - I$.

Il risultato può essere controllato anche con un disegno: la direzione di u è la diagonale che forma angoli uguali con i due assi. La riflessione deve quindi scambiare il ruolo della direzione orizzontale e di quella verticale. Dopo aver trovato il vettore finale, verificare che la sua norma sia 1 e che una seconda riflessione restituisca v. La soluzione seguente confronta il calcolo algebrico con questa previsione.

$$u = \frac{1}{\sqrt{2}}(1,1), \qquad v = (1,0)$$

Richiami $v_{\parallel} = u\langle u, v\rangle$ e $v_{\perp} = v - v_{\parallel}$. La riflessione è $R_u v = v_{\parallel} - v_{\perp}$.

Soluzione della riflessione lungo una diagonale

Il prodotto scalare vale $1/\sqrt{2}$. Moltiplicandolo per u otteniamo $v_{\|} = (1/2, 1/2)$. Sottraendo questa proiezione da $v = (1, 0)$ otteniamo $v_{\perp} = (1/2, -1/2)$. La riflessione conserva la prima parte e cambia segno alla seconda.

Il risultato è $(1/2, 1/2) - (1/2, -1/2) = (0, 1)$, di norma 1. La matrice della riflessione è $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ e una seconda applicazione restituisce $(1, 0)$. È la stessa matrice della porta X su un qubit. L'esempio mostra che una trasformazione già nota come scambio delle due ampiezze può anche essere letta come riflessione geometrica rispetto alla direzione uniforme del piano.

$$R_u = 2 \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - I = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

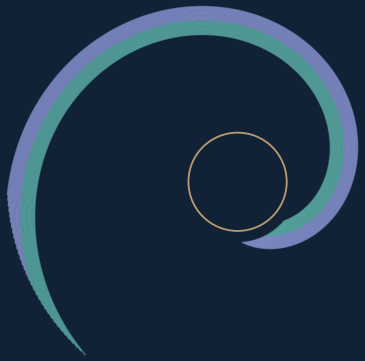
$$R_u(1, 0) = (0, 1), \qquad R_u(0, 1) = (1, 0)$$

Richiami La direzione uniforme su due coordinate è $(1, 1)/\sqrt{2}$. La porta X scambia le ampiezze di $|0\rangle$ e $|1\rangle$.

CAPITOLO 07

La rotazione di Grover

Il piano degli stati
e la probabilità di successo



L'algoritmo di Grover

Le soluzioni e gli altri candidati

Il predicato divide i candidati in due insiemi: quelli per cui $f(x) = 1$ sono soluzioni, quelli per cui $f(x) = 0$ non lo sono. Indichiamo con M il numero delle soluzioni e con $N - M$ il numero degli altri candidati. Per costruire la descrizione geometrica supponiamo inizialmente $0 < M < N$, così entrambi gli insiemi sono presenti.

Nell'esempio con otto candidati e bersaglio 101, il primo insieme contiene un solo elemento e il secondo ne contiene sette. Durante l'algoritmo le ampiezze dei sette elementi rimangono uguali fra loro. Possiamo raccoglierle in un'unica direzione dello spazio degli stati. Una seconda direzione descriverà le soluzioni. La riduzione a due coordinate dipende da questa simmetria, pur lasciando invariati gli otto possibili esiti della misura.

8 candidati, 3 bit

000 $f = 0$	001 $f = 0$	010 $f = 0$	011 $f = 0$
100 $f = 0$	101 $f = 1$	110 $f = 0$	111 $f = 0$

Ocra: $f(x) = 1$

Il colore distingue il bersaglio dagli altri sette candidati. Le posizioni nel disegno servono a elencare le stringhe, senza attribuire loro una distanza.

Richiami $N = 2^n$ è il numero dei candidati codificati da n qubit; M conta quelli che soddisfano $f(x) = 1$.

Lo stato uniforme delle soluzioni

Costruiamo un vettore sommando gli stati di base che corrispondono alle soluzioni. Se le soluzioni sono M, la somma contiene M termini ortogonali. Per ottenere un vettore di norma 1 assegniamo a ogni termine il coefficiente $1/\sqrt{M}$. Chiameremo questo stato $|g\rangle$, scegliendo la lettera dalla parola inglese good.

Con un solo bersaglio 101, la somma ha un termine e $|g\rangle = |101\rangle$. Con due soluzioni, per esempio 011 e 101, i coefficienti sono invece entrambi $1/\sqrt{2}$. Misurare questo secondo stato produce ciascuna soluzione con probabilità $1/2$. Il simbolo $|g\rangle$ indica dunque un vettore normalizzato, che può comprendere molti stati di base, e non necessariamente una singola stringa.

$$|g\rangle = \frac{1}{\sqrt{M}} \sum_{x : f(x)=1} |x\rangle$$

$$M = 2 : \qquad |g\rangle = \frac{|011\rangle + |101\rangle}{\sqrt{2}}$$

Richiami Gli stati di base distinti sono ortogonali; la probabilità di una stringa è il modulo quadro del suo coefficiente.

Lo stato uniforme dei candidati non validi

Raccogliamo ora i candidati che non soddisfano il predicato. Sono $N - M$ e ricevono tutti il coefficiente $1/\sqrt{N-M}$. Il vettore così normalizzato si indica con $|b\rangle$, dalla parola inglese bad. Le lettere descrivono la validità rispetto al predicato, non una diversa natura fisica dei qubit.

Per $N = 8$ e bersaglio 101, lo stato $|b\rangle$ comprende le sette stringhe rimanenti. Ciascuna ha probabilità $1/7$ se misuriamo proprio questo stato. Questa distribuzione va distinta dallo stato iniziale della ricerca, nel quale anche il bersaglio è presente e tutte le otto probabilità valgono $1/8$. Lo stato $|b\rangle$ serve come direzione di riferimento per descrivere l'evoluzione, senza richiedere che l'algoritmo sappia prepararlo separatamente.

$$|b\rangle = \frac{1}{\sqrt{N-M}} \sum_{x:f(x)=0} |x\rangle$$

$$N = 8,\ M = 1: \qquad \frac{1}{7} + \cdots + \frac{1}{7} = 1$$

Richiami La normalizzazione impone che la somma delle probabilità degli esiti valga 1.

Due direzioni ortogonali

Nessuna stringa compare contemporaneamente in $|g\rangle$ e in $|b\rangle$. Quando calcoliamo il loro prodotto scalare, ogni termine mette quindi a confronto due stati di base distinti e vale zero. I due vettori hanno norma 1 e sono ortogonali: costituiscono una base ortonormale del piano che ci interessa.

Per esempio, con bersaglio 101 il prodotto $\langle 101|b\rangle$ si annulla perché il coefficiente di $|101\rangle$ dentro $|b\rangle$ è zero. Un vettore $u|b\rangle + v|g\rangle$ ha perciò norma quadrata $|u|^2 + |v|^2$. Le due quantità sono rispettivamente la probabilità totale di ottenere un candidato non valido e quella di ottenere una soluzione. Non occorre sommare ancora una volta sul numero dei candidati, già incorporato nella normalizzazione dei due vettori.

$$\langle b|g\rangle = 0, \qquad \langle b|b\rangle = \langle g|g\rangle = 1$$

$$\Big\| u|b\rangle + v|g\rangle \Big\|^2 = |u|^2 + |v|^2$$

Richiami Per vettori normalizzati e ortogonali, le probabilità delle due componenti si sommano senza termini incrociati.

Le due componenti dello stato iniziale

Lo stato uniforme $|s\rangle$ contiene tutte le stringhe con coefficiente $1/\sqrt{N}$. Separiamo la somma nelle soluzioni e nei candidati non validi. La somma delle soluzioni è $\sqrt{M}|g\rangle$, mentre quella degli altri candidati è $\sqrt{N-M}|b\rangle$. Dividendo entrambe per $\sqrt{N}$ otteniamo le due coordinate dello stato iniziale.

Il coefficiente di $|g\rangle$ vale dunque $\sqrt{M/N}$ e il suo quadrato è M/N, la probabilità di successo di una misura immediata. Il coefficiente di $|b\rangle$ vale $\sqrt{1-M/N}$. La loro norma quadrata complessiva è 1. Da questo momento descriveremo Grover seguendo come cambiano questi due coefficienti, anziché riscrivere ogni volta tutte le N ampiezze.

$$|s\rangle = \sqrt{\frac{N-M}{N}}|b\rangle + \sqrt{\frac{M}{N}}|g\rangle$$

$$\frac{N-M}{N} + \frac{M}{N} = 1$$

Richiami $|s\rangle = N^{-1/2}\sum_x |x\rangle$ assegna inizialmente probabilità $1/N$ a ogni candidato.

Una coordinata per sette ampiezze

Nel caso con otto candidati e una soluzione, il coefficiente di $|b\rangle$ è $\sqrt{7/8}$, circa $0{,}9354$. Questo numero può sembrare incompatibile con le sette ampiezze iniziali $1/\sqrt{8}$, circa $0{,}3536$. La differenza si spiega ricordando che ogni termine dentro $|b\rangle$ ha già il coefficiente $1/\sqrt{7}$.

Moltiplicando $\sqrt{7/8}$ per $1/\sqrt{7}$ recuperiamo infatti $1/\sqrt{8}$ per ciascuna stringa non valida. La coordinata aggregata ha quadrato $7/8$, pari alla loro probabilità complessiva. Per il bersaglio, invece, la coordinata aggregata coincide con la singola ampiezza, perché $|g\rangle$ contiene un solo elemento. Questa distinzione evita di confondere le altezze dei grafici delle ampiezze con le coordinate del diagramma geometrico.

$$\sqrt{\frac{7}{8}}\,\frac{1}{\sqrt{7}} = \frac{1}{\sqrt{8}}$$

$$|s\rangle = \sqrt{\frac{7}{8}}|b\rangle + \frac{1}{\sqrt{8}}|g\rangle$$

Richiami Il coefficiente di $|b\rangle$ raccoglie $N - M$ ampiezze uguali; la singola ampiezza si ottiene dividendo per $\sqrt{N - M}$.

L’evoluzione rimane nello stesso piano

Per uno stato del piano, l’oracolo cambia soltanto il segno della componente sulle soluzioni: $u|b\rangle + v|g\rangle$ diventa $u|b\rangle - v|g\rangle$. Il risultato rimane quindi nel piano generato dai due vettori. Anche il diffusore lo conserva. Infatti $D|\psi\rangle = 2|s\rangle\langle s|\psi\rangle - |\psi\rangle$ combina soltanto $|s\rangle$ e lo stato d’ingresso, entrambi appartenenti al piano.

Partendo da $|s\rangle$, nessuna iterazione introduce altre direzioni. Si dice che questo piano è invariante rispetto all’azione degli operatori. Inoltre i coefficienti iniziali sono reali e le due operazioni mantengono reali i coefficienti. Possiamo quindi rappresentare l’intera traiettoria mediante due assi cartesiani reali, anche quando il registro ha molti qubit e uno spazio degli stati molto più grande.

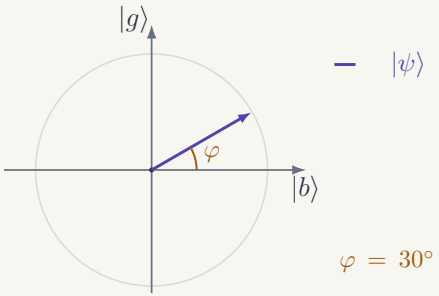
$$O_f\Big(u|b\rangle + v|g\rangle\Big) = u|b\rangle - v|g\rangle$$

Richiami O_f cambia il segno delle componenti con $f(x) = 1$; $D = 2|s\rangle\langle s| - I$ riflette rispetto alla direzione $|s\rangle$.

Le coordinate sul cerchio unitario

Disegniamo $|b\rangle$ lungo l'asse orizzontale e $|g\rangle$ lungo quello verticale. Uno stato normalizzato con coordinate reali (u, v) soddisfa $u^2 + v^2 = 1$ e si trova quindi sul cerchio di raggio 1. L'angolo φ, misurato dall'asse orizzontale in senso antiorario, permette di scriverne le coordinate come $u = \cos\varphi$ e $v = \sin\varphi$.

Il coseno è la coordinata orizzontale e il seno quella verticale. A 0 gradi lo stato coincide con $|b\rangle$, mentre a 90 gradi coincide con $|g\rangle$. Gli angoli oltre 90 gradi hanno coseno negativo, come le ampiezze non valide osservate dopo due iterazioni nell'esempio su otto candidati. Il disegno rappresenta due direzioni dell'intero registro, non la sfera di Bloch di un singolo qubit.



Piano generato da $|b\rangle$ e $|g\rangle$

Il vettore ha lunghezza 1. Le sue proiezioni sugli assi sono le coordinate dello stato, mentre i loro quadrati danno le due probabilità complessive.

Richiami Nel piano ortonormale $(|b\rangle, |g\rangle)$, la norma quadrata è $u^2 + v^2$.

Gradi e radianti

Un giro completo misura 360 gradi oppure 2π radianti. Un radiante è l'angolo che, su un cerchio di raggio 1, individua un arco lungo 1. Questa unità collega direttamente angoli e lunghezze e rende più semplici le formule trigonometriche. Un angolo retto corrisponde quindi a 90 gradi oppure a $\pi/2$ radianti.

Per convertire i gradi in radianti moltiplichiamo per $\pi/180$. L'angolo di 30 gradi vale perciò $\pi/6$, circa $0{,}5236$ radianti. Le formule che seguiranno usano radianti, salvo quando compare esplicitamente il simbolo dei gradi. I diagrammi possono mostrare una misura in gradi per facilitarne la lettura, senza cambiare l'angolo rappresentato. Occorre prestare attenzione anche alla modalità della calcolatrice quando si valutano seno e coseno.

$$\varphi_{\text{rad}} = \varphi_{\text{gradi}} \frac{\pi}{180}$$

$$30^\circ = \frac{\pi}{6} \text{ radianti}$$

Richiami Un giro completo vale 2π radianti; $\pi/2$ radianti corrispondono a 90°.

L'angolo della preparazione uniforme

Indichiamo con θ l'angolo dello stato iniziale $|s\rangle$. La sua coordinata verticale è $\sqrt{M/N}$, quindi $\sin\theta = \sqrt{M/N}$. Poiché abbiamo supposto $0 < M < N$, l'angolo iniziale è compreso fra 0 e $\pi/2$. In questo intervallo il seno cresce e individua un unico angolo.

La funzione $\arcsin$ restituisce proprio l'angolo il cui seno è il numero assegnato. Per una soluzione fra quattro candidati, $\sin\theta = 1/2$ e $\theta = \pi/6 = 30°$. Per una soluzione fra otto candidati, $\sin\theta = 1/\sqrt{8}$ e θ è circa $0{,}3614$ radianti, ossia $20{,}7048°$. Una frazione iniziale di soluzioni più piccola corrisponde a un vettore più vicino all'asse dei candidati non validi.

$$\theta = \arcsin\sqrt{\frac{M}{N}}$$

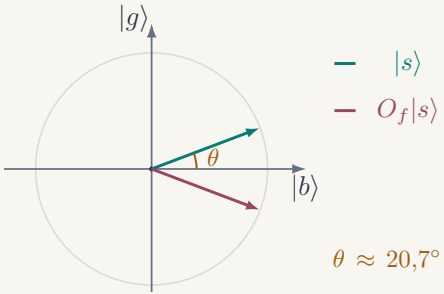
$$|s\rangle = \cos\theta|b\rangle + \sin\theta|g\rangle$$

Richiami La coordinata di $|s\rangle$ lungo $|g\rangle$ è $\sqrt{M/N}$; il suo quadrato è la probabilità iniziale di successo.

La riflessione prodotta dall'oracolo

Nel piano, l'oracolo lascia invariata la coordinata orizzontale e cambia il segno di quella verticale. Un vettore che forma un angolo φ con $|b\rangle$ passa così all'angolo $-\varphi$. Si tratta di una riflessione rispetto all'asse orizzontale, coerente con la marcatura di fase delle soluzioni.

Partendo da $|s\rangle$, che forma l'angolo θ, arriviamo a un vettore sotto l'asse, con angolo $-\theta$. La coordinata verticale cambia segno ma il suo quadrato rimane uguale. La probabilità di successo, immediatamente dopo questa operazione, è ancora M/N. Il diagramma rende visibile l'informazione di fase che una misura diretta non distingue e che il diffusore userà nel passaggio successivo.



Piano generato da $|b\rangle$ e $|g\rangle$

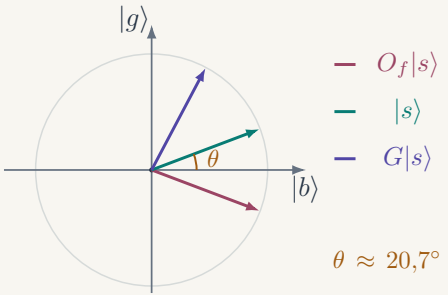
Il disegno usa $N = 8$ e $M = 1$. I due vettori hanno coordinate verticali opposte e uguale probabilità di successo. L'oracolo riflette rispetto all'asse orizzontale.

Richiami $O_f|g\rangle = -|g\rangle$ e $O_f|b\rangle = |b\rangle$; un cambio di segno conserva il modulo quadro.

La riflessione rispetto allo stato iniziale

Il diffusore riflette rispetto alla direzione di $|s\rangle$, che forma l'angolo θ con l'asse orizzontale. Se il vettore d'ingresso ha angolo φ, la sua distanza angolare dall'asse di riflessione è $\varphi - \theta$. Dopo la riflessione questa distanza cambia segno: il nuovo angolo è $\theta - (\varphi - \theta) = 2\theta - \varphi$.

All'uscita dell'oracolo il vettore iniziale si trova a $-\theta$. Applicando il diffusore otteniamo dunque $2\theta - (-\theta) = 3\theta$. Con quattro candidati l'angolo passa da $-30°$ a $90°$, raggiungendo esattamente la direzione delle soluzioni. La regola delle ampiezze $2\bar{a} - a_x$ e questa riflessione descrivono la stessa operazione, espressa in due basi diverse.



Piano generato da $|b\rangle$ e $|g\rangle$

Il disegno usa $N = 8$ e $M = 1$. La direzione centrale è l'asse del diffusore: ingresso e uscita formano con questo asse due angoli di uguale ampiezza.

Richiami $D = 2|s\rangle\langle s| - I$ conserva la componente parallela a $|s\rangle$ e cambia segno a quella perpendicolare.

Due riflessioni producono una rotazione

Consideriamo ora un generico stato del piano, con angolo φ. Prima agisce l'oracolo, che porta l'angolo a $-\varphi$. Poi agisce il diffusore, che lo trasforma in $2\theta - (-\varphi) = \varphi + 2\theta$. Un'iterazione completa aggiunge quindi sempre lo stesso angolo 2θ.

L'ordine delle operazioni è essenziale. Scriviamo $G = DO_f$ perché, nel prodotto fra operatori, quello più a destra agisce per primo sullo stato. Se applicassimo prima D e poi O_f, otterremmo invece l'angolo $\varphi - 2\theta$, una rotazione nel verso opposto. Il calcolo chiarisce sia il segno sia l'ampiezza della rotazione, senza ricorrere a un'analogia puramente visiva. Le lunghezze dei vettori rimangono invariate perché entrambe le riflessioni conservano la norma.

$$G = DO_f : \quad \varphi \longmapsto \varphi + 2\theta$$

$$O_f D : \quad \varphi \longmapsto \varphi - 2\theta$$

Richiami L'oracolo manda φ in $-\varphi$; il diffusore manda un angolo α in $2\theta - \alpha$.

La matrice dell'iterazione nel piano

Ordiniamo le coordinate come componente su $|b\rangle$ e componente su $|g\rangle$. In questa base l'oracolo ha matrice diagonale con elementi 1 e -1. La matrice del diffusore si ricava da $2|s\rangle\langle s| - I$, usando le coordinate $(\cos\theta, \sin\theta)$ dello stato iniziale.

Il prodotto DO_f cambia il segno della seconda colonna del diffusore e dà la consueta matrice di rotazione. Le identità $2\sin\theta\cos\theta = \sin 2\theta$ e $2\cos^2\theta - 1 = \cos 2\theta$ permettono di riconoscerne gli elementi. Applicandola a $(\cos\varphi, \sin\varphi)$ si ottiene $(\cos(\varphi + 2\theta), \sin(\varphi + 2\theta))$. Questa rappresentazione vale sul piano invariante, non descrive da sola tutte le altre direzioni dello spazio a N dimensioni.

$$D = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}, \quad O_f = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$G = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}$$

Richiami La base ordinata è $(|b\rangle, |g\rangle)$; le colonne di una matrice descrivono l'immagine dei vettori di base.

Lo stato dopo k iterazioni

Lo stato iniziale forma l'angolo θ. Ogni iterazione aggiunge 2θ, quindi dopo k iterazioni l'angolo è $\theta + 2k\theta = (2k+1)\theta$. Il caso $k = 0$ indica la sola preparazione, prima di interrogare l'oracolo. Il caso $k = 1$ corrisponde invece a una marcatura seguita da un diffusore.

Le coordinate finali si ottengono prendendo coseno e seno del nuovo angolo. La formula permette di seguire la ricerca con un numero qualsiasi di iterazioni senza moltiplicare ripetutamente matrici a N dimensioni. Non significa che il circuito conosca $|g\rangle$ o $|b\rangle$: questi vettori servono all'analisi. Il circuito continua ad applicare il verificatore reversibile e il diffusore, entrambi definiti senza elencare le soluzioni trovate.

$$|\psi_k\rangle = G^k|s\rangle$$

$$|\psi_k\rangle = \cos((2k+1)\theta)|b\rangle + \sin((2k+1)\theta)|g\rangle$$

Richiami k conta le applicazioni di $G = DO_f$; θ è l'angolo della preparazione uniforme.

La probabilità di misurare una soluzione

La componente sulle soluzioni ha coefficiente $\sin((2k+1)\theta)$. Poiché $|g\rangle$ è normalizzato, il modulo quadro di questo coefficiente è già la probabilità totale di successo. Otteniamo quindi $P_k = \sin^2((2k+1)\theta)$. La probabilità di ottenere un candidato non valido è il coseno quadro dello stesso angolo.

All'interno dello stato $|g\rangle$, ciascuna delle M soluzioni ha coefficiente $1/\sqrt{M}$. La singola ampiezza dopo k iterazioni è pertanto $\sin((2k+1)\theta)/\sqrt{M}$ e la sua probabilità è P_k/M. Per $M = 1$ queste due descrizioni coincidono. Per più soluzioni è invece necessario distinguere il successo complessivo dalla probabilità di una particolare stringa, che può essere assai più piccola.

$$P_k = \sin^2((2k+1)\theta)$$

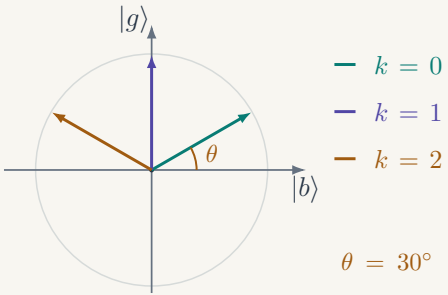
$$\Pr(\text{esito } x) = \frac{P_k}{M} \quad \text{se } f(x) = 1$$

Richiami $|g\rangle$ è la sovrapposizione uniforme normalizzata delle M soluzioni.

La soluzione esatta su quattro candidati

Con $N = 4$ e $M = 1$, l'angolo iniziale è $\theta = \pi/6$. Dopo una iterazione l'angolo diventa $3\theta = \pi/2$, perciò la componente sulle soluzioni vale 1 e quella sugli altri candidati vale zero. Il risultato coincide con il calcolo delle quattro ampiezze svolto in precedenza.

Se proseguiamo, la rotazione continua anche dopo aver raggiunto il bersaglio. Alla seconda iterazione l'angolo è $5\pi/6$ e la probabilità torna a $1/4$. Alla terza arriva a $7\pi/6$, con la stessa probabilità ma con una componente sulle soluzioni negativa. L'esito certo della prima iterazione dipende dall'allineamento esatto degli angoli: tre volte 30 gradi formano un angolo retto.



Piano generato da $|b\rangle$ e $|g\rangle$

Dopo una iterazione il vettore coincide con l'asse delle soluzioni. La seconda iterazione lo porta oltre quell'asse, facendo diminuire la probabilità di successo.

Richiami $P_k = \sin^2((2k+1)\theta)$, con $\sin\theta = \sqrt{M/N}$.

Gli angoli dell'esempio su otto candidati

Con $N = 8$ e $M = 1$, l'angolo iniziale è circa $20{,}7048°$. Ogni iterazione aggiunge circa $41{,}4096°$. Dopo il primo passo il vettore forma circa $62{,}1144°$ con l'asse orizzontale, mentre dopo il secondo arriva a circa $103{,}5241°$. Il secondo vettore è più vicino all'asse verticale.

Il terzo passo conduce a circa $144{,}9337°$ e allontana nuovamente il vettore dalla direzione delle soluzioni. Il seno quadro restituisce esattamente le probabilità già ottenute dalle ampiezze: $25/32$, $121/128$ e $169/512$. Le misure decimali degli angoli sono arrotondate, mentre queste frazioni sono esatte. La geometria spiega la diminuzione al terzo passo che il solo nome amplificazione potrebbe far trascurare.

k	**Angolo**	P_k
0	$20{,}7048°$	$1/8$
1	$62{,}1144°$	$25/32$
2	$103{,}5241°$	$121/128$
3	$144{,}9337°$	$169/512$

La probabilità è calcolata con l'angolo esatto. Gli angoli stampati in tabella servono soltanto alla lettura geometrica e non sostituiscono le frazioni esatte.

Richiami Per una soluzione fra otto candidati, $\theta = \arcsin(1/\sqrt{8})$.

Le ampiezze negative oltre l'asse verticale

Dopo due iterazioni su otto candidati il vettore ha superato i 90 gradi. La coordinata orizzontale, uguale al coseno dell'angolo, è quindi negativa. Non si tratta di una probabilità negativa: il coseno è un'ampiezza aggregata, il cui quadrato determina la probabilità totale degli esiti non validi.

Dai calcoli precedenti, ognuno dei sette candidati non validi ha ampiezza $-1/(8\sqrt{2})$. Moltiplicando per $\sqrt{7}$ otteniamo la coordinata su $|b\rangle$. Il bersaglio ha ampiezza $11/(8\sqrt{2})$. I due quadrati sono $7/128$ e $121/128$ e sommano a 1. I segni spiegano dunque la posizione del vettore e influenzano l'iterazione successiva, anche se la misura legge soltanto i moduli quadrati.

$$|\psi_2\rangle = -\frac{\sqrt{7}}{8\sqrt{2}}|b\rangle + \frac{11}{8\sqrt{2}}|g\rangle$$

$$\frac{7}{128} + \frac{121}{128} = 1$$

Richiami La coordinata su $|b\rangle$ è $\sqrt{N-M}$ volte la singola ampiezza non valida, quando tali ampiezze sono uguali.

Il primo massimo della curva continua

Per ottenere probabilità 1 vorremmo che il vettore raggiungesse l'angolo $\pi/2$. Risolviamo quindi $(2k+1)\theta = \pi/2$ rispetto a k. Si ottiene il valore reale $k_* = \pi/(4\theta) - 1/2$. Di solito questo numero non è intero e non corrisponde a un numero eseguibile di iterazioni.

Il valore k_* individua il primo massimo se immaginiamo di variare continuamente il numero dei passi. Nel circuito possiamo confrontare gli interi immediatamente sotto e sopra questo valore. Il primo massimo è quello utile a raggiungere presto una probabilità elevata. Non stiamo affermando che nessun numero di iterazioni molto più grande possa produrre una probabilità ancora più vicina a 1: le rotazioni successive possono tornare vicino all'asse verticale, ma richiedono altro lavoro.

$$(2k_* + 1)\theta = \frac{\pi}{2}$$

$$k_* = \frac{\pi}{4\theta} - \frac{1}{2}$$

Richiami Il successo ha probabilità $\sin^2((2k+1)\theta)$; il primo valore angolare con successo certo è $\pi/2$.

L’arrotondamento e la probabilità garantita

Scegliamo un intero k più vicino a k_*. La differenza fra i due numeri non supera $1/2$. Poiché ogni passo corrisponde a un angolo 2θ, la distanza angolare dal primo massimo non supera θ. Se i due interi sono equidistanti, danno la stessa probabilità e conviene quello più piccolo.

Chiamando ε questa distanza, la probabilità è $\sin^2(\pi/2 + \varepsilon) = \cos^2 \varepsilon$. Nell’intervallo considerato risulta almeno $\cos^2 \theta = 1 - M/N$. Per una sola soluzione fra molti candidati il limite è vicino a 1. Per esempio, con $N = 8$ garantisce almeno $7/8$, mentre il valore effettivo dopo due iterazioni è $121/128$, che è maggiore. Il limite è una garanzia, non una previsione esatta del successo.

$$|k - k_*| \leq \frac{1}{2} \quad \Longrightarrow \quad |\varepsilon| \leq \theta$$

$$P_k = \cos^2 \varepsilon \geq 1 - \frac{M}{N}$$

Richiami $\sin^2 \theta = M/N$ e $\sin^2 \theta + \cos^2 \theta = 1$.

La scelta di due iterazioni

Per una soluzione fra otto candidati, il primo massimo continuo si trova a $k_* \approx 1{,}6734$. Gli interi da confrontare sono 1 e 2. Le probabilità corrispondenti sono $25/32 \approx 0{,}7813$ e $121/128 \approx 0{,}9453$, quindi scegliamo due iterazioni.

Il grafico mostra anche i passi successivi: la probabilità oscilla e non aumenta indefinitamente. La misura va eseguita dopo il numero di passi scelto, poi il candidato ottenuto va controllato con il predicato. Se non è una soluzione, una nuova esecuzione deve ripartire dallo stato iniziale. Proseguire il circuito dopo aver misurato non equivale a continuare la stessa rotazione, perché la misura ha modificato lo stato. La scelta del numero di iterazioni presuppone qui che M sia noto.

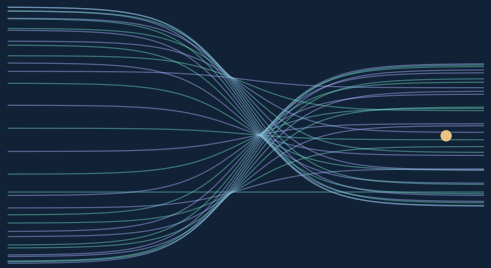


I punti interi corrispondono ai numeri di iterazioni eseguibili. La curva tra i punti aiuta a riconoscere l'oscillazione, senza rappresentare passi frazionari del circuito.

Richiami $k_* = \pi/(4\theta) - 1/2$ individua il primo picco continuo; nel circuito k deve essere un intero non negativo.

CAPITOLO 08

Circuiti e ausiliari



La costruzione dell’oracolo e del diffusore

L’algoritmo di Grover

L’ordine dei fili e delle operazioni

In un circuito ogni linea orizzontale rappresenta un qubit e il tempo procede da sinistra verso destra. Una porta modifica lo stato dei fili sui quali è disegnata. Per tre qubit usiamo dall’alto verso il basso l’ordine q_2, q_1, q_0, così il filo più alto corrisponde al bit di peso maggiore.

La stringa 101 significa quindi $q_2 = 1$, $q_1 = 0$, $q_0 = 1$ e codifica il numero $4 + 0 + 1 = 5$. Questa convenzione permette di confrontare i disegni con i vettori scritti nell’ordine $000, 001, \ldots, 111$. Quando compaiono fili ausiliari, li aggiungiamo sotto quelli dei dati e li nominiamo separatamente. Il numero dei fili aumenta, ma il candidato continua a essere identificato soltanto dai qubit del registro dei dati.

La stringa 101

x_2	x_1	x_0
1	0	1
$2^2 = 4$	$2^1 = 2$	$2^0 = 1$

$$4 + 0 + 1 = 5$$

Il bit più significativo è a sinistra.

I pesi dei bit collegano la stringa al numero intero. Nei circuiti il bit di peso 4 occuperà il filo più alto fra quelli dei dati.

Richiami Una stringa $q_2 q_1 q_0$ rappresenta l’intero $4q_2 + 2q_1 + q_0$.

Le porte X e Z

La porta X scambia $|0\rangle$ e $|1\rangle$. Su una sovrapposizione scambia quindi le due ampiezze: $a|0\rangle + b|1\rangle$ diventa $b|0\rangle + a|1\rangle$. Il suo effetto sui valori di base coincide con la negazione di un bit classico, ma l'azione lineare la rende applicabile anche quando il qubit non ha un valore definito.

La porta Z lascia invariato $|0\rangle$ e cambia il segno di $|1\rangle$. Le probabilità nella base computazionale rimangono uguali, mentre cambia la fase relativa delle componenti. Per esempio $Z(|0\rangle + |1\rangle)/\sqrt{2} = (|0\rangle - |1\rangle)/\sqrt{2}$. Entrambe le porte sono la propria inversa: applicarle due volte ripristina qualunque stato d'ingresso. Le useremo per spostare una configurazione da marcare e per applicare la marcatura di fase.

$$X|0\rangle = |1\rangle, \quad X|1\rangle = |0\rangle$$

$$Z|0\rangle = |0\rangle, \quad Z|1\rangle = -|1\rangle$$

Richiami La linearità consente di ricavare l'azione su una sovrapposizione conoscendo l'azione sugli stati di base.

La preparazione con Hadamard

La porta Hadamard, indicata con H, manda $|0\rangle$ nello stato $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ e $|1\rangle$ nello stato $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Per linearità, una seconda applicazione restituisce rispettivamente $|0\rangle$ e $|1\rangle$. Abbiamo dunque $H^2 = I$, dove I è l'identità.

Applicando H a ogni qubit inizializzato a zero otteniamo tutte le stringhe con la stessa ampiezza. Su tre qubit ciascuna ampiezza è il prodotto di tre fattori $1/\sqrt{2}$, quindi vale $1/\sqrt{8}$. La notazione $H^{\otimes 3}$ indica tre Hadamard indipendenti sui tre fili, non tre applicazioni successive sullo stesso qubit. Lo stesso insieme di porte, applicato nuovamente, riporta lo stato uniforme a $|000\rangle$.

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \qquad H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$H^{\otimes n}|0^n\rangle = |s\rangle$$

Richiami $\otimes$ combina gli stati o le operazioni di registri distinti; $H^2 = I$.

La negazione controllata

La porta CNOT agisce su due qubit: un controllo c e un bersaglio t. Se il controllo vale 0, il bersaglio rimane invariato. Se vale 1, il bersaglio viene negato. Il controllo non cambia. La regola compatta è $|c, t\rangle \mapsto |c, t \oplus c\rangle$, dove $\oplus$ indica lo XOR.

Nei circuiti il controllo si rappresenta con un punto pieno, collegato al bersaglio rappresentato da un cerchio con un segno più. Quest'ultimo simbolo indica uno XOR, non una somma ordinaria che possa produrre 2. Applicare due volte la stessa CNOT annulla le due negazioni eventuali e ripristina lo stato. La tabella descrive gli ingressi di base; su una loro sovrapposizione la porta agisce contemporaneamente sui coefficienti secondo la linearità.

Ingresso	Uscita
$\|00\rangle$	$\|00\rangle$
$\|01\rangle$	$\|01\rangle$
$\|10\rangle$	$\|11\rangle$
$\|11\rangle$	$\|10\rangle$

Il primo bit è il controllo e il secondo il bersaglio. Le quattro uscite sono distinte, perciò la trasformazione non perde l'informazione sull'ingresso.

Richiami Lo XOR vale 1 quando i due bit sono diversi; $t \oplus c \oplus c = t$.

La porta Toffoli e la congiunzione

La porta Toffoli ha due controlli c_1, c_2 e un bersaglio t. Nega il bersaglio soltanto quando entrambi i controlli valgono 1. La sua azione è $t \mapsto t \oplus (c_1 \wedge c_2)$, lasciando invariati i controlli. Il simbolo $\wedge$ indica la congiunzione logica AND.

Se inizializziamo il bersaglio a zero, l'uscita contiene proprio $c_1 \wedge c_2$. Conservare i controlli rende questa operazione reversibile, mentre sostituire due bit con il solo valore del loro AND cancellerebbe informazione. La tabella mostra il caso con bersaglio iniziale zero. Con bersaglio iniziale 1, la stessa porta nega il valore della congiunzione tramite XOR. Una seconda Toffoli sugli stessi fili ripristina il bersaglio, proprietà che useremo per pulire i registri ausiliari.

c_1	c_2	t iniziale	t finale
0	0	0	0
0	1	0	0
1	0	0	0
1	1	0	1

Il bersaglio registra la congiunzione solo perché parte da zero. I due controlli rimangono disponibili e consentono di invertire il calcolo.

Richiami $c_1 \wedge c_2 = 1$ soltanto se entrambi i bit sono 1; lo XOR con lo stesso bit ripetuto due volte si annulla.

La fase controllata su tutti i bit

La porta Z controllata su più fili cambia il segno soltanto della componente in cui tutti quei qubit valgono 1. Con due qubit si chiama CZ e marca $|11\rangle$. Con tre qubit si indica spesso con CCZ e marca $|111\rangle$. Useremo MCZ come nome generico della versione su n qubit.

Il valore delle stringhe non cambia: a cambiare è il loro coefficiente nello stato. Per esempio, CCZ manda $(|110\rangle + |111\rangle)/\sqrt{2}$ in $(|110\rangle - |111\rangle)/\sqrt{2}$. La sua rappresentazione come singolo simbolo nel circuito descrive un'operazione logica. Una macchina che offre soltanto porte a uno e due qubit deve decomporla in una sequenza di porte disponibili, il cui costo non va ignorato quando si stimano le risorse.

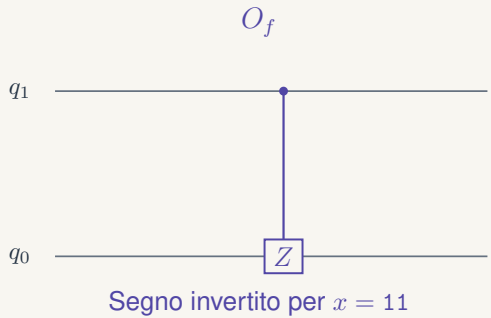
$$\mathrm{MCZ}|x\rangle = \begin{cases} -|x\rangle, & x = 11\cdots1, \\ |x\rangle, & \text{altrimenti.} \end{cases}$$

Richiami Una marcatura di fase moltiplica per -1 le ampiezze selezionate e lascia invariati i moduli quadrati.

L’oracolo per il bersaglio 11

Nel primo esempio vogliamo cambiare segno soltanto alla stringa 11. La porta CZ compie esattamente questa trasformazione. Applicata allo stato uniforme di due qubit, lascia positive le ampiezze di 00, 01 e 10 e rende negativa quella di 11.

Questo circuito elementare rende concreta la marcatura già usata nei calcoli. Conoscere la porta ci permette anche di verificare l’azione su ciascuno dei quattro ingressi. In un problema di ricerca reale, tuttavia, non costruiamo il circuito partendo dalla soluzione che dobbiamo ancora trovare. Disponiamo di un predicato che riconosce le soluzioni. Il bersaglio noto serve qui a controllare il meccanismo dell’algoritmo su un caso interamente calcolabile.



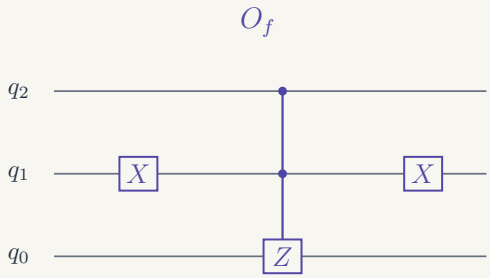
La fase negativa compare soltanto quando entrambi i fili valgono 1. Nessun bit cambia valore durante l’applicazione dell’oracolo.

Richiami $O_f|x\rangle = (-1)^{f(x)}|x\rangle$; in questo esempio $f(11) = 1$ e gli altri tre valori sono zero.

La marcatura della stringa 101

Per marcare 101 possiamo usare CCZ, purché trasformiamo temporaneamente quel bersaglio in 111. Basta applicare X al qubit centrale q_1, che nel bersaglio vale zero. Dopo la porta di fase, applichiamo di nuovo X allo stesso filo per ripristinare la codifica originale.

La composizione cambia quindi soltanto il segno di $|101\rangle$. Per esempio, la stringa 111 diventa temporaneamente 101, non riceve la fase di CCZ e torna a 111. Questo controllo mostra che la sequenza non marca accidentalmente anche la stringa tutta a 1. In generale, per un bersaglio fissato, le porte X vanno poste prima e dopo MCZ sui fili in cui il bersaglio contiene uno zero.



Segno invertito per $x = 101$

Le due porte X sul filo centrale trasformano e poi ripristinano la codifica. La fase acquisita dalla componente bersaglio rimane dopo il ripristino.

Richiami $X^2 = I$; CCZ cambia segno soltanto a $|111\rangle$.

Il verificatore come descrizione del problema

Gli oracoli costruiti a partire da 11 o 101 contengono esplicitamente il bersaglio. Sono utili per studiare il circuito, ma non rappresentano un compito di ricerca in cui la risposta è ancora sconosciuta al progettista. In quel caso il dato iniziale è una regola di verifica: per esempio, una formula booleana che deve risultare vera.

Il circuito reversibile valuta la regola su un candidato e ne usa il risultato per applicare una fase. Scrivere la regola può essere semplice anche quando trovare un ingresso che la soddisfa richiede molti tentativi. Per distinguere le due situazioni, costruiremo ora un piccolo predicato composto da XOR e AND. Potremo realizzarne l'oracolo seguendo le operazioni logiche, senza cominciare dall'elenco dei suoi ingressi validi. L'elenco servirà soltanto a verificare il risultato dell'esempio.

Richiami Un predicato f restituisce 1 per le soluzioni e 0 per gli altri candidati.

La riflessione rispetto al registro azzerato

Per costruire il diffusore partiamo dall'operatore $R_0 = 2|0^n\rangle\langle 0^n| - I$. Esso conserva la componente su $|0^n\rangle$ e cambia segno a tutte le componenti ortogonali. Lo stato uniforme si ottiene applicando $H^{\otimes n}$ a $|0^n\rangle$, quindi le Hadamard permettono di trasportare questa riflessione sulla direzione $|s\rangle$.

Il prodotto $H^{\otimes n} R_0 H^{\otimes n}$ è infatti $2|s\rangle\langle s| - I = D$. La prima serie di Hadamard, quella a destra nella formula, esprime lo stato nella base in cui la direzione da conservare è $|0^n\rangle$. La riflessione agisce in quella base e la seconda serie riporta alla base originale. Questa costruzione non dipende dal bersaglio né dal numero delle soluzioni.

$$R_0 = 2|0^n\rangle\langle 0^n| - I$$

$$D = H^{\otimes n} R_0 H^{\otimes n}$$

Richiami $H^{\otimes n}|0^n\rangle = |s\rangle$ e $(H^{\otimes n})^2 = I$.

La marcatura dello zero e il suo segno

Una porta X su ogni filo scambia $00\cdots0$ con $11\cdots1$. Possiamo perciò cambiare segno alla sola stringa nulla usando $X^{\otimes n}\mathrm{MCZ}X^{\otimes n}$. Le porte X esterne ripristinano tutte le stringhe dopo la marcatura.

L'operazione ottenuta è $S_0 = I - 2|0^n\rangle\langle 0^n|$. Agisce con segno negativo sullo zero e positivo sugli altri stati di base. Il confronto con R_0 mostra che $S_0 = -R_0$: i due operatori hanno tutti i segni opposti. Questo meno globale non cambia le probabilità quando l'intera operazione agisce senza controllo su uno stato. Va comunque conservato nel calcolo, soprattutto se in seguito vogliamo rendere la stessa operazione condizionata a un altro qubit.

$$S_0 = X^{\otimes n}\mathrm{MCZ}X^{\otimes n}$$

$$S_0 = I - 2|0^n\rangle\langle 0^n| = -R_0$$

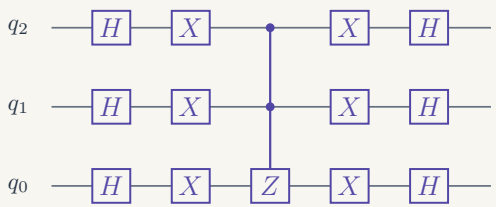
Richiami $R_0 = 2|0^n\rangle\langle 0^n| - I$ conserva lo zero e cambia segno alle componenti ortogonali.

Il circuito usuale realizza meno D

Inseriamo la marcatura dello zero fra due serie di Hadamard. Da sinistra verso destra il circuito esegue H su tutti i fili, poi X su tutti i fili, la porta MCZ, ancora X e infine H. Il suo operatore è $H^{\otimes n} S_0 H^{\otimes n} = -D$.

Il segno è importante per confrontare le ampiezze del circuito con quelle calcolate usando D. Dopo una sola applicazione tutte le ampiezze sono opposte, ma le probabilità coincidono. Nell'algoritmo ordinario possiamo usare questo circuito per il diffusore, purché manteniamo chiara la convenzione. L'etichetta nel diagramma riporta quindi $-D$. Una decomposizione in porte elementari deve preservare lo stesso operatore, oppure dichiarare e gestire la fase globale introdotta.

La rete realizza $-D$



Le colonne H e X agiscono su tutti i qubit. La sequenza realizza meno D: lo stato differisce da D per un unico segno globale.

Richiami $S_0 = -R_0$ e $D = H^{\otimes n} R_0 H^{\otimes n}$.

Il segno globale nelle iterazioni ordinarie

Se sostituiamo D con $-D$, un'iterazione diventa $(-D)O_f = -G$. Dopo k iterazioni il risultato è $(-G)^k|s\rangle = (-1)^k G^k|s\rangle$. Il fattore $(-1)^k$ moltiplica l'intero stato, senza dipendere dal candidato.

Per k pari gli stati coincidono esattamente. Per k dispari tutte le ampiezze hanno segno opposto, ma ogni probabilità rimane identica. Anche applicando altre operazioni lineari all'intero stato, quel fattore resta comune a tutte le componenti e non diventa osservabile. È per questo che la sequenza usuale di porte è adatta alla ricerca di Grover non controllata. La situazione cambia se solo una parte di una sovrapposizione riceve quel segno, perché allora si modifica una fase relativa.

$$((-D)O_f)^k|s\rangle = (-1)^k G^k|s\rangle$$

Richiami Gli stati $|\psi\rangle$ e $-|\psi\rangle$ differiscono per una fase globale e producono le stesse probabilità.

Il segno diventa relativo sotto controllo

Consideriamo un qubit di controllo preparato in $|+\rangle$ e il registro dei dati preparato in $|s\rangle$. L'operazione controllata applica il diffusore soltanto quando il controllo vale 1. Siccome $D|s\rangle = |s\rangle$, il diffusore controllato lascia invariato lo stato complessivo $|+\rangle|s\rangle$.

Se il ramo controllato applica invece $-D$, lo stato finale è $|-\rangle|s\rangle$: solo la componente con controllo 1 riceve il meno. Una Hadamard sul controllo distingue i due casi, producendo rispettivamente 0 oppure 1 con certezza. La differenza è dunque osservabile. Una porta Z sul controllo corregge il segno relativo. Questo esempio spiega perché un'identità valida a meno di fase globale richieda attenzione quando viene incorporata in un'operazione controllata.

$$C(D)|+\rangle|s\rangle = |+\rangle|s\rangle$$

$$C(-D)|+\rangle|s\rangle = |-\rangle|s\rangle$$

Richiami $D|s\rangle = |s\rangle$, $H|+\rangle = |0\rangle$, $H|-\rangle = |1\rangle$ e $Z|+\rangle = |-\rangle$.

Un predicato composto da XOR e AND

Consideriamo $f(x_2, x_1, x_0) = (x_2 \oplus x_1) \wedge x_0$. Il primo calcolo controlla se i due bit più significativi sono diversi. Il secondo richiede anche che il bit meno significativo valga 1. Una soluzione deve quindi avere $x_2 \neq x_1$ e $x_0 = 1$.

Per 011, lo XOR dei primi due bit vale 1 e l'AND con l'ultimo bit vale ancora 1. Per 010 lo XOR vale ugualmente 1, ma l'ultimo bit è zero e il predicato restituisce 0. La formula fornisce una procedura di verifica senza richiedere una comparazione esplicita con una stringa bersaglio. Nei passaggi successivi realizzeremo questa procedura con porte reversibili, poi elimineremo i risultati temporanei mantenendo la fase necessaria a Grover.

$$f(x_2, x_1, x_0) = (x_2 \oplus x_1) \wedge x_0$$

Richiami $u \oplus v = 1$ quando $u \neq v$; $u \wedge v = 1$ soltanto quando entrambi valgono 1.

Il predicato quando il primo bit vale zero

Quando $x_2 = 0$, lo XOR $x_2 \oplus x_1$ coincide con x_1. Il predicato si riduce perciò a $x_1 \wedge x_0$. Fra le quattro stringhe che iniziano con zero, soltanto 011 ha entrambi questi bit uguali a 1.

La tabella separa il risultato intermedio dello XOR dal valore finale del predicato. Questa separazione è utile per progettare il circuito: un primo ausiliario conserverà temporaneamente lo XOR e un secondo conserverà il risultato dell'AND. Gli ingressi rimarranno sui loro tre fili. L'uscita zero di f non implica che tutti i risultati intermedi siano zero: per 010, per esempio, lo XOR vale 1. Anche questi dati temporanei dovranno essere eliminati reversibilmente dopo la marcatura.

Stringa	$x_2 \oplus x_1$	f
000	0	0
001	0	0
010	1	0
011	1	1

Le quattro righe coprono tutti gli ingressi con primo bit zero. La colonna centrale mostra il dato temporaneo che il circuito dovrà poi ripristinare.

Richiami $f = (x_2 \oplus x_1) \wedge x_0$; il risultato intermedio può essere diverso dal valore finale.

Il predicato quando il primo bit vale uno

Quando $x_2 = 1$, lo XOR $x_2 \oplus x_1$ vale 1 se $x_1 = 0$ e vale 0 se $x_1 = 1$. Per ottenere una soluzione occorre inoltre $x_0 = 1$. Fra le quattro stringhe che iniziano con uno, soltanto 101 soddisfa entrambe le condizioni.

Unendo questa tabella alla precedente troviamo due soluzioni, 011 e 101, fra otto candidati. Il circuito che costruiremo seguirà comunque la formula del predicato, attraverso XOR e AND. L'elenco delle soluzioni ci servirà per controllare che la fase negativa compaia esattamente sui due ingressi previsti. Le righe 110 e 111 mostrano anche che avere diversi bit uguali a 1 non basta: conta la relazione logica stabilita da f.

Stringa	$x_2 \oplus x_1$	f
100	1	0
101	1	1
110	0	0
111	0	0

La seconda metà della tabella aggiunge la soluzione 101. Insieme alle quattro righe precedenti permette di verificare l'oracolo su tutti gli stati di base.

Richiami Le soluzioni sono gli ingressi per cui il predicato restituisce 1; qui $N = 8$ e $M = 2$.

Il calcolo reversibile dello XOR

Aggiungiamo due qubit ausiliari, chiamati a e t, inizialmente a zero. Il primo dovrà contenere $x_2 \oplus x_1$. Applichiamo una CNOT con controllo x_2 e bersaglio a, poi una CNOT con controllo x_1 e lo stesso bersaglio. Il valore finale è $0 \oplus x_2 \oplus x_1$.

Per l'ingresso 101, il primo controllo porta a a 1 e il secondo lo lascia invariato. Per 111, entrambe le CNOT agiscono e a torna a zero. I tre bit del candidato non cambiano. Il circuito memorizza quindi un risultato intermedio senza cancellare gli ingressi dai quali lo ha ricavato. Su una sovrapposizione, l'ausiliario si correla con ciascuna stringa secondo la stessa regola lineare.

$$|x\rangle|0\rangle_a|0\rangle_t \longmapsto |x\rangle|x_2 \oplus x_1\rangle_a|0\rangle_t$$

Richiami Una CNOT sostituisce il bersaglio a con $a \oplus c$ e conserva il controllo c.

Il calcolo del valore del predicato

Il primo ausiliario contiene ora $a = x_2 \oplus x_1$. Applichiamo una Toffoli con controlli a e x_0 e bersaglio t, inizializzato a zero. Il bersaglio riceve quindi $a \wedge x_0 = f(x)$. Indichiamo con V l'insieme delle operazioni di calcolo appena descritte.

Per ogni ingresso di base otteniamo $|x\rangle|a(x)\rangle|f(x)\rangle$. I risultati temporanei dipendono dall'ingresso e possono quindi essere correlati con il registro dei dati. Non possiamo ancora considerarli qubit indipendenti e ignorarli. Il loro ruolo è rendere disponibile il valore di f come controllo della fase. Dopo aver usato quel valore dovremo invertire il calcolo, affinché gli ausiliari tornino a uno stato comune a tutti gli ingressi.

Calcolo del predicato



Le CNOT calcolano lo XOR nel primo ausiliario. La Toffoli usa quel risultato e il bit meno significativo per scrivere il predicato nel secondo ausiliario.

Richiami Con bersaglio iniziale zero, una Toffoli registra l'AND dei due controlli senza modificarli.

La fase applicata al qubit di esito

Applichiamo ora una porta Z al qubit t che contiene $f(x)$. Se il candidato non è valido, $t = 0$ e la componente rimane invariata. Se il candidato è una soluzione, $t = 1$ e la componente acquista un segno negativo. Il valore del qubit t non cambia durante l'applicazione di Z.

Per l'ingresso 101 lo stato diventa $-|101\rangle|1\rangle_a|1\rangle_t$. Per 010, invece, rimane $|010\rangle|1\rangle_a|0\rangle_t$. La fase desiderata è già presente, ma gli ausiliari conservano ancora dati dipendenti dall'ingresso. La trasformazione complessiva non è ancora l'oracolo pulito O_f sul solo registro dei candidati. Il passo successivo eliminerà questi dati temporanei senza annullare il segno appena applicato.

Marcatura della risposta



$$Z_t|t\rangle = (-1)^t|t\rangle$$

La porta Z agisce sul qubit che contiene il valore del predicato. Le componenti con esito 1 ricevono il segno negativo richiesto dall'oracolo.

Richiami $Z|f(x)\rangle = (-1)^{f(x)}|f(x)\rangle$; la porta cambia la fase, non il valore di base del qubit.

Il ripristino dei due ausiliari

Invertiamo ora il calcolo V. La prima porta da ripetere è l'ultima eseguita, cioè la Toffoli. Essa sostituisce $t = f(x)$ con $f(x) \oplus f(x) = 0$. Poi ripetiamo le due CNOT nell'ordine inverso, riportando a a $x_2 \oplus x_1 \oplus x_1 \oplus x_2 = 0$.

La fase $(-1)^{f(x)}$ rimane come fattore del vettore. Per linearità, ripristinare i bit ausiliari non rimuove quel coefficiente. L'operazione completa è quindi $V^\dagger Z_t V$, dove $V^\dagger$ indica l'inversa unitaria di V. Su ausiliari inizialmente nulli essa applica proprio O_f ai dati e restituisce entrambi gli ausiliari a zero. Questa inversione del calcolo temporaneo è spesso chiamata uncompute.

Ripristino degli ausiliari



$a, t \longmapsto 0, 0$

Le porte inverse cancellano prima il valore del predicato e poi lo XOR intermedio. I qubit dei dati conservano la fase ricevuta durante il calcolo.

Richiami CNOT e Toffoli sono le proprie inverse; l'inversa di una sequenza esegue le porte inverse nell'ordine opposto.

La traccia completa per il candidato 101

Seguiamo il candidato 101, che soddisfa il predicato. I dati rimangono 101 durante tutta la sequenza. I due ausiliari, scritti nell'ordine a, t, passano da 00 a 10 e poi a 11. La porta Z introduce un meno quando il secondo ausiliario vale 1.

La Toffoli inversa riporta gli ausiliari a 10 e le CNOT inverse li riportano a 00. Il risultato finale è $-|101\rangle|00\rangle$. Il meno non va interpretato come un nuovo valore di un bit: è il coefficiente della componente quantistica. Se eseguissimo la stessa traccia su 010, gli ausiliari passerebbero da 00 a 10 e poi tornerebbero a 00, senza ricevere alcun segno negativo perché t rimarrebbe sempre zero.

Passaggio	a, t	Fattore
Inizio	00	$+1$
XOR	10	$+1$
AND	11	$+1$
Z	11	-1
AND inverso	10	-1
XOR inverso	00	-1

La tabella distingue i valori degli ausiliari dal coefficiente di fase. Il ripristino dei valori a zero lascia intatto il coefficiente negativo.

Richiami $a = x_2 \oplus x_1$ e $t = a \wedge x_0$; Z modifica il segno soltanto quando $t = 1$.

Il ripristino vale per tutta la sovrapposizione

Abbiamo verificato l'azione del circuito su ciascuno stato di base. La linearità estende il risultato a uno stato arbitrario $\sum_x c_x|x\rangle$: ogni componente riceve il fattore $(-1)^{f(x)}$ e termina con gli ausiliari nello stesso stato $|00\rangle$. Possiamo quindi raccogliere questo stato ausiliario come fattore comune.

Il registro dei dati contiene esattamente $O_f \sum_x c_x|x\rangle$ e gli ausiliari sono disaccoppiati. Applicare il diffusore ai dati produce allora l'interferenza prevista dall'analisi matematica. Non è sufficiente che gli ausiliari abbiano probabilità elevate di tornare a zero: il ripristino ideale deve valere coerentemente per tutte le componenti. Conoscere l'azione completa, comprese le fasi, su tutti gli stati di base ne garantisce l'estensione alle sovrapposizioni.

$$\sum_x c_x|x\rangle|00\rangle \longmapsto \left(\sum_x (-1)^{f(x)} c_x|x\rangle\right)|00\rangle$$

Richiami Una trasformazione lineare agisce su una somma applicando la stessa trasformazione a ogni termine.

Un ausiliario può impedire l'interferenza

Un esempio con due qubit mostra il problema dei dati temporanei residui. Prepariamo il qubit dei dati in $|+\rangle$ e l'ausiliario in $|0\rangle$. Una CNOT dai dati all'ausiliario produce $(|00\rangle + |11\rangle)/\sqrt{2}$. I due valori dei dati sono ora accompagnati da stati ausiliari distinguibili.

Applichiamo H soltanto ai dati. Lo stato diventa $(|00\rangle + |10\rangle + |01\rangle - |11\rangle)/2$, dove il primo bit identifica i dati. La probabilità di misurare zero su quel qubit è $1/4 + 1/4 = 1/2$. Senza la correlazione con l'ausiliario, H avrebbe invece riportato $|+\rangle$ a $|0\rangle$ con certezza. I contributi con ausiliari diversi non si cancellano fra loro quando leggiamo soltanto i dati.

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \xrightarrow{H\otimes I} \frac{|00\rangle + |10\rangle + |01\rangle - |11\rangle}{2}$$

$$\Pr(\text{dato} = 0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Richiami Per la probabilità di un esito sui dati si sommano le probabilità dei diversi valori dell'ausiliario non osservato.

L’inversa ricostruisce l’interferenza

Ripartiamo dallo stato correlato $(|00\rangle + |11\rangle)/\sqrt{2}$. Prima di applicare H ai dati, ripetiamo la CNOT usata per scrivere l’ausiliario. Il termine $|00\rangle$ rimane invariato, mentre $|11\rangle$ diventa $|10\rangle$. Otteniamo quindi $(|0\rangle + |1\rangle)|0\rangle/\sqrt{2} = |+\rangle|0\rangle$.

Una Hadamard sui dati produce ora $|00\rangle$ con certezza. L’inversa ha fatto coincidere gli stati ausiliari dei due contributi, consentendo la loro interferenza. Misurare o azzerare in modo irreversibile l’ausiliario non equivale a questa operazione: può lasciare una traccia dell’ingresso nell’apparato o nell’ambiente. Il ripristino usato nei circuiti di Grover deve quindi essere un’inversione coerente del calcolo, non una semplice cancellazione classica di memoria.

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \xrightarrow{\text{CNOT}} |+\rangle|0\rangle \xrightarrow{H\otimes I} |00\rangle$$

Richiami La CNOT è la propria inversa; $H|+\rangle = |0\rangle$.

I contributi separati dagli ausiliari

Per calcolare l'effetto di un ripristino omesso, dividiamo le componenti secondo gli ausiliari a, t. Quattro ingressi hanno ausiliari 00: $000, 001, 110, 111$. Due hanno 10: $010, 100$. Le due soluzioni $011, 101$ hanno 11 e ricevono il segno negativo. Inizialmente ogni ampiezza dei dati ha modulo $1/\sqrt{8}$.

Nel diffusore su otto candidati, il coefficiente che collega un ingresso diverso dall'uscita è $2/8 = 1/4$, mentre quello sulla diagonale è $1/4 - 1 = -3/4$. Fissiamo come uscita una soluzione, per esempio 011. Il settore 00 contribuisce con quattro termini positivi da $1/(4\sqrt{8})$. Il settore 10 ne contribuisce con due. Nel settore 11, il termine diagonale e quello dell'altra soluzione partono entrambi negativi e producono i coefficienti riportati nella formula. I tre risultati rimangono associati ad ausiliari distinti.

$$a_{00} = \frac{4}{4\sqrt{8}} = \frac{1}{\sqrt{8}}, \qquad a_{10} = \frac{2}{4\sqrt{8}} = \frac{1}{2\sqrt{8}}$$

$$a_{11} = \left(-\frac{3}{4}\right)\left(-\frac{1}{\sqrt{8}}\right) + \frac{1}{4}\left(-\frac{1}{\sqrt{8}}\right) = \frac{1}{2\sqrt{8}}$$

Richiami Per $N = 8$, gli elementi diagonali di $D = 2|s\rangle\langle s| - I$ valgono $-3/4$ e quelli fuori diagonale valgono $1/4$.

Il costo di un ripristino omesso

Nel predicato su tre bit, omettiamo il ripristino dopo la porta Z e applichiamo D soltanto ai dati. Per uno dei due esiti validi, 011 oppure 101, rimangono tre componenti con ausiliari diversi. Le loro ampiezze sono riportate nella tabella e le loro probabilità vanno sommate, perché gli stati ausiliari sono ortogonali.

La probabilità di ciascuna soluzione vale così $1/8 + 1/32 + 1/32 = 3/16$. Il successo complessivo è $3/8$, mentre il circuito con ausiliari puliti raggiungerà 1 nello stesso esempio. Il calcolo assume lo stato iniziale uniforme e un'unica marcatura seguita da D. Mostra quantitativamente che ignorare gli ausiliari residui altera l'algoritmo, anche quando il valore del predicato era stato calcolato correttamente.

Ausiliari a, t	Ampiezza	Probabilità
00	$1/\sqrt{8}$	$1/8$
10	$1/(2\sqrt{8})$	$1/32$
11	$1/(2\sqrt{8})$	$1/32$

La tabella riguarda un singolo esito valido dopo il diffusore. L'altro esito valido ha gli stessi coefficienti, per simmetria del predicato e della preparazione.

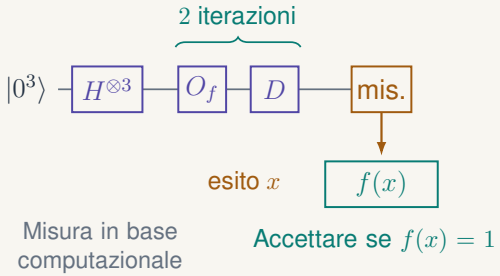
Richiami Con ausiliari distinti, si sommano i moduli quadrati dei coefficienti, non i coefficienti stessi.

Il circuito della ricerca e la verifica finale

Il circuito completo prepara il registro dei dati nello stato uniforme e ripete il blocco oracolo-diffusore per il numero di iterazioni stabilito. L'oracolo deve restituire puliti i propri ausiliari, così il blocco successivo agisce sullo stato dei dati previsto dal modello. Nell'esempio con una soluzione fra otto candidati scegliamo due iterazioni.

Alla fine misuriamo i dati e controlliamo classicamente il candidato con f. La misura produce una stringa, non il valore della probabilità di successo. Un esito non valido richiede una nuova preparazione se decidiamo di riprovare. Le due interrogazioni quantistiche sono applicazioni dell'intera marcatura e non includono la verifica classica finale. Il numero di porte necessarie per realizzare una marcatura dipende dal verificatore e può essere molto maggiore di due.

Preparazione, iterazioni e verifica



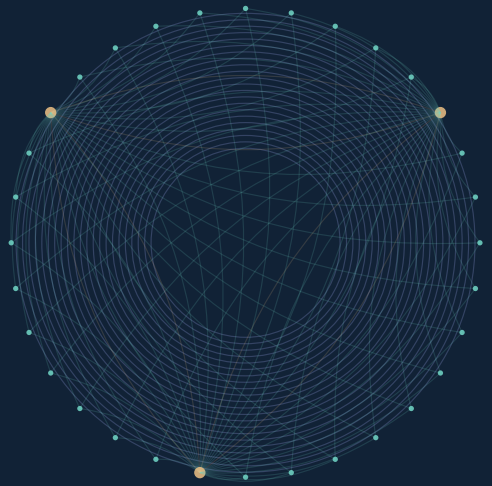
La parte quantistica termina con la misura. Il controllo del candidato usa il predicato sul risultato classico, senza recuperare le ampiezze dello stato precedente.

Richiami Un'iterazione è $G = DO_f$, con oracolo per primo. La misura restituisce una stringa che va verificata.

CAPITOLO 09

Più soluzioni

La probabilità complessiva e i singoli esiti



L’algoritmo di Grover

La ricerca di una qualunque soluzione

Quando il predicato ammette più soluzioni, il compito di ricerca consiste anzitutto nel restituirne una. Misurare una qualsiasi stringa con $f(x) = 1$ conta come successo. Se la preparazione è uniforme, ciascuna stringa ha probabilità $1/N$ e le M soluzioni hanno probabilità complessiva M/N.

L'oracolo cambia segno a tutte le soluzioni e il diffusore ne amplifica la componente comune. La formula geometrica continua a valere, usando $\sin\theta = \sqrt{M/N}$. Un numero maggiore di soluzioni modifica quindi l'angolo iniziale e la rotazione di ogni iterazione. Il numero di passi scelto per un unico bersaglio non può essere riutilizzato automaticamente. Studieremo prima il predicato con due soluzioni, poi casi in cui la frazione delle soluzioni è più grande.

Richiami Per $0 < M < N$, $P_k = \sin^2((2k+1)\theta)$ con $\sin\theta = \sqrt{M/N}$.

Le due soluzioni del predicato su tre bit

Il predicato $(x_2 \oplus x_1) \wedge x_0$ ha due soluzioni, 011 e 101. La loro sovrapposizione uniforme normalizzata è $|g\rangle = (|011\rangle + |101\rangle)/\sqrt{2}$. Le altre sei stringhe formano lo stato $|b\rangle$, con coefficienti $1/\sqrt{6}$.

La preparazione uniforme su otto candidati ha componente $\sqrt{2/8} = 1/2$ lungo $|g\rangle$ e componente $\sqrt{6/8} = \sqrt{3}/2$ lungo $|b\rangle$. L'angolo iniziale è quindi $\theta = \pi/6$. È lo stesso angolo incontrato con una soluzione fra quattro candidati, perché in entrambi i casi le soluzioni occupano un quarto dello spazio di ricerca. La dinamica nel piano dipende da questa frazione, benché il numero degli esiti della misura sia diverso.

8 candidati, 3 bit

000 $f = 0$	001 $f = 0$	010 $f = 0$	011 $f = 1$
100 $f = 0$	101 $f = 1$	110 $f = 0$	111 $f = 0$

Ocra: $f(x) = 1$

Le due stringhe evidenziate soddisfano lo stesso predicato. Il successo della ricerca comprende entrambi gli esiti, senza richiedere che il circuito ne privilegi uno.

Richiami $|g\rangle$ contiene soltanto le soluzioni ed è normalizzato; il suo coefficiente quadrato dà la probabilità totale di successo.

La prima iterazione con due soluzioni

Le otto ampiezze iniziali valgono $1/\sqrt{8}$. Dopo l'oracolo, due diventano negative e sei rimangono positive. La media è quindi $(6-2)/(8\sqrt{8}) = 1/(2\sqrt{8})$. Il doppio della media vale $1/\sqrt{8}$.

Il diffusore manda ciascuna ampiezza non valida in $1/\sqrt{8} - 1/\sqrt{8} = 0$. Per ciascuna soluzione sottrae invece l'ampiezza negativa e ottiene $2/\sqrt{8} = 1/\sqrt{2}$. Lo stato finale contiene dunque soltanto 011 e 101. Il calcolo assume che gli ausiliari dell'oracolo siano stati ripristinati: in caso contrario i contributi possono rimanere separati, come mostrato nel capitolo precedente. La normalizzazione finale si verifica sommando due probabilità uguali a $1/2$.

$$\bar{a} = \frac{6-2}{8\sqrt{8}} = \frac{1}{2\sqrt{8}}$$

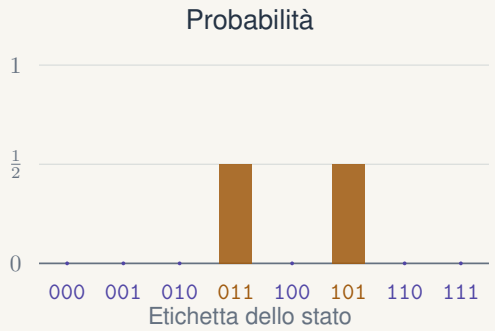
$$a_{\text{sol}} = \frac{2}{\sqrt{8}} = \frac{1}{\sqrt{2}}, \qquad a_{\text{altre}} = 0$$

Richiami Il diffusore trasforma ogni ampiezza a_x in $2\bar{a} - a_x$, con $\bar{a}$ media delle ampiezze correnti.

Successo certo con due esiti possibili

Dopo una iterazione, lo stato dei dati è $(|011\rangle + |101\rangle)/\sqrt{2}$. La misura restituisce 011 con probabilità $1/2$ e 101 con probabilità $1/2$. Ogni risultato possibile soddisfa il predicato, quindi la probabilità totale di successo è 1.

Il successo certo non significa che la misura restituisca entrambe le soluzioni. Una singola esecuzione produce una sola stringa. Anche ripetendo la procedura possiamo ottenere più volte la stessa soluzione. La verifica classica conferma che il risultato è valido, senza ricostruire dall'esito l'intera sovrapposizione precedente. Nel caso ideale qui calcolato non compaiono altri candidati, mentre un dispositivo reale potrebbe introdurre errori che questo modello non include.



Le due barre sommano a 1 e corrispondono a due esiti alternativi. Il grafico rappresenta le probabilità ideali, non un campione di misure sperimentali.

Richiami Le probabilità di esiti distinti si sommano; misurare un registro produce una sola stringa per esecuzione.

Una seconda iterazione riduce il successo

Con $N = 8$ e $M = 2$ abbiamo $\theta = \pi/6$. Dopo una iterazione il vettore raggiunge $3\theta = \pi/2$ e il successo è certo. Dopo due iterazioni arriva invece a $5\theta = 5\pi/6$, per cui $P_2 = \sin^2(5\pi/6) = 1/4$.

Le due soluzioni hanno allora ampiezza $1/\sqrt{8}$, mentre i sei candidati non validi hanno ampiezza $-1/\sqrt{8}$. Tutte le probabilità individuali tornano a $1/8$, pur con segni differenti dallo stato iniziale. Il numero di iterazioni che era opportuno per un unico bersaglio fra otto candidati è dunque sbagliato per questo predicato. Il circuito usa lo stesso tipo di operazioni, ma il valore di M cambia la scelta del momento in cui misurare.



Il primo passo raggiunge probabilità 1. Il passo successivo torna a 1/4: l'amplificazione standard non conserva automaticamente una soluzione già raggiunta.

Richiami $P_k = \sin^2((2k+1)\theta)$; per $M/N = 1/4$, $\theta = \pi/6$.

La ricorrenza per le due ampiezze

Finché le soluzioni hanno tutte la stessa ampiezza, possiamo indicarla con a_g; chiamiamo a_b l'ampiezza comune degli altri candidati. Sono ampiezze di singole stringhe, non le coordinate aggregate del piano. La normalizzazione richiede $M|a_g|^2 + (N-M)|a_b|^2 = 1$.

L'oracolo sostituisce a_g con $-a_g$ e lascia invariato a_b. La media dopo la marcatura è perciò $\mu = ((N-M)a_b - Ma_g)/N$. Applicando il diffusore otteniamo le nuove ampiezze $a_g' = 2\mu + a_g$ e $a_b' = 2\mu - a_b$. Il segno più nella prima formula deriva dal fatto che stiamo sottraendo l'ampiezza già negata dall'oracolo. Questa ricorrenza permette di verificare i calcoli geometrici usando soltanto somme e prodotti.

$$\mu = \frac{(N-M)a_b - Ma_g}{N}$$

$$a_g' = 2\mu + a_g, \qquad a_b' = 2\mu - a_b$$

Richiami a_g e a_b sono ampiezze individuali; le coordinate aggregate sono rispettivamente $\sqrt{M}\,a_g$ e $\sqrt{N-M}\,a_b$.

Tre soluzioni fra sedici candidati

Consideriamo ora un predicato su quattro bit con tre soluzioni. Per verificare i numeri possiamo sceglierle come 0010, 0111 e 1100; il calcolo dell'amplificazione dipende dal loro numero, non dalle etichette. La preparazione uniforme assegna a ogni stringa ampiezza $1/4$.

Dopo la marcatura rimangono tredici ampiezze positive e tre negative. La media vale $(13-3)/(16\cdot 4) = 5/32$, quindi il doppio della media è $5/16$. Ogni soluzione riceve ampiezza $5/16 + 1/4 = 9/16$, mentre gli altri candidati ricevono $5/16 - 1/4 = 1/16$. Una sola iterazione concentra quindi gran parte della probabilità sulle tre soluzioni, pur senza eliminare completamente gli esiti non validi.

$$\bar{a} = \frac{13(1/4) + 3(-1/4)}{16} = \frac{5}{32}$$

$$a_{\text{sol}} = \frac{9}{16}, \qquad a_{\text{altre}} = \frac{1}{16}$$

Richiami La preparazione uniforme su $N = 16$ assegna ampiezza $1/\sqrt{16} = 1/4$ a ogni candidato.

Il conteggio delle probabilità nei due gruppi

Nell'esempio con tre soluzioni fra sedici candidati, ogni soluzione ha probabilità $(9/16)^2 = 81/256$. Le tre probabilità si sommano e danno $P_1 = 243/256$, circa $0{,}9492$. Ciascuno dei tredici candidati non validi ha invece probabilità $(1/16)^2 = 1/256$.

La probabilità totale di fallimento è perciò $13/256$. Sommando successo e fallimento otteniamo $256/256 = 1$, come richiesto dalla normalizzazione. Il controllo distingue due operazioni diverse: elevare al quadrato l'ampiezza di un singolo candidato e moltiplicare poi per il numero dei candidati del gruppo. Elevare al quadrato la somma delle tre ampiezze valide darebbe un risultato sbagliato, perché le tre stringhe sono esiti ortogonali della misura.

Gruppo	Numero	Probabilità per esito
Soluzioni	3	$81/256$
Altri candidati	13	$1/256$

$$P_1 = \frac{243}{256}$$

$$\frac{243}{256} + \frac{13}{256} = 1$$

Il numero degli elementi moltiplica la probabilità individuale. La somma dei due contributi verifica che nessuna probabilità sia stata persa o contata due volte.

Richiami Per esiti distinti si sommano i moduli quadrati; non si eleva al quadrato la somma delle loro ampiezze.

La distribuzione fra le soluzioni

La preparazione uniforme assegna la stessa ampiezza a tutte le soluzioni. L'oracolo applica loro lo stesso cambio di segno e il diffusore usa la stessa media per ciascuna. Questa simmetria si conserva a ogni iterazione. Ogni soluzione ha quindi probabilità P_k/M.

Se $P_k > 0$ e consideriamo soltanto le esecuzioni riuscite, ciascuna delle M soluzioni compare con probabilità $1/M$. Nel caso con tre soluzioni appena studiato, il rapporto fra $81/256$ e $243/256$ vale infatti $1/3$. Il condizionamento significa che scartiamo gli esiti non validi e rinormalizziamo le probabilità rimanenti. Questa uniformità dipende dalla preparazione uniforme; una preparazione iniziale differente richiederà l'analisi dell'amplificazione di ampiezza nel capitolo dedicato.

$$\frac{P_k/M}{P_k} = \frac{1}{M} \quad (P_k > 0)$$

$$\frac{81/256}{243/256} = \frac{1}{3}$$

Richiami Condizionare al successo significa dividere la probabilità di un esito valido per la probabilità totale P_k, purché $P_k > 0$.

Il caso con metà dei candidati validi

Se $M = N/2$, la probabilità iniziale di successo è già $1/2$ e $\theta = \pi/4$. La formula dà $P_k = \sin^2((2k+1)\pi/4)$. Per ogni intero k, l'angolo è un multiplo dispari di 45 gradi e il seno quadro vale sempre $1/2$.

Le iterazioni standard non migliorano il successo. Possiamo campionare subito un candidato e verificarlo. Indichiamo con E il numero medio di tentativi fino al primo successo, dove ogni tentativo indipendente riesce con probabilità p. Il primo tentativo costa 1. Con probabilità $1-p$ fallisce e ripartiamo dalla stessa situazione, che richiede ancora in media E tentativi. Pertanto $E = 1 + (1-p)E$ e, risolvendo, $E = 1/p$. Con $p = 1/2$ otteniamo due tentativi in media, senza amplificazione.

$$P_k = \sin^2\left((2k+1)\frac{\pi}{4}\right) = \frac{1}{2}$$

$$E = 1 + (1-p)E \quad \Longrightarrow \quad E = \frac{1}{p}$$

Richiami Un tentativo uniforme ha probabilità M/N di successo; per tentativi indipendenti con successo $p > 0$ il numero medio è $1/p$.

Un predicato senza soluzioni

Se $M = 0$, il predicato restituisce zero su ogni candidato. L'oracolo di fase è allora l'identità: nessuna componente riceve un meno. Lo stato uniforme rimane invariato anche sotto D, poiché il diffusore conserva la direzione $|s\rangle$. Qualunque numero di iterazioni lascia quindi inalterata la distribuzione uniforme.

Ogni candidato misurato fallisce la verifica, perché una soluzione non esiste. Il vettore $|g\rangle$ non è definito in questo caso: la sua formula dividerebbe per $\sqrt{M} = 0$. Per questo lo trattiamo separatamente dalla descrizione geometrica con due gruppi non vuoti. Finché non abbiamo controllato l'intero insieme dei candidati, una sequenza di tentativi falliti non certifica che $M = 0$, salvo altre informazioni sul problema. La procedura per un numero ignoto di soluzioni dovrà specificare questa limitazione.

$$M = 0 : \qquad O_f = I, \quad G^k|s\rangle = |s\rangle, \quad P_k = 0$$

Richiami $D|s\rangle = |s\rangle$; $|g\rangle = M^{-1/2} \sum_{f(x)=1} |x\rangle$ richiede $M > 0$.

Un predicato soddisfatto da ogni candidato

Se $M = N$, ogni stringa è una soluzione e la preparazione uniforme ha già probabilità di successo 1. L'oracolo cambia segno a tutte le componenti, quindi è $-I$. Dopo l'oracolo il diffusore agisce su $-|s\rangle$ e lo lascia su quella stessa direzione.

Dopo k iterazioni lo stato è $(-1)^k|s\rangle$: cambia soltanto una fase globale, mentre ogni misura continua a produrre una soluzione. Non occorre eseguire alcuna iterazione per raggiungere il risultato. Il vettore $|b\rangle$ non è definito perché non esistono candidati non validi e la sua normalizzazione richiederebbe una divisione per zero. Anche questo caso va separato dalle formule che assumono esplicitamente $0 < M < N$.

$$M = N: \qquad O_f = -I, \quad G^k|s\rangle = (-1)^k|s\rangle, \quad P_k = 1$$

Richiami Una fase globale non modifica le probabilità; $|b\rangle$ richiede $N - M > 0$.

L’amplificazione può peggiorare un caso facile

Consideriamo sei soluzioni fra otto candidati. La misura immediata ha successo con probabilità $6/8 = 3/4$. L’angolo iniziale soddisfa $\sin\theta = \sqrt{3}/2$, quindi $\theta = \pi/3$. Dopo una iterazione l’angolo è $3\theta = \pi$ e la probabilità di successo scende a zero.

Il calcolo delle ampiezze conferma questo risultato: dopo l’oracolo la media è $-1/(2\sqrt{8})$. Il diffusore porta le sei ampiezze valide a zero e le due non valide a $-1/\sqrt{2}$. La rotazione standard ha superato la direzione utile. Il primo massimo continuo si trova a $k_* = 1/4$, perciò l’intero più vicino è $k = 0$. Per questo caso conviene misurare subito e verificare, anziché applicare automaticamente una iterazione.

$$P_0 = \frac{3}{4}, \qquad P_1 = \sin^2\pi = 0$$

Richiami $k_* = \pi/(4\theta) - 1/2$; scegliere $k = 0$ significa usare soltanto la preparazione iniziale.

Trovare una soluzione e trovarle tutte

La ricerca di Grover produce una soluzione per esecuzione riuscita. Se vogliamo ottenere tutte le soluzioni, le ripetizioni possono restituire stringhe già osservate. Con due soluzioni equiprobabili, dopo averne trovata una la successiva esecuzione riuscita ha probabilità $1/2$ di mostrarci l'altra.

In questo esempio occorrono in media altre due esecuzioni riuscite per vedere la soluzione mancante, quindi tre in totale partendo da zero. Il conteggio riguarda gli esiti validi e non include eventuali tentativi falliti. Un'altra possibilità è modificare il predicato per escludere le soluzioni già note, ma ciò cambia il numero M e può aumentare il costo del verificatore. Enumerare tutte le soluzioni richiede pertanto una procedura e un'analisi distinte dal semplice compito di trovarne una.

Richiami Con preparazione uniforme, le M soluzioni sono equiprobabili condizionatamente a un successo di probabilità non nulla.

Esercizio con due soluzioni fra sedici candidati

Un predicato su quattro bit ha esattamente due soluzioni. Il registro parte dallo stato uniforme e gli ausiliari dell'oracolo vengono ripristinati correttamente. Determinare l'angolo iniziale, scegliere il numero di iterazioni vicino al primo massimo e calcolare la probabilità totale di successo.

Calcolare inoltre la probabilità di ciascuna soluzione e quella di ciascuno dei quattordici candidati non validi al termine delle iterazioni scelte. Per il primo passaggio è utile confrontare $M/N = 2/16$ con l'esempio già svolto su otto candidati. Per distribuire la probabilità fra le singole stringhe occorre invece usare i nuovi numeri degli elementi dei due gruppi. La soluzione nella pagina seguente separa esplicitamente questi due aspetti.

Richiami $\sin\theta = \sqrt{M/N}$; una coordinata aggregata si divide per la radice del numero di elementi del gruppo per ottenere l'ampiezza individuale.

Soluzione dell'esercizio con due candidati validi

La frazione $M/N = 2/16 = 1/8$ coincide con quella dell'esempio con un bersaglio fra otto candidati. Abbiamo quindi lo stesso angolo $\theta = \arcsin(1/\sqrt{8})$ e scegliamo ancora due iterazioni. La probabilità totale di successo è $121/128$ e quella di fallimento è $7/128$.

Le due soluzioni condividono la probabilità di successo: ciascuna ha $121/256$. I quattordici candidati non validi condividono invece $7/128$, quindi ciascuno ha $1/256$. Le rispettive ampiezze sono $11/16$ e $-1/16$, coerenti con il segno del coseno dopo la seconda iterazione. La normalizzazione si verifica come $2(121/256) + 14(1/256) = 1$. La stessa traiettoria nel piano può quindi corrispondere a distribuzioni individuali differenti, perché cambia il numero degli stati di base in ciascun gruppo.

$$\frac{121/128}{2} = \frac{121}{256}, \qquad \frac{7/128}{14} = \frac{1}{256}$$

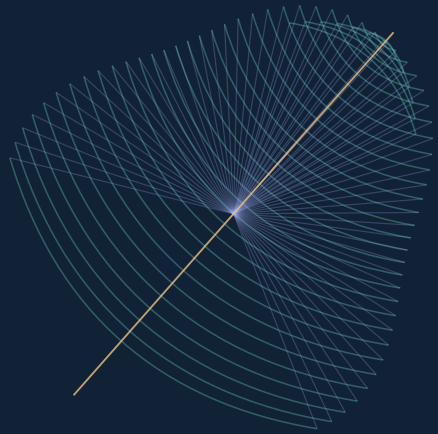
$$2\left(\frac{11}{16}\right)^2 + 14\left(-\frac{1}{16}\right)^2 = 1$$

Richiami P_k è la probabilità totale sulle soluzioni; la probabilità di una singola soluzione, con preparazione uniforme, vale P_k/M.

CAPITOLO 10

Il numero di soluzioni ignoto

Conteggi casuali, verifiche e nuovi tentativi



L’algoritmo di Grover

La ricerca senza conoscere il numero di soluzioni

La scelta del numero di iterazioni dipende da M, perché l'angolo iniziale soddisfa $\sin^2\theta = M/N$. In un problema di ricerca conosciamo però il verificatore f e possiamo ignorare quante stringhe lo soddisfino. Contare tutte le soluzioni prima di cercarne una richiederebbe, in generale, un altro algoritmo. La formula del primo picco non risolve questa difficoltà: contiene proprio il dato che manca.

Occorre quindi scegliere la durata dei tentativi senza pretendere di conoscere l'angolo. Manteniamo la preparazione uniforme, l'oracolo e il diffusore già costruiti. Modifichiamo il controllo classico che decide quante iterazioni eseguire prima della misura. Ogni risultato viene verificato e un tentativo fallito riparte da una nuova preparazione.

Richiami N conta i candidati e M quelli per cui $f(x) = 1$. Dopo k iterazioni, $P_k = \sin^2((2k+1)\theta)$ con $\sin^2\theta = M/N$.

La stessa durata può favorire problemi diversi

Consideriamo otto candidati. Con una soluzione, due iterazioni danno probabilità $121/128$, quindi una scelta molto favorevole. Con due soluzioni, invece, l'angolo è $\pi/6$ e le stesse due iterazioni danno $\sin^2(5\pi/6) = 1/4$. In quest'ultimo problema sarebbe bastata una sola iterazione per raggiungere probabilità 1.

La differenza riguarda il numero delle soluzioni, anche se il registro ha sempre tre qubit e il circuito usa gli stessi tipi di operazione. Scegliere sistematicamente la durata adatta a una soluzione può quindi oltrepassare un picco relativo a un altro valore di M. I tentativi di durata variabile servono a evitare questa dipendenza da una scelta unica.

M	P_1	P_2
1	$25/32$	$121/128$
2	1	$1/4$

Le probabilità sono teoriche. Il confronto mantiene fisso il numero di candidati e cambia soltanto il numero delle soluzioni.

Richiami Per $N = 8, M = 1$, il primo picco discreto si ottiene con $k = 2$. Per $M/N = 1/4$, $\theta = \pi/6$ e $P_1 = 1$.

Un tentativo con un numero casuale di iterazioni

Fissiamo un intero positivo J e scegliamo un intero j uniformemente fra $0, 1, \ldots, J-1$. Uniformemente significa che ogni scelta ha probabilità $1/J$. Prepariamo lo stato uniforme, eseguiamo j iterazioni e misuriamo. Il caso $j = 0$ è incluso: misura direttamente la preparazione iniziale, senza alcuna amplificazione.

La scelta di j avviene classicamente prima del circuito. Non richiede un registro quantistico che contenga contemporaneamente tutte le durate. Dopo la misura calcoliamo $f(x)$ sul risultato ottenuto. Se il risultato non è una soluzione, il tentativo successivo usa un nuovo stato iniziale e una nuova scelta casuale. Le durate diverse contribuiscono così alla probabilità complessiva attraverso una media.

Richiami Un'iterazione è $G = DO_f$: prima la marcatura, poi la riflessione rispetto allo stato uniforme. La verifica accetta x soltanto quando $f(x) = 1$.

La probabilità media di un tentativo

Per una durata fissata j, la probabilità di trovare una soluzione è P_j. Se tutte le J durate sono equiprobabili, la probabilità del tentativo casuale è la somma dei prodotti $(1/J)P_j$. Indichiamo questa media con $\overline{P}_J$. Qui la barra rappresenta una media di probabilità, distinta dalla media delle ampiezze usata dal diffusore.

Per esempio, scegliendo fra zero e una iterazione nel problema con otto candidati e una soluzione, otteniamo $(1/8 + 25/32)/2 = 29/64$. Il valore supera il successo della misura iniziale, pur senza scegliere soltanto la durata migliore. Questa media descrive un insieme di esperimenti alternativi: non si sommano le ampiezze di circuiti diversi.

$$\overline{P}_J = \frac{1}{J} \sum_{j=0}^{J-1} P_j$$

$$\overline{P}_2 = \frac{1}{2}\left(\frac{1}{8} + \frac{25}{32}\right) = \frac{29}{64}$$

Richiami J è il numero intero di durate possibili, da 0 a $J-1$. La media di probabilità pesa ogni esperimento con la probabilità di sceglierlo.

Quattro durate nel problema su otto candidati

Con $J = 4$ scegliamo fra zero, una, due e tre iterazioni. Le probabilità calcolate nei capitoli precedenti permettono una verifica esatta della media. Il risultato è $1117/2048$, circa $0{,}5454$. La probabilità media resta quindi superiore a un mezzo, benché una delle quattro durate abbia già oltrepassato il primo picco.

Questo esempio non prescrive $J = 4$ per ogni problema. Mostra soltanto come una scelta casuale possa combinare tentativi molto favorevoli e tentativi meno favorevoli. Se M è ignoto, serve ancora un metodo per stabilire quanto ampia debba diventare la finestra delle durate. La formula della media ci permetterà di giustificare una crescita graduale.

j	P_j
0	$1/8$
1	$25/32$
2	$121/128$
3	$169/512$

$$\overline{P}_4 = \frac{1117}{2048}$$

Richiami Le probabilità si riferiscono a nuove preparazioni indipendenti dello stesso stato uniforme. $\overline{P}_J = (P_0 + \cdots + P_{J-1})/J$.

La media espressa attraverso l'angolo iniziale

Sostituiamo nella media la formula $P_j = \sin^2((2j+1)\theta)$. L'identità $\sin^2 u = (1-\cos 2u)/2$ separa ogni termine in una parte costante e una parte oscillante. La media delle parti costanti vale $1/2$. Resta da calcolare la somma dei coseni, i cui angoli aumentano ogni volta di 4θ.

Il contributo medio delle oscillazioni non diminuisce necessariamente a ogni ampliamento della finestra. Il suo valore assoluto è però limitato da una quantità che decresce al crescere di J, come mostrerà il calcolo seguente. Non occorre che tutte le durate siano ugualmente buone. Supponiamo $0 < M < N$, così l'angolo è compreso strettamente fra zero e $\pi/2$.

$$\overline{P}_J = \frac{1}{2} - \frac{1}{2J} \sum_{j=0}^{J-1} \cos((4j+2)\theta)$$

Richiami $\sin^2 u + \cos^2 u = 1$ e $\cos(2u) = 1 - 2\sin^2 u$. $0 < M < N$ implica $0 < \theta < \pi/2$.

La cancellazione dei termini interni

Usiamo l'identità $2\sin v\cos u = \sin(u+v) - \sin(u-v)$ con $v = 2\theta$ e $u = (4j+2)\theta$. Ogni coseno diventa una differenza fra due seni consecutivi. Sommando per tutti i valori di j, il termine negativo di una differenza cancella il termine positivo della precedente. Rimangono soltanto $\sin(4J\theta)$ e $-\sin 0$.

Poiché $\sin 0 = 0$, possiamo sostituire il risultato nella media. Il denominatore $\sin(2\theta)$ è diverso da zero sotto l'ipotesi $0 < M < N$. La formula ottenuta vale per un numero intero positivo J di scelte. I casi senza soluzioni o con tutti i candidati validi si trattano direttamente e non richiedono questa divisione.

$$2\sin(2\theta)\sum_{j=0}^{J-1}\cos((4j+2)\theta) = \sin(4J\theta)$$

$$\overline{P}_J = \frac{1}{2} - \frac{\sin(4J\theta)}{4J\sin(2\theta)}$$

Richiami Una somma è telescopica quando i termini interni si cancellano a coppie. J conta le durate equiprobabili, mentre θ dipende da M/N.

Una finestra sufficientemente ampia

Il seno al numeratore è sempre compreso fra -1 e 1. Possiamo quindi limitare la distanza della media da $1/2$ senza conoscere il valore preciso dell'oscillazione. Quando $J \sin(2\theta) \geq 1$, tale distanza non supera $1/4$. Ne segue una probabilità di successo almeno uguale a $1/4$.

Per otto candidati e una soluzione, $\sin\theta = 1/\sqrt{8}$ e $\cos\theta = \sqrt{7/8}$. Pertanto $\sin(2\theta) = \sqrt{7}/4$ e la soglia teorica è $4/\sqrt{7} \approx 1{,}512$. Già $J = 2$ la supera. La media esatta $29/64$ soddisfa il limite, che serve come garanzia prudente e non come previsione esatta del successo.

$$\left|\overline{P}_J - \frac{1}{2}\right| \leq \frac{1}{4J\sin(2\theta)}$$

$$J \geq \frac{1}{\sin(2\theta)} \quad \Longrightarrow \quad \overline{P}_J \geq \frac{1}{4}$$

Richiami $\sin(2\theta) = 2\sin\theta\cos\theta$. Una garanzia inferiore può essere minore della probabilità effettiva.

La finestra adattiva e il numero delle scelte

Introduciamo un parametro reale $m \geq 1$, che controllerà l'ampiezza della finestra. Scegliamo j fra gli interi non negativi strettamente minori di m. Il loro numero è $J = \lceil m \rceil$, dove il soffitto indica il più piccolo intero maggiore o uguale a m. Per $m = 1{,}44$, le scelte sono 0 e 1, dunque $J = 2$.

Questa distinzione evita un'ambiguità nei calcoli. Il parametro m cresce moltiplicativamente e può non essere intero. La formula della probabilità media usa invece J, perché somma un numero intero di esperimenti. Per $m = 2$, le scelte rimangono 0 e 1. Per $m = 2{,}01$, si aggiunge la durata 2 e le scelte diventano tre.

$$j \in \{0, 1, \ldots, J-1\}, \qquad J = \lceil m \rceil$$

Richiami $\lceil m \rceil$ arrotonda m per eccesso al primo intero disponibile. La probabilità media $\overline{P}_J$ richiede J intero positivo.

La scelta progressiva della durata

La procedura di Boyer, Brassard, Høyer e Tapp parte da $m = 1$ e usa un fattore di crescita $\lambda = 6/5$. A ogni tentativo sceglie uniformemente un intero $j < m$, prepara lo stato uniforme ed esegue j iterazioni. Dopo la misura verifica il candidato. Se la verifica riesce, termina restituendo quel candidato.

Se la verifica fallisce, sostituisce m con il minimo fra λm e $\sqrt{N}$, quindi avvia un nuovo tentativo. La crescita consente di raggiungere una finestra adatta anche a soluzioni rare. Il limite $\sqrt{N}$ impedisce alle durate di crescere indefinitamente. Nessun passo di questa procedura richiede di conoscere M o di stimarlo prima della ricerca.

$$m \leftarrow 1, \qquad \lambda = \frac{6}{5}$$

$$m \leftarrow \min\{\lambda m, \sqrt{N}\} \quad \text{dopo un fallimento}$$

Richiami Il tentativo usa $J = \lceil m \rceil$ durate equiprobabili. Ogni nuovo tentativo riparte da $|s\rangle$, non dallo stato dopo la misura.

La crescita della finestra per otto candidati

Per $N = 8$, il limite della finestra è $\sqrt{8} \approx 2{,}828$. Se i primi tentativi falliscono, il parametro percorre i valori riportati nella tabella. I valori non descrivono una sequenza obbligatoria di risultati: il procedimento termina appena una verifica riesce. La tabella segue soltanto il ramo in cui continuano i fallimenti.

Inizialmente si misura senza iterazioni. Quando m supera 1, diventano possibili zero o una iterazione. Superato 2, si può scegliere anche la durata 2, favorevole nel caso di una sola soluzione. Raggiunto il limite, si continua a estrarre nuove durate nella stessa finestra. Il numero delle scelte dipende dal soffitto di m, non dal suo arrotondamento all'intero più vicino.

m	$J = \lceil m \rceil$
1	1
$1{,}2$	2
$1{,}44$	2
$1{,}728$	2
$2{,}0736$	3
$2{,}48832$	3
$\sqrt{8}$	3

Richiami j è estratto fra 0 e $J - 1$, estremi inclusi. m cresce solo dopo un tentativo fallito.

Un test iniziale per le soluzioni frequenti

Quando le soluzioni costituiscono gran parte dei candidati, un test classico uniforme ha già una probabilità elevata di successo. Possiamo perciò anteporre a ogni tentativo un campionamento classico: scegliamo una stringa uniforme e calcoliamo f. Se il valore è 1, terminiamo subito. Altrimenti eseguiamo il tentativo quantistico previsto dalla finestra corrente.

Con $M > 3N/4$, questo test riesce con probabilità maggiore di $3/4$ a ogni giro. Il suo costo atteso rimane costante. Per le soluzioni rare, aggiunge una sola verifica per giro senza cambiare l'ordine asintotico della procedura adattiva. L'aggiunta permette di usare una sola procedura anche quando ignoriamo se il problema abbia poche o molte soluzioni.

$$P(\text{test uniforme riuscito}) = \frac{M}{N}$$

Richiami Un campione uniforme assegna probabilità $1/N$ a ciascun candidato. La verifica classica restituisce esattamente $f(x)$.

Il costo atteso della crescita

Per $1 \leq M \leq 3N/4$, una finestra dell'ordine di $1/\sin(2\theta)$ è sufficiente a garantire una probabilità costante di successo. Le finestre precedenti crescono geometricamente, quindi la somma dei loro costi rimane proporzionale all'ultima scala raggiunta. Dopo tale soglia, la probabilità di dover continuare per altri u tentativi è al più $(3/4)^u$.

Il costo dei tentativi può crescere come λ^u. Il prodotto è una progressione di ragione al più $3\lambda/4$, inferiore a 1 per $\lambda = 6/5$. Questi contributi spiegano il costo atteso $O(\sqrt{N/M})$: la notazione O indica un limite superiore proporzionale alla funzione scritta, a meno di costanti nel regime considerato. Atteso significa mediato sulle scelte casuali e sugli esiti delle misure, non limitato deterministicamente in ogni esecuzione.

$$\frac{1}{\sin(2\theta)} = \frac{N}{2\sqrt{M(N-M)}}$$

$$\frac{3\lambda}{4} = \frac{9}{10} < 1$$

Richiami Una progressione geometrica di ragione r con $0 \leq r < 1$ ha somma finita. Il valore atteso pesa ogni costo con la sua probabilità.

I fallimenti e l'assenza di soluzioni

Se $M = 0$, ogni verifica restituisce 0 e una ricerca che attende soltanto il successo non termina. Occorre quindi stabilire un limite alle risorse. Raggiunto il limite, il risultato corretto da comunicare è che nessuna soluzione è stata trovata entro quel budget. Un fallimento isolato non dimostra che il predicato sia falso su tutti i candidati.

Per decidere probabilisticamente se esistano soluzioni, il budget deve derivare da una garanzia di errore sotto l'ipotesi $M \geq 1$. Una risposta negativa può allora mancare una soluzione, mentre ogni soluzione restituita viene verificata. Si parla di errore unilaterale perché può essere sbagliata soltanto la risposta negativa. Essa resta diversa da una dimostrazione deterministica di assenza, come quella ottenuta verificando tutti i candidati.

Richiami $M = 0$ significa $f(x) = 0$ per ogni candidato x. Un errore unilaterale ammette false risposte negative, mentre ogni soluzione restituita viene verificata.

Il controllo della probabilità di fallimento

Supponiamo di usare tentativi indipendenti ciascuno con successo almeno $1/4$. La probabilità che falliscano tutti e L è allora al più $(3/4)^L$. Per esempio, 17 tentativi danno un limite inferiore a $0{,}01$. Per applicare questo conto occorre però aver raggiunto una finestra che garantisca quella probabilità, oppure usare una procedura la cui analisi fornisca un limite equivalente.

Un limite prudente basato su $M \geq 1$ porta a un costo dell'ordine di $\sqrt{N}$ per una probabilità costante di successo, con un fattore aggiuntivo dell'ordine di $\log(1/\delta)$ per ridurre l'errore sotto δ. Questa garanzia si riferisce all'algoritmo ideale. Errori del circuito o del verificatore richiederebbero un'analisi ulteriore.

$$P(L \text{ fallimenti}) \leq \left(\frac{3}{4}\right)^L$$

$$\left(\frac{3}{4}\right)^{17} \approx 0{,}00752$$

Richiami δ indica la soglia ammessa per la probabilità di mancare una soluzione esistente. L'indipendenza richiede nuove preparazioni e nuove scelte a ogni tentativo.

Esercizio sulla scelta casuale della durata

Riprendiamo il problema con otto candidati e una soluzione. Un tentativo sceglie uniformemente j fra 0, 1, 2 e 3. Le probabilità corrispondenti sono $1/8$, $25/32$, $121/128$ e $169/512$. Si calcolino il numero medio di iterazioni e il numero medio di interrogazioni contando anche una verifica finale.

Si confronti inoltre la probabilità di successo del tentativo con quella della durata fissa $j = 2$. Il confronto deve distinguere il costo di un tentativo dal costo necessario per ottenere effettivamente una soluzione. Nel secondo caso bisogna considerare anche le eventuali ripetizioni. Per questo esercizio si assume che ogni applicazione dell'oracolo di fase e ogni verifica classica costino una interrogazione.

Richiami $E[j] = \sum_j j\, P(j)$ è il valore atteso della durata. La verifica finale costa una chiamata al predicato anche quando restituisce 0.

Soluzione: costo di un tentativo e costo dei riavvii

La durata media è $(0+1+2+3)/4 = 3/2$. Aggiungendo la verifica si ottengono $5/2$ interrogazioni per tentativo. Il successo medio è $1117/2048$, inferiore al $121/128$ ottenuto scegliendo sempre due iterazioni. Quando possiamo conoscere $M = 1$, la durata fissa risulta qui più favorevole sia nel successo sia nel costo medio fino alla soluzione.

Con riavvii indipendenti, il costo atteso complessivo è il costo medio di un tentativo diviso per la sua probabilità di successo. La scelta casuale dà $5120/1117 \approx 4{,}58$ interrogazioni, mentre la durata fissa ne dà $384/121 \approx 3{,}17$. La procedura casuale serve a gestire l'informazione mancante su M, non a migliorare ogni caso in cui M sia già noto.

$$\frac{5/2}{1117/2048} = \frac{5120}{1117}$$

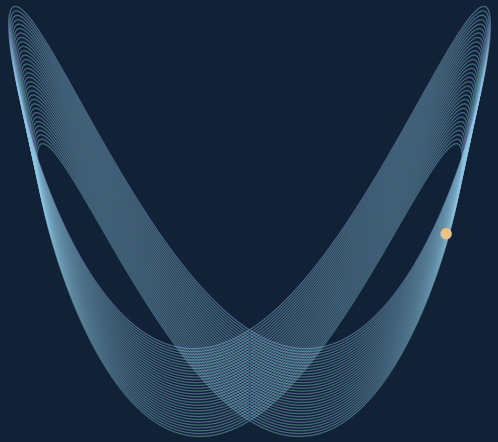
$$\frac{3}{121/128} = \frac{384}{121}$$

Richiami Per tentativi indipendenti ripetuti fino al successo, $E[C_{\text{totale}}] = E[C_{\text{tentativo}}]/p$. p è la probabilità di successo di un tentativo completo.

CAPITOLO 11

Amplificazione di ampiezza

La ricerca con una preparazione non uniforme



L’algoritmo di Grover

Una preparazione iniziale diversa da Hadamard

La preparazione uniforme assegna lo stesso peso a ogni candidato. In alcuni problemi disponiamo invece di un procedimento che produce più spesso candidati promettenti. Descriviamo la sua versione coerente mediante un operatore unitario A, applicato a un registro inizialmente nello stato zero. Lo stato preparato si indica con $|\psi\rangle = A|0\rangle$. Qui $|0\rangle$ rappresenta tutti i qubit del registro nello stato zero.

Il verificatore continua a distinguere le soluzioni mediante $f(x)$. La probabilità iniziale di successo è la somma dei moduli quadrati delle ampiezze delle soluzioni e viene indicata con p. Con una preparazione non uniforme, p non coincide necessariamente con la frazione M/N. L'amplificazione può essere costruita rispetto a questo nuovo stato iniziale.

$$|\psi\rangle = \sum_x \alpha_x |x\rangle, \qquad p = \sum_{f(x)=1} |\alpha_x|^2$$

Richiami Un operatore unitario conserva la norma e ammette un inverso. La probabilità totale di un insieme di esiti è la somma delle probabilità dei suoi elementi.

Una soluzione con probabilità un sedicesimo

Consideriamo quattro stringhe e supponiamo che l'unica soluzione sia 11. Un circuito di preparazione produce ampiezza $\sqrt{5}/4$ per ciascuna delle altre tre stringhe e ampiezza $1/4$ per 11. Lo stato è normalizzato, perché la somma delle probabilità è $3 \cdot 5/16 + 1/16 = 1$.

In questo esempio $M/N = 1/4$, mentre la probabilità iniziale di successo è $p = 1/16$. Il numero dei candidati non basta dunque a stabilire l'angolo dell'amplificazione generale. Serve il peso effettivamente assegnato alle soluzioni dalla preparazione. Assumiamo qui che il circuito A sia disponibile e invertibile, per studiare l'effetto dell'amplificazione separatamente dal costo necessario a realizzarlo.

$$|\psi\rangle = \frac{\sqrt{5}}{4}(|00\rangle + |01\rangle + |10\rangle) + \frac{1}{4}|11\rangle$$

Richiami p è il successo di una misura immediata dopo A. $\sqrt{5}/4$ è un'ampiezza: il suo modulo quadro è $5/16$.

Ripetere la preparazione e verificare il risultato

Una prima strategia consiste nel preparare $A|0\rangle$, misurare e verificare, ripetendo tutto in caso di fallimento. Se ogni tentativo ha successo con probabilità $p > 0$, il numero medio di tentativi fino al primo successo è $1/p$. Per $p = 1/16$, occorrono quindi in media sedici preparazioni, ciascuna seguita da una misura e da una verifica.

La probabilità di fallire i primi L tentativi è $(1-p)^L$, perché le preparazioni ripetute sono indipendenti. L'amplificazione modifica invece lo stato prima della misura finale, usando anche l'inverso della preparazione. Il confronto riguarda pertanto due procedure con accessi differenti: la ripetizione richiede soltanto di preparare e misurare, mentre l'amplificazione richiede controllo coerente sul circuito e sulla sua inversione.

$$E[\text{tentativi}] = \frac{1}{p}, \qquad P(L \text{ fallimenti}) = (1-p)^L$$

Richiami Con tentativi indipendenti, la probabilità di una successione di fallimenti è il prodotto delle singole probabilità. E indica un valore atteso, non un numero obbligatorio di tentativi.

Le due componenti della preparazione

Separiamo la somma che definisce $|\psi\rangle$ in due parti: quella con $f(x) = 1$ e quella con $f(x) = 0$. Le due parti hanno supporti disgiunti, quindi sono ortogonali. Le loro norme sono rispettivamente $\sqrt{p}$ e $\sqrt{1-p}$. Dividendo ogni parte per la propria norma otteniamo due stati normalizzati, che chiamiamo ancora $|g\rangle$ e $|b\rangle$.

Questi stati dipendono ora dalle ampiezze preparate da A. Possono avere coefficienti diversi fra le varie soluzioni. Sotto l'ipotesi $0 < p < 1$, la decomposizione riassume comunque tutta la preparazione in un piano: una componente lungo $|g\rangle$ e una lungo $|b\rangle$. Le definizioni non richiedono che il circuito conosca le singole soluzioni.

$$|g\rangle = \frac{1}{\sqrt{p}} \sum_{f(x)=1} \alpha_x |x\rangle$$

$$|b\rangle = \frac{1}{\sqrt{1-p}} \sum_{f(x)=0} \alpha_x |x\rangle$$

$$|\psi\rangle = \sqrt{p}\,|g\rangle + \sqrt{1-p}\,|b\rangle$$

Richiami Due vettori con supporti disgiunti nella base computazionale sono ortogonali. Le definizioni di entrambe le componenti normalizzate richiedono $0 < p < 1$.

L’inverso del circuito di preparazione

Per un operatore unitario, l’inverso coincide con l’aggiunto, indicato con $A^\dagger$. In una rappresentazione matriciale, l’aggiunto si ottiene trasponendo la matrice e coniugando i coefficienti. La sua proprietà operativa è $A^\dagger A = I$: applicare A e poi $A^\dagger$ riporta qualunque stato al punto di partenza.

Se A è descritto da una sequenza di porte, si costruisce l’inverso applicando le porte inverse nell’ordine opposto. Ad esempio, per una preparazione che applica prima H e poi una rotazione R, l’operatore è $A = RH$ e l’inverso è $A^\dagger = HR^\dagger$. Il primo passo temporale dell’inverso è dunque $R^\dagger$. Misurare e ripreparare il registro non realizza questa inversione coerente.

$$A^\dagger A = AA^\dagger = I$$

$$(RH)^\dagger = HR^\dagger$$

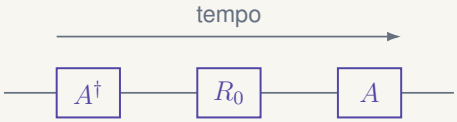
Richiami Nel prodotto UV, l’operatore V agisce per primo. La porta Hadamard soddisfa $H^\dagger = H$ e $H^2 = I$.

La riflessione rispetto allo stato preparato

Per riflettere rispetto a $|\psi\rangle$, utilizziamo l'operatore $D_A = 2|\psi\rangle\langle\psi| - I$. Possiamo realizzarlo riportando prima il riferimento $|\psi\rangle$ allo stato zero con $A^\dagger$, applicando la riflessione $R_0 = 2|0\rangle\langle 0| - I$ e infine riapplicando A. Il prodotto complessivo è quindi $AR_0A^\dagger$.

Il procedimento conserva la componente parallela a $|\psi\rangle$ e cambia segno alla componente ortogonale, esattamente come faceva il diffusore rispetto allo stato uniforme. La nuova riflessione dipende dalla preparazione scelta. Riutilizzare senza modifiche il diffusore uniforme farebbe invece ruotare lo stato rispetto a un riferimento diverso e non darebbe la costruzione che stiamo analizzando.

Riflessione attorno allo stato preparato



$$R_0 = 2|0\rangle\langle 0| - I$$

$$D_A = A\, R_0\, A^\dagger$$

Il tempo procede da sinistra a destra. La sequenza applica prima l'inverso della preparazione, riflette rispetto allo zero e ripristina la preparazione.

Richiami $2|v\rangle\langle v| - I$ riflette rispetto al vettore normalizzato $|v\rangle$. $A|0\rangle = |\psi\rangle$ e $A^\dagger|\psi\rangle = |0\rangle$.

La verifica algebrica della riflessione

Sostituiamo $R_0 = 2|0\rangle\langle 0| - I$ nel prodotto $AR_0A^\dagger$. La linearità permette di distribuire gli operatori sui due termini. Nel primo compare $A|0\rangle = |\psi\rangle$ e, prendendo l'aggiunto, $\langle 0|A^\dagger = \langle\psi|$. Nel secondo compare $AA^\dagger = I$. Otteniamo perciò esattamente $2|\psi\rangle\langle\psi| - I$.

Questa identità spiega la costruzione del circuito, senza richiedere una lista delle ampiezze o delle soluzioni al suo interno. Il circuito usa A come operazione disponibile. La notazione con il proiettore descrive matematicamente la sua azione. Come per il diffusore uniforme, il quadrato della riflessione è l'identità, dunque applicarla due volte annulla il suo effetto.

$$AR_0A^\dagger = 2A|0\rangle\langle 0|A^\dagger - AA^\dagger$$

$$= 2|\psi\rangle\langle\psi| - I = D_A$$

Richiami Per un vettore normalizzato, $P = |\psi\rangle\langle\psi|$ soddisfa $P^2 = P$. Da $P^2 = P$ segue $(2P - I)^2 = I$.

L’iterazione rispetto alla preparazione scelta

L’oracolo O_f continua a cambiare segno alle componenti valide. Lo facciamo seguire dalla riflessione D_A e definiamo $Q_A = D_A O_f$. Nel piano generato da $|b\rangle$ e $|g\rangle$, questa composizione ripete la geometria già studiata. L’angolo iniziale soddisfa ora $\sin^2\theta = p$, perché la probabilità delle soluzioni dipende dalla preparazione.

Dopo k iterazioni, la componente valida vale $\sin((2k+1)\theta)$, mentre quella non valida vale $\cos((2k+1)\theta)$. La probabilità totale di successo è ancora il quadrato del seno. Per la preparazione uniforme, $p = M/N$ e si recupera Grover. La generalizzazione conserva quindi la struttura del ragionamento, sostituendo lo stato rispetto al quale viene eseguita la riflessione.

$$Q_A = D_A O_f, \qquad \sin^2\theta = p$$

$$Q_A^k|\psi\rangle = \cos((2k+1)\theta)|b\rangle + \sin((2k+1)\theta)|g\rangle$$

Richiami $|b\rangle$ e $|g\rangle$ sono le componenti normalizzate della preparazione corrente. Il numero k conta le coppie oracolo–riflessione.

Un successo iniziale di un quarto

Un esempio con un solo qubit permette di vedere l'amplificazione senza altri registri. Il verificatore accetta 1 e la preparazione produce $|\psi\rangle = (\sqrt{3}/2)|0\rangle + (1/2)|1\rangle$. La probabilità iniziale di successo è $1/4$, quindi $\theta = \pi/6$. Una iterazione porta l'angolo a $3\theta = \pi/2$ e il risultato diventa $|1\rangle$.

La preparazione può essere descritta dalla matrice reale riportata sotto: la prima colonna è proprio lo stato ottenuto da $|0\rangle$. Le colonne hanno norma 1 e prodotto scalare nullo, quindi la matrice è unitaria. Questo esempio distingue il numero dei candidati dalla probabilità iniziale: i candidati sono due, ma la preparazione assegna alla soluzione probabilità $1/4$, non $1/2$.

$$A = \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}$$

$$p = \frac{1}{4}, \qquad P_1 = \sin^2\left(3\frac{\pi}{6}\right) = 1$$

Richiami Le colonne di una matrice unitaria formano una base ortonormale. L'angolo iniziale dipende da p, non dal solo numero di qubit.

La traccia dell’amplificazione

Torniamo alla preparazione su quattro stringhe con peso $1/16$ sulla soluzione 11. L’angolo è $\arcsin(1/4)$ e il primo picco continuo corrisponde a circa $2{,}608$ iterazioni. I due interi adiacenti sono 2 e 3. Il confronto delle probabilità favorisce tre iterazioni, con successo $63001/65536$, circa $0{,}9613$.

Il numero dei candidati è quattro, ma la curva coincide con quella di Grover uniforme su sedici candidati con una soluzione. I due problemi hanno infatti lo stesso peso iniziale $p = 1/16$. I loro circuiti e i loro costi di preparazione possono essere diversi. La coincidenza riguarda soltanto la rotazione nel piano e la probabilità totale di successo.

k	P_k
0	$1/16$
1	$121/256$
2	$3721/4096$
3	$63001/65536$
4	$609961/1048576$

Richiami $P_k = \sin^2((2k+1)\theta)$ con $\sin\theta = 1/4$. Il primo picco continuo è $\pi/(4\theta) - 1/2$ e va confrontato con durate intere.

Due soluzioni con pesi iniziali differenti

Consideriamo ora valide le stringhe 10 e 11. La preparazione assegna probabilità $1/20$ alla prima e $3/20$ alla seconda. Le probabilità delle stringhe non valide 00 e 01 sono entrambe $2/5$. La somma vale 1 e il successo totale iniziale è $p = 1/5$.

All'interno della parte valida, 11 pesa tre volte 10. Lo stato valido normalizzato è quindi $|g\rangle = (1/2)|10\rangle + (\sqrt{3}/2)|11\rangle$. La componente non valida è $|b\rangle = (|00\rangle + |01\rangle)/\sqrt{2}$. L'amplificazione modifica i pesi complessivi delle due componenti, ma la rotazione resta nel piano che esse generano. Non introduce una nuova ripartizione fra le due soluzioni.

$$|\psi\rangle = \sqrt{\frac{2}{5}}(|00\rangle + |01\rangle) + \frac{1}{\sqrt{20}}|10\rangle + \sqrt{\frac{3}{20}}|11\rangle$$

$$|g\rangle = \frac{1}{2}|10\rangle + \frac{\sqrt{3}}{2}|11\rangle$$

Richiami La normalizzazione della parte valida divide le sue ampiezze per $\sqrt{p}$. Le probabilità condizionate delle due soluzioni sono $1/4$ e $3/4$.

La distribuzione fra le soluzioni rimane la stessa

Dopo l'amplificazione, la parte valida dello stato è un unico coefficiente moltiplicato per $|g\rangle$. Tutte le ampiezze delle soluzioni ricevono quindi lo stesso fattore. Se la misura trova una soluzione, la probabilità condizionata di ciascun valore rimane quella definita dalla preparazione iniziale.

Per una soluzione x, la probabilità iniziale condizionata al successo è $|\alpha_x|^2/p$. La stessa frazione si ottiene dopo l'amplificazione, purché la probabilità totale di successo in quel momento sia diversa da zero. Nel nostro esempio, ogni misura riuscita restituisce 10 con probabilità $1/4$ e 11 con probabilità $3/4$. L'amplificazione aumenta la probabilità di entrare nell'insieme valido, senza renderne automaticamente uniforme il campionamento.

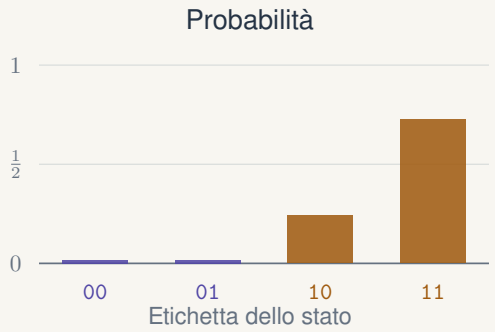
$$P(\text{esito } x \mid \text{successo}) = \frac{|\alpha_x|^2}{p}$$

Richiami $P(\text{esito } x \mid \text{successo})$ indica una probabilità condizionata all'evento di successo. Il coefficiente comune alle ampiezze valide si semplifica nel rapporto fra probabilità.

La distribuzione dopo una iterazione

Con $p = 1/5$, l'identità $\sin(3\theta) = 3\sin\theta - 4\sin^3\theta$ dà $P_1 = p(3-4p)^2 = 121/125$. Le due soluzioni si dividono questo peso nei rapporti $1/4$ e $3/4$. Le loro probabilità sono quindi $121/500$ e $363/500$.

La probabilità residua $4/125$ si divide in parti uguali fra i due candidati non validi, ciascuno con peso $2/125$. Il grafico permette di distinguere l'aumento del successo complessivo dalla distribuzione dei singoli esiti. Il valore 11 rimane tre volte più probabile di 10, perché questa preferenza era già presente nella componente valida della preparazione.



Il grafico mostra probabilità teoriche. Le due barre valide crescono insieme, conservando il rapporto fra i pesi delle due soluzioni.

Richiami $p = 1/5$ è il successo iniziale e $P_1 = 121/125$ quello dopo una iterazione. Le probabilità di tutti gli esiti devono sommare a 1.

Una soluzione priva di peso iniziale

Se A assegna ampiezza zero a tutte le soluzioni, allora $p = 0$. L'oracolo non modifica lo stato preparato, perché non vi compaiono componenti valide. Anche la riflessione D_A lascia invariato $|\psi\rangle$. Ripetere l'iterazione non crea dunque un peso iniziale che manca completamente.

Può inoltre accadere che alcune soluzioni abbiano ampiezza zero e altre no. L'amplificazione aumenta soltanto la componente valida presente nella preparazione: una soluzione assente da quella componente resta assente. All'estremo opposto, se $p = 1$, una misura immediata ha già successo. Questi casi spiegano le ipotesi $0 < p < 1$ usate per definire entrambe le componenti normalizzate e per disegnare il piano della rotazione.

$$p = 0 : \quad Q_A^k |\psi\rangle = |\psi\rangle$$

$$p = 1 : \quad P(\text{successo immediato}) = 1$$

Richiami L'amplificazione resta nel piano generato dalle parti valida e non valida dello stato iniziale. Una componente di norma zero non può essere normalizzata.

Il ruolo dei registri di lavoro nella preparazione

Per usare $A^\dagger$, la preparazione deve essere disponibile come evoluzione unitaria sull'intero sistema considerato. Se un algoritmo usa registri di lavoro, bisogna includerli nello stato e nell'operazione inversa. Il verificatore può dipendere soltanto dai bit del candidato, ma le altre correlazioni appartengono comunque alla preparazione su cui viene costruita la riflessione.

Le misure intermedie e la cancellazione irreversibile di informazioni non si possono invertire semplicemente ripercorrendo le istruzioni. In alcuni casi si può costruire una versione coerente di un procedimento classico conservando i dati di lavoro e ripulendoli reversibilmente. Questa trasformazione ha però un costo. Il vantaggio nell'amplificazione delle probabilità non rende gratuita la costruzione di A, del suo inverso o del verificatore.

Richiami $A^\dagger$ deve invertire A su tutti i registri su cui A agisce. La cancellazione reversibile degli ausiliari conserva le correlazioni necessarie all'interferenza.

Il numero di usi della preparazione e del suo inverso

Una esecuzione con k iterazioni usa una prima applicazione di A per preparare lo stato. Ogni riflessione $D_A = AR_0A^\dagger$ richiede poi una applicazione di $A^\dagger$ e una di A. Il totale è quindi $2k+1$ usi dei due circuiti di preparazione, oltre a k marcature e alle riflessioni rispetto allo zero.

Per l'esempio con $p = 1/16$, tre iterazioni richiedono sette usi di A o $A^\dagger$ e raggiungono probabilità circa $0{,}9613$. La semplice ripetizione aveva bisogno in media di sedici preparazioni fino al successo. I numeri non sono ancora un confronto di tempi reali, perché l'amplificazione usa anche il verificatore coerente e altre porte. Per p piccolo, il numero di iterazioni cresce come $1/\sqrt{p}$.

$$\#(A \text{ o } A^\dagger) = 2k+1, \qquad \#O_f = k$$

Richiami Il conteggio delle chiamate deve includere la preparazione iniziale e tutte le operazioni inverse. Per p piccolo, $\arcsin\sqrt{p} \approx \sqrt{p}$.

La preparazione generale con successo ignoto

Anche per una preparazione generale, scegliere la durata del primo picco richiede di conoscere il successo iniziale p. Se tale dato manca, le durate casuali e i riavvii verificati possono essere adattati all'amplificazione. La geometria che determina la media dei successi dipende da θ, con $\sin^2\theta = p$, e conserva la struttura già usata per la ricerca uniforme.

Il risultato generale garantisce, sotto le ipotesi di accesso coerente a A, al suo inverso e al verificatore, un costo atteso proporzionale a $1/\sqrt{p}$ per $p > 0$. Questa garanzia non equivale a conoscere preventivamente la durata dell'esecuzione. Se si vuole imporre un budget deterministico o dichiarare un errore massimo, occorre aggiungere un'analisi dei riavvii e delle ipotesi disponibili su p.

Richiami p è la probabilità che la preparazione $A|0\rangle$ produca una soluzione alla misura. Il costo atteso tiene conto della distribuzione delle durate casuali.

Esercizio sulla probabilità delle singole soluzioni

Una preparazione assegna probabilità $1/20$ a una soluzione u e $3/20$ a una soluzione v. Non ci sono altre soluzioni. Dopo una iterazione, il successo totale è $P_1 = p(3-4p)^2 = 121/125$, come calcolato nell'esempio precedente. Quali sono le probabilità finali di misurare u e v? Qual è la probabilità che una misura riuscita restituisca v?

Per rispondere non occorre conoscere la distribuzione fra i candidati non validi. Occorre invece distinguere tre quantità: il successo iniziale, il successo totale dopo l'amplificazione e la probabilità condizionata a una misura riuscita. Si verifichi infine che la somma delle probabilità finali di u e v coincida con il successo totale assegnato nel testo.

Richiami La componente valida normalizzata conserva i rapporti fra le probabilità delle soluzioni. La probabilità iniziale totale è $p = 1/20 + 3/20 = 1/5$.

Soluzione: le probabilità dei due esiti

Il peso condizionato di u è $(1/20)/(1/5) = 1/4$, mentre quello di v è $3/4$. Dopo l'amplificazione moltiplichiamo questi pesi per il nuovo successo totale $121/125$. Otteniamo $P(u) = 121/500$ e $P(v) = 363/500$. La loro somma è $484/500 = 121/125$, come richiesto.

Sapendo che la misura ha prodotto una soluzione, la probabilità che si tratti di v rimane $3/4$. La probabilità assoluta $363/500$ comprende invece anche il rischio che la misura cada nell'insieme non valido. La distinzione è importante quando si utilizza Grover come parte di un procedimento più ampio: trovare una soluzione con alta probabilità non implica campionare in modo uniforme tutte le soluzioni.

$$P(u) = \frac{121}{125} \cdot \frac{1}{4} = \frac{121}{500}, \qquad P(v) = \frac{121}{125} \cdot \frac{3}{4} = \frac{363}{500}$$

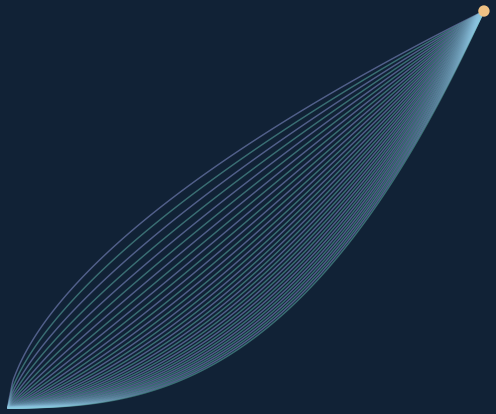
$$P(v \mid \text{successo}) = \frac{363/500}{121/125} = \frac{3}{4}$$

Richiami Una probabilità condizionata divide la probabilità dell'esito per quella dell'evento su cui si condiziona. La formula richiede che la probabilità dell'evento condizionante sia positiva.

CAPITOLO 12

Costo e uso dell’algoritmo

Risorse, limiti
ed esercizi svolti



L’algoritmo di Grover

I dati necessari alla ricerca

La procedura completa parte da un insieme di candidati codificati in un registro e da un predicato f che ne verifica la validità. Per la versione uniforme assumiamo $N = 2^n$. Occorre disporre dell'oracolo coerente O_f e di una realizzazione della riflessione rispetto a $|s\rangle$. Il numero delle soluzioni M può essere noto oppure ignoto, ma i due casi richiedono controlli classici differenti.

Con M noto possiamo scegliere una durata vicina al primo picco. Con M ignoto utilizziamo durate adattive e verifiche, come descritto nel capitolo 10. In entrambi i casi l'algoritmo restituisce una stringa candidata, che viene accettata solo dopo aver verificato $f(x) = 1$. Le etichette marcate mostrate negli esempi erano dati della dimostrazione, non informazioni che il programma di ricerca dovesse conoscere.

Richiami $O_f|x\rangle = (-1)^{f(x)}|x\rangle$ marca le soluzioni mediante una fase. $|s\rangle = N^{-1/2}\sum_x |x\rangle$ è lo stato uniforme.

La scelta della durata quando M è noto

Se $M = 0$ è noto, non esistono soluzioni da cercare. Se $M = N$ è noto, basta scegliere un candidato. Nel caso $0 < M < N$, calcoliamo $\theta = \arcsin\sqrt{M/N}$ e il valore reale $k_* = \pi/(4\theta) - 1/2$ che porterebbe esattamente al primo picco. Il numero di iterazioni deve però essere un intero non negativo.

Confrontiamo quindi le probabilità dei due interi adiacenti a k_*, usando la formula esatta P_k. Questo confronto sceglie una buona durata vicino al primo picco, senza sostenere che sia la migliore fra tutti i tempi possibili o per ogni criterio di costo. Quando le soluzioni sono numerose, anche una misura immediata può già essere adeguata. La scelta deve riflettere la probabilità di successo desiderata e il costo delle ripetizioni.

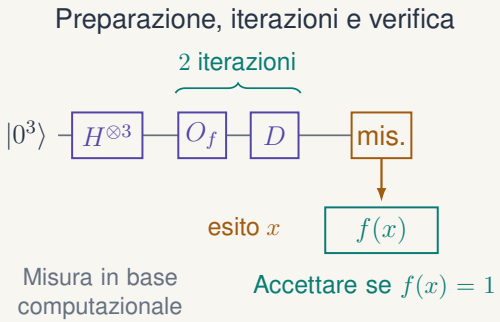
$$k_* = \frac{\pi}{4\theta} - \frac{1}{2}, \qquad P_k = \sin^2((2k+1)\theta)$$

Richiami k è un intero non negativo: si confrontano i due interi adiacenti al valore continuo k_*. k conta iterazioni complete DO_f.

La preparazione, le iterazioni e la verifica

Prepariamo n qubit nello stato zero e applichiamo Hadamard a ciascuno. Il registro contiene così lo stato uniforme. Eseguiamo il numero scelto di iterazioni, applicando ogni volta prima l'oracolo di fase e poi il diffusore. La misura finale restituisce una stringa di n bit, che viene passata al verificatore classico.

Se la verifica riesce, restituiamo la stringa. Se fallisce e il budget consente altri tentativi, inizializziamo nuovamente il registro. Non proseguiamo semplicemente dall'uscita misurata: la misura ha modificato lo stato su cui si basa la rotazione. Nel diagramma sono rappresentate due iterazioni, adatte all'esempio con otto candidati e una soluzione.



La parte quantistica termina con la misura. La verifica è un calcolo classico sul solo candidato ottenuto e determina se occorra un nuovo tentativo.

Richiami Il prodotto $G = DO_f$ applica O_f prima di D. La misura dà un candidato, non la lista delle ampiezze del registro.

Un’esecuzione completa per otto candidati

Per l’esempio con unica soluzione 101, abbiamo $N = 8$, $M = 1$ e $\theta \approx 0{,}36137$ radianti. Il valore continuo del primo picco è circa $1{,}6734$. Confrontiamo una e due iterazioni: il successo passa da $25/32$ a $121/128$, quindi scegliamo due.

Dopo la misura, il verificatore accetta 101 e rifiuta ogni altra stringa. La tabella riassume la distribuzione teorica del candidato. Il segno negativo delle altre sette ampiezze, pari a $-1/(8\sqrt{2})$, non produce probabilità negative: ogni probabilità è $1/128$. Questa distribuzione descrive molte possibili esecuzioni del circuito ideale, non un risultato deterministico di ogni singola misura.

Esito	Probabilità
101	$121/128$
Ciascun altro esito	$1/128$
Tutti gli altri esiti	$7/128$

Richiami La probabilità è il modulo quadro dell’ampiezza. La somma $121/128 + 7/128 = 1$ verifica la normalizzazione.

Due tentativi e la probabilità di insuccesso

Il fallimento di un tentativo nell'esempio precedente ha probabilità $7/128$. Se ripreparimo lo stato ed eseguiamo un secondo tentativo indipendente, la probabilità di fallire entrambi è $(7/128)^2 = 49/16384$. La probabilità di ottenere almeno una soluzione nei due tentativi è quindi $16335/16384$, circa $0{,}9970$.

La verifica impedisce di accettare un candidato sbagliato. Può restare soltanto il rischio di non trovare alcun candidato valido entro i due tentativi. Questa distinzione è utile per specificare l'uscita dell'algoritmo: o una soluzione verificata, oppure un esito di mancato ritrovamento. Le probabilità dichiarate presuppongono oracoli, porte e misure ideali.

$$P(\text{due fallimenti}) = \frac{49}{16384}$$

$$P(\text{almeno un successo}) = 1 - \frac{49}{16384}$$

Richiami Il complemento dell'evento «tutti i tentativi falliscono» è «almeno uno riesce». I tentativi devono essere indipendenti per moltiplicare le probabilità di fallimento.

Il costo medio fino a una soluzione verificata

Contiamo una interrogazione per ciascun oracolo di fase e una per la verifica finale del candidato. Un tentativo con due iterazioni costa così tre interrogazioni. Ripetendolo fino al successo, il numero medio di tentativi è $128/121$ e il costo medio è $384/121$, circa $3{,}17$ interrogazioni.

Questo conto non comprende ancora le porte del diffusore, la preparazione, la misura o l'accesso ai dati. Inoltre, la durata che massimizza il successo di un tentativo non minimizza necessariamente il costo medio con riavvii: un tentativo più breve può costare meno anche se riesce meno spesso. Per questo motivo occorre dichiarare quale quantità viene ottimizzata quando si parla di numero ottimale di iterazioni.

$$E[\text{interrogazioni fino al successo}] = \frac{k+1}{P_k}$$

$$\frac{2+1}{121/128} = \frac{384}{121}$$

Richiami Il termine $+1$ conta la verifica classica finale. Il valore atteso dei tentativi fino al primo successo è $1/P_k$.

L’unità di costo nel modello a oracolo

Nel modello a oracolo, una interrogazione significa un accesso all’operazione che contiene il predicato. Se scegliamo come primitiva O_f, una iterazione di Grover ne usa una. Questa convenzione permette di studiare quante volte sia necessario interagire con il problema, separando tale domanda dalla realizzazione interna del verificatore.

Un oracolo di fase può essere costruito da un oracolo XOR con l’ausiliario nello stato meno. Può anche essere realizzato calcolando il predicato su un bit di lavoro, applicando una fase e cancellando il calcolo a ritroso. Quest’ultima costruzione usa un calcolo e una cancellazione reversibile. Contare una interrogazione astratta non significa quindi contare una sola porta elementare o una sola istruzione di un programma classico.

Richiami L’oracolo XOR agisce come $U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$. L’oracolo di fase applica il segno $(-1)^{f(x)}$ alla componente di ingresso.

La dipendenza dalla radice quadrata

Quando M/N è piccolo, l'angolo iniziale è piccolo e $\theta = \arcsin\sqrt{M/N} \approx \sqrt{M/N}$. Per arrivare vicino al primo picco occorrono circa $\pi/(4\theta)$ iterazioni. Sostituendo l'approssimazione si ottiene un numero proporzionale a $\sqrt{N/M}$.

La notazione $O(\sqrt{N/M})$ esprime un limite superiore asintotico, trascurando fattori costanti e termini inferiori nel regime considerato. Il risultato si riferisce alle interrogazioni e a una probabilità di successo fissata, non a un tempo in secondi. Con un campionamento classico uniforme indipendente, la probabilità di successo è M/N e il numero medio di verifiche è invece N/M.

$$k \approx \frac{\pi}{4}\sqrt{\frac{N}{M}}$$

$$E[\text{verifiche classiche con reinserimento}] = \frac{N}{M}$$

Richiami Per u piccolo, $\arcsin u \approx u$ quando gli angoli sono espressi in radianti. Con reinserimento, lo stesso candidato può essere scelto più volte.

Una soluzione fra circa un milione di candidati

Un registro di venti qubit codifica $N = 2^{20} = 1\,048\,576$ candidati. Se esiste una sola soluzione, la radice di N è 1024 e la stima $\pi\sqrt{N}/4$ vale circa $804{,}25$. Il confronto vicino al primo picco sceglie 804 iterazioni. Nel circuito ideale la probabilità risultante è circa $0{,}999999757$.

Per un confronto classico esplicito, interrogare $524\,288$ candidati distinti dà probabilità $1/2$ di aver trovato l'unica soluzione, se la sua posizione è uniforme. Le soglie di successo qui sono diverse: il dato classico riguarda un mezzo, quello quantistico una probabilità vicina a 1. Entrambi rendono visibile la diversa scala delle interrogazioni, ma non costituiscono ancora un confronto fra tempi di esecuzione su due macchine reali.

$$N = 2^{20}, \qquad \sqrt{N} = 2^{10} = 1024$$

$$P_{804} = \sin^2\!\left(1609 \arcsin \frac{1}{1024}\right)$$

Richiami Con una soluzione uniforme, q test classici distinti hanno successo q/N. Una verifica finale del risultato si aggiunge alle 804 interrogazioni di fase.

La crescita rispetto al numero di bit

Il numero dei candidati è esponenziale nel numero di bit: $N = 2^n$. Per una soluzione, il costo di Grover cresce come $\sqrt{N} = 2^{n/2}$. È molto minore di 2^n, ma rimane esponenziale in n. Se raddoppiamo il numero di bit, la radice dello spazio di ricerca viene elevata al quadrato.

Per esempio, passando da 20 a 40 bit, la quantità $\sqrt{N}$ passa da $2^{10} = 1024$ a $2^{20} = 1\,048\,576$. Il vantaggio quadratico va quindi interpretato rispetto al numero dei candidati, non come trasformazione generale di un problema esponenziale in uno polinomiale. Il confronto con Shor richiede proprio questa attenzione alla misura della lunghezza dell'ingresso.

$$\sqrt{2^n} = 2^{n/2}$$

Richiami n è il numero di bit necessari per descrivere un candidato. Un costo polinomiale in n cresce come una potenza fissata di n, non come 2^{cn}.

Dieci candidati in un registro da quattro qubit

Un problema può avere dieci candidati, mentre un registro da quattro qubit dispone di sedici etichette. Possiamo associare i candidati alle etichette da 0 a 9 e definire il predicato in modo che rifiuti sempre le etichette da 10 a 15. Le sei etichette aggiuntive fanno parte dello spazio su cui agiscono le porte, anche se non rappresentano oggetti del problema.

Se prepariamo uniformemente tutti i quattro qubit, la probabilità iniziale di successo è $M/16$. Nelle formule di Grover dobbiamo quindi usare $N = 16$, non 10. Per usare una distribuzione uniforme soltanto sui dieci candidati servirebbe una preparazione diversa e la riflessione corrispondente. Il suo costo e la sua inversione rientrerebbero nell'analisi dell'amplificazione generale.

$$f(x) = 0 \quad \text{per } 10 \le x \le 15, \qquad p = \frac{M}{16}$$

Richiami Con quattro qubit, la base computazionale contiene $2^4 = 16$ stati. La probabilità iniziale dipende dalla preparazione effettivamente realizzata.

L’ottimalità riguarda la ricerca senza struttura

Per la ricerca senza struttura con accesso a oracolo, anche gli algoritmi quantistici richiedono un numero di interrogazioni dell’ordine di $\sqrt{N}$ quando la soluzione è unica e si vuole una probabilità costante di successo. Il simbolo $\Omega(\sqrt{N})$ indica un limite inferiore di questo ordine. Insieme al limite superiore di Grover, dà la dipendenza asintotica $\Theta(\sqrt{N})$.

Il risultato riguarda algoritmi che apprendono il predicato attraverso gli accessi previsti dal modello. Non dimostra che ogni problema con N candidati richieda quel costo. Un problema può offrire struttura algebrica, ordinamento o altre informazioni sfruttabili da un algoritmo specializzato. La dimostrazione completa del limite inferiore è un risultato ulteriore: i piccoli esempi e i diagrammi di rotazione ne illustrano il contesto, ma non la sostituiscono.

ricerca a oracolo con una soluzione: $\Theta(\sqrt{N})$

Richiami O indica un limite superiore asintotico, Ω un limite inferiore e Θ la coincidenza dell’ordine. Il modello stabilisce quali informazioni sull’ingresso sono accessibili.

Il costo delle porte del verificatore

Indichiamo con C_f il numero di porte necessario a calcolare reversibilmente il predicato e con C_D il costo del diffusore. Se l'oracolo di fase usa calcolo, una porta di fase e cancellazione del calcolo, il suo costo è circa $2C_f + C_Z$, dove C_Z rappresenta la marcatura. Le porte di cancellazione sono necessarie quanto quelle del calcolo iniziale.

Un tentativo con k iterazioni costa quindi la preparazione iniziale più k volte la somma dei costi dell'oracolo e del diffusore, oltre alla misura e alla verifica finale. Il vantaggio nelle interrogazioni può essere importante, ma deve essere confrontato con questi costi. In particolare, un verificatore classico economico non garantisce automaticamente una realizzazione reversibile altrettanto economica.

$$C_{O_f} \approx 2C_f + C_Z$$

$$C_{\text{circuito}} \approx C_{\text{prep}} + k(C_{O_f} + C_D)$$

Richiami Il conteggio delle porte qui riguarda un circuito logico ideale. La formula del circuito non include ancora misura, verifica classica e correzione degli errori.

L’accesso ai dati del problema

La preparazione uniforme di un registro di indirizzi richiede Hadamard sui suoi qubit, ma non carica automaticamente una tabella di dati nel circuito. Se il predicato dipende da valori memorizzati altrove, l’oracolo deve poter accedere a quei valori in modo coerente con gli indirizzi in sovrapposizione. Tale accesso è una parte della realizzazione del problema.

Quando invece il predicato calcola una proprietà della stringa candidata, si costruisce un circuito reversibile per quel calcolo. Le due situazioni possono avere costi molto diversi, pur condividendo la stessa formula per il numero delle interrogazioni. Per valutare un’applicazione occorre specificare dove risiedono i dati, come vengono consultati e quale preparazione sia necessaria. Il termine ricerca in un database, da solo, non fornisce questa descrizione.

Richiami $H^{\otimes n}|0^n\rangle$ prepara indirizzi uniformi, senza fornire i valori di un archivio esterno. L’oracolo incorpora il verificatore e l’accesso alle informazioni di cui esso ha bisogno.

Le probabilità teoriche e l'esecuzione di un circuito

I valori esatti ottenuti per quattro e otto candidati presuppongono porte unitarie corrette, registri ausiliari ripuliti e una misura ideale. Se una porta introduce errori, lo stato reale può allontanarsi dal piano e dalla rotazione previsti. Aumentare il numero delle iterazioni prolunga inoltre il circuito, anche quando il calcolo ideale prevederebbe una probabilità elevata.

Una valutazione sperimentale deve quindi confrontare la distribuzione osservata con quella teorica e tenere conto del costo dell'affidabilità. Le frequenze di un numero finito di esecuzioni oscillano anche in assenza di errori, perché le misure sono casuali. I grafici di questo manuale riportano distribuzioni teoriche, non dati raccolti su un dispositivo. Non consentono, da soli, di stimare un vantaggio prestazionale su hardware.

Richiami La frequenza di un esito è il rapporto fra il numero delle sue osservazioni e il numero delle esecuzioni. La probabilità teorica appartiene al modello del circuito, non a una singola esecuzione.

Simon, Shor e Grover nel percorso del corso

Simon sfrutta una promessa di periodicità rispetto allo XOR e ricava vincoli sulla stringa nascosta. Shor riconduce la fattorizzazione alla ricerca di un periodo aritmetico, usando campionamento di Fourier e ricostruzione classica. In entrambi i casi, la struttura della funzione determina il tipo di informazione estratta.

Grover richiede invece un verificatore di soluzioni e aumenta il peso delle componenti che lo soddisfano. La misura cerca direttamente un candidato valido, senza ricostruire un periodo. Tutti e tre gli algoritmi controllano l'interferenza, ma questo elemento comune non rende intercambiabili le loro ipotesi. Anche i confronti di costo vanno riferiti ai rispettivi problemi e modelli di accesso.

Algoritmo	Informazione ricercata
Simon	Vincoli sulla stringa di periodo XOR
Shor	Periodo delle potenze modulari
Grover	Un candidato accettato dal predicato

Richiami Una promessa è un'ipotesi sulla funzione di ingresso che l'algoritmo può sfruttare. In Grover la marcatura dipende da $f(x)$, senza una promessa di periodicità.

Esercizio: una soluzione fra sedici candidati

Un verificatore accetta una sola stringa fra le sedici rappresentabili con quattro qubit. Si parte dalla preparazione uniforme. Si calcolino la probabilità iniziale, l'angolo θ e i due interi adiacenti alla posizione continua del primo picco. Si scelga poi il numero di iterazioni che dà il successo maggiore fra questi due valori.

Per controllare il calcolo si può usare la formula trigonometrica oppure seguire le ampiezze comuni ai candidati validi e non validi. Nella seconda via, dopo l'oracolo bisogna ricalcolare la media prima di applicare $2\bar{a} - a_x$. Si indichino infine il numero delle interrogazioni di fase e il costo aggiuntivo della verifica classica, senza confonderli con il numero delle porte del circuito.

Richiami Con una soluzione e $N = 16$, l'ampiezza iniziale di ogni candidato è $1/4$. $k_* = \pi/(4\theta) - 1/2$ e $\sin\theta = 1/4$.

Soluzione: tre iterazioni e verifica finale

La probabilità iniziale è $1/16$ e l'angolo è $\theta = \arcsin(1/4) \approx 0{,}25268$ radianti. La posizione continua del primo picco è circa $2{,}6083$. I due interi da confrontare sono quindi 2 e 3. Le probabilità sono rispettivamente $3721/4096 \approx 0{,}9084$ e $63001/65536 \approx 0{,}9613$, perciò scegliamo tre iterazioni.

Al termine della terza iterazione, l'ampiezza della soluzione è $251/256$ e quella di ciascuno degli altri quindici candidati è $-13/256$. La somma dei moduli quadrati vale 1. Il tentativo usa tre interrogazioni di fase e una verifica classica finale. L'esempio coincide nella probabilità con la preparazione generale a successo $1/16$, pur usando qui uno stato uniforme su un registro più ampio.

$$\left(\frac{251}{256}\right)^2 + 15\left(\frac{-13}{256}\right)^2 = 1$$

$$P_3 = \frac{63001}{65536}$$

Richiami Il segno delle ampiezze non valide influisce sulle successive interferenze, mentre il loro quadrato determina la probabilità. La verifica finale costa una chiamata distinta.

Esercizio sul costo della realizzazione reversibile

Un verificatore reversibile richiede 120 porte per calcolare il predicato su un ausiliario pulito. La marcatura richiede una porta e il calcolo a ritroso ha lo stesso costo del calcolo iniziale. Il diffusore richiede 40 porte e la preparazione iniziale 10. Si stimi il costo logico di un tentativo con tre iterazioni, escludendo misura e verifica classica finale.

Si calcoli inoltre la frazione del costo delle iterazioni dovuta all'oracolo di fase. Il confronto serve a distinguere il numero delle interrogazioni dal costo della loro implementazione. Tutti i valori forniti sono ipotesi numeriche dell'esercizio: non descrivono una particolare piattaforma o una decomposizione ottimizzata delle porte.

Richiami Un oracolo compute–fase–uncompute costa $C_{O_f} = 2C_f + C_Z$. Il costo stimato è $C_{\text{prep}} + k(C_{O_f} + C_D)$.

Soluzione: l'oracolo domina il conteggio delle porte

Il calcolo e la sua cancellazione richiedono complessivamente 240 porte. Aggiungendo la marcatura, una interrogazione di fase costa 241 porte. Ogni iterazione completa ne richiede $241 + 40 = 281$. Le tre iterazioni costano dunque 843 porte e, aggiungendo la preparazione, il totale diventa 853.

Delle 843 porte impiegate nelle iterazioni, 723 appartengono alle tre realizzazioni dell'oracolo. La loro frazione è $723/843 = 241/281$, circa $85{,}8\%$. Dire che il tentativo usa tre interrogazioni nasconde dunque gran parte del lavoro necessario a realizzarle. Un confronto applicativo dovrebbe aggiungere anche i costi qui esclusi e confrontarsi con un procedimento classico adatto allo stesso problema.

$$C_{O_f} = 2 \cdot 120 + 1 = 241$$

$$C_{\text{totale}} = 10 + 3(241 + 40) = 853$$

Richiami Il numero delle porte dipende dalla realizzazione del verificatore e dal repertorio di porte disponibile. Il conteggio a oracolo descrive una risorsa astratta diversa dal tempo fisico.

Il risultato della ricerca e le sue condizioni

Grover costruisce un'evoluzione che concentra probabilità sull'insieme riconosciuto dal verificatore. La marcatura cambia le fasi, la riflessione rispetto alla preparazione converte quelle differenze in una rotazione e la misura produce un candidato. La verifica finale collega il risultato probabilistico del circuito alla condizione classica che definisce la soluzione.

L'efficacia del procedimento dipende dalla durata scelta, dal peso iniziale delle soluzioni e dalla disponibilità di un verificatore coerente. Gli esempi numerici hanno mostrato sia l'aumento del successo sia il suo successivo calo. L'analisi dei costi ha poi distinto il vantaggio nelle interrogazioni dal lavoro necessario a costruire e usare il circuito. Sono queste condizioni, considerate insieme, a determinare l'utilità della ricerca quantistica in un problema concreto.

Richiami Il successo dopo k iterazioni è $\sin^2((2k+1)\theta)$, con $\sin^2\theta$ uguale al successo iniziale. Ogni candidato restituito viene accettato soltanto dopo la verifica del predicato.

Riferimenti per la ricerca e la sua interpretazione

N. David Mermin, *Quantum Computer Science: An Introduction*, Cambridge University Press, 2007, capitolo 4. Il capitolo sviluppa la ricerca quantistica mediante la costruzione dell'iterazione e la sua interpretazione geometrica. Sono disponibili anche le dispense dell'autore, *Searching with a Quantum Computer*, revisione del 30 marzo 2006: capitolo IV delle lecture notes di Cornell.

Lov K. Grover, *A fast quantum mechanical algorithm for database search*, Proceedings of STOC, 1996, pp. 212–219: arXiv:quant-ph/9605043. Il lavoro introduce l'algoritmo di ricerca e l'interpretazione dell'inversione rispetto alla media. La presentazione adotta esempi numerici e derivazioni originali, mantenendo separati ampiezze, probabilità e costi del modello a oracolo.

Richiami La riflessione uniforme agisce sulle ampiezze mediante $a_x \mapsto 2\bar{a} - a_x$. La probabilità dell'esito x è $|a_x|^2$.

Riferimenti per le estensioni e i limiti di costo

M. Boyer, G. Brassard, P. Høyer e A. Tapp, *Tight bounds on quantum searching*, Fortschritte der Physik 46, 493–506, 1998: arXiv:quant-ph/9605034. Riferimento per la probabilità esatta, le soluzioni multiple e la ricerca con numero di soluzioni ignoto. C. Zalka, *Grover's quantum searching algorithm is optimal*, Physical Review A 60, 2746–2751, 1999: arXiv:quant-ph/9711070, per i limiti del modello a oracolo.

G. Brassard, P. Høyer, M. Mosca e A. Tapp, *Quantum Amplitude Amplification and Estimation*, Contemporary Mathematics 305, 53–74, 2002: arXiv:quant-ph/0005055. Riferimento per la preparazione generale e per l'amplificazione del suo successo. La stima di ampiezza, trattata nello stesso lavoro, è un problema ulteriore rispetto alla ricerca sviluppata qui.

Richiami Per una preparazione generale, p è il peso iniziale dell'insieme valido. Il risultato $O(1/\sqrt{p})$ richiede accesso coerente alla preparazione, al suo inverso e alla marcatura.