1.4 Il problema delle approssimazioni

A causa del numero limitato di cifre a disposizione, un numero si può rappresentare esattamente in un calcolatore solo se la sua parte frazionaria è esprimibile come somma di un numero limitato di potenze di 2. Diversamente lo si approssima a quello a esso piú vicino. Mentre 22.75 era rappresentabile esattamente in altri casi, ad esempio 0.1, non è cosí. Per convincervene supponete di avere a disposizione un computer a 32 bit e provate a scrivere questo numero in notazione binaria.

Per lo stesso motivo è evidentemente impossibile rappresentare i numeri irrazionali (come $\sqrt{2}$, π, etc.): questi hanno infinite cifre dopo la virgola perciò non possono trovare posto nella memoria di un calcolatore. Si possono solo approssimare al numero razionale piú vicino.

Questa limitazione conduce a un problema con il quale occorre misurarsi ogni volta che si debbano eseguire calcoli con un computer: l'errore di arrotondamento. In pratica tutti i numeri non interi si approssimano al numero razionale piú vicino che ammette una rappresentazione finita, con una precisione dell'ordine di una parte su 2^{-n_m}, dove n_m è il numero di bit riservati alla mantissa: due numeri che differiscono tra loro per meno di questa quantità si considerano uguali su un calcolatore e un numero minore del piú piccolo numero rappresentabile equivale a 0 (in questo caso si parla di errore di *underflow*).

Occorre sempre prestare molta attenzione a questo tipo di approssimazione perché, benché in molti casi appaia di piccola entità, l'errore può propagarsi in maniera catastrofica in alcuni algoritmi, specie se di tipo iterativo, e diventare importante (un esempio concreto di questo effetto è descritto nel Capitolo 4). Un caso frequente di errore catastrofico si verifica quando gli operandi sono tra loro molto diversi o molto simili. Nel primo caso il piú piccolo dei due numeri perde precisione a causa dello spostamento della mantissa necessario a portarlo a una rappresentazione con lo stesso esponente del piú grande. Consideriamo la somma tra i numeri $a = 68\,833\,152$, che nella rappresentazione IEEE 754 si scrive 0 1001 1001 000 0011 0100 1001 1111 0000 e $b = 2309\,657\,318\,129\,664$ (0 1011 0010 000 0011 0100 1001 1111 0000). Per poter eseguire questa somma occorre incolonnare i numeri, pertanto il piú piccolo (a) deve essere espresso con lo stesso esponente di 2 del piú grande (b) nella notazione IEEE 754. L'esponente di a è 26, mentre quello di b è 51. La differenza tra questi due numeri è 25. La mantissa di a, dunque, si dovrebbe riscrivere spostandone le cifre di altrettanti posti verso destra. Poiché il numero massimo di cifre per la mantissa è 23, il risultato netto sarà che la mantissa di a sarà rappresentata come una sequenza di zeri:

$$\begin{array}{ll}
0\ \ 1011\,0010\ \ 000\,0000\,0000\,0000\,0000\,0000 & + \ \ldots \\
\underline{0\ \ 1011\,0010\ \ 000\,0011\,0100\,1001\,1111\,0000} & = \\
0\ \ 1011\,0010\ \ 000\,0011\,0100\,1001\,1111\,0000 &
\end{array}$$

e la somma risulta essere $a + b = b$, un'evidente assurdità!

Il secondo genere d'errore si presenta spesso nel calcolo della differenza tra due numeri vicini. Lo si comprende bene attraverso semplici esempi che svolgiamo nella base 10 solo perché piú familiare e dunque di comprensione piú immediata. Supponiamo di rappresentare i numeri in virgola mobile con 3 cifre significative dopo la virgola. Il numero 1000 si rappresenta come $a = 1.000 \times 10^3$, mentre il numero 999.8 come $b = 9.998 \times 10^2$.

La differenza $(a - b) = 0.2$, ma quando questa si calcola in virgola mobile il numero piú piccolo si deve esprimere con una potenza di 10 uguale a quella del piú grande, con conseguente perdita di cifre significative, quindi $b \to b' = 0.999 \times 10^3$ e la differenza diventa

$$(1.000 - 0.999) \times 10^3 = 10^{-3} \times 10^3 = 1.0\,.$$

Il valore approssimato è piú grande di un fattore 5 rispetto al valore vero, benché la precisione sul singolo numero sia dell'ordine di 10^{-3}! In taluni casi questo genere d'errore si può evitare (o almeno ridurre) riformulando l'espressione da calcolare. Supponiamo, dati $x = 3.451 \times 10^0$ e $y = 3.45 \times 10^0$, di dover eseguire l'operazione $\Delta = \left(x^2 - y^2\right) = 0.006\,901 = 6.901 \times 10^{-3}$. Se si calcolano prima i quadrati $x^2 = 11.909\,401 = 1.190 \times 10^1$ e $y^2 = 11.9025 = 1.190 \times 10^1$, si trova che $\Delta = 0$. Un risultato davvero disastroso, specialmente se Δ rappresenta il denominatore di un'espressione. Se si riformula l'espressione come $\Delta = \left(x^2 - y^2\right) = (x - y)\,(x + y)$, invece, si trova che $(x - y) = (3.451 - 3.450) \times 10^0 = 0.001 \times 10^0 = 1.000 \times 10^{-3}$, mentre $(x + y) = (3.451 + 3.450) \times 10^0 = 6.901 \times 10^0$, quindi $\Delta = 1.000 \times 10^{-3} \times 6.901 \times 10^0 = 6.901 \times 10^{-3}$.

Nel calcolo numerico, diversamente da quello analitico, l'ordine in cui si eseguono le operazioni è di estrema importanza, cosí come l'ordine di grandezza dei termini soggetti alle operazioni aritmetiche che non devono essere troppo diversi tra loro, né troppo simili, secondo le operazioni svolte su di essi.

1.5 Non-numeri sul calcolatore

Lo svolgimento di un compito, sia esso a carattere scientifico o no, può richiedere la manipolazione di dati di natura diversa dai numeri, come immagini, suoni, etc. Oggi questo può sembrare ovvio perché costantemente sugli schermi dei computer si possono osservare immagini o seguire un film; attraverso gli altoparlanti si può ascoltare la musica; con un microfono si può comunicare con la voce con altri utenti remoti attraverso la rete; con una telecamera si può fare una videoconferenza.

Quando nacquero, negli anni '40 del secolo scorso, i computer servivano esclusivamente a eseguire calcoli. I numeri rappresentavano praticamente l'unica forma d'informazione disponibile. I programmi si impostavano azionando interruttori e i risultati erano visualizzati da lampadine il cui stato (acceso/spento) indicava la sequenza di bit nella rappresentazione binaria. Quando fu introdotta la possibilità di produrre i risultati in forma scritta (su carta e, successivamente, su schermo) e d'inserire i comandi mediante una tastiera si pose il problema di rappresentare altri tipi d'informazione come i caratteri. La necessità di rappresentare informazioni di natura diversa aumentò costantemente nel