DENTRO LA SCATOLA

Rubrica a cura di

Fabio A. Schreiber

Il Consiglio Scientifico della rivista ha pensato di attuare un'iniziativa culturalmente utile presentando in ogni numero di Mondo Digitale un argomento fondante per l'Informatica e le sue applicazioni; in tal modo, anche il lettore curioso, ma frettoloso, potrà rendersi conto di che cosa sta "dentro la scatola". È infatti diffusa la sensazione che lo sviluppo formidabile assunto dal settore e di conseguenza il grande numero di persone di diverse estrazioni culturali che - a vario titolo - si occupano dei calcolatori elettronici e del loro mondo, abbiano nascosto dietro una cortina di nebbia i concetti basilari che lo hanno reso possibile. La realizzazione degli articoli è affidata ad autori che uniscono una grande autorevolezza scientifica e professionale a una notevole capacità divulgativa.

L'aritmetica dei residui e il suo uso per la realizzazione di unità aritmetiche specializzate particolarmente veloci

Renato Stefanelli

1. Introduzione

Una contadina desiderava sapere quante fossero le uova contenute in un cesto, ma sapeva contare solo fino a sette, e le uova erano sicuramente più di sette; allora dal cesto ne tolse tre, poi altre tre, poi ancora altre tre finché non ne rimasero che due; rimise le uova nel cesto e poi da questo tolse a più riprese gruppi di cinque uova fino a che non ne rimasero che quattro; poi rifece l'operazione togliendone a gruppi di sette finché non ne rimase che una.

Il matematico cinese Sun Tsu Suan-Ching, circa 1650 anni fa si domandò se tali risultati definissero in modo univoco il numero di uova nel cesto e arrivò alla seguente conclusione:

il cesto contiene 29 uova. Infatti, chiamando Q_1, Q_2, Q_3 *i gruppi di tre, cinque e sette uova, e chiamando N il numero di uova nel cesto, si ha*: $N = 3\,Q_1 + 2$; $N = 5\,Q_2 + 4$; $N = 7\,Q_3 + 1$ *e il numero* 29, *diviso per* 3, 5, 7 *fornisce resti* 2, 4, 1.

Per ottenere questo risultato egli definì il concetto di "residuo", ne studiò le proprietà ed enunciò e dimostrò il "teorema cinese sui residui", la base per la moderna aritmetica dei residui. Ovviamente Sun Tsu Suan-Ching non immaginava certo l'applicazione che le sue idee avrebbero avuto 1650 anni dopo. Nell'ultimo mezzo secolo, infatti, si è verificato come tali idee potessero essere applicate con successo alla realizzazione di unità aritmetiche particolarmente veloci.

Questo articolo tratta, in primo luogo, la definizione di residuo e le sue proprietà aritmetiche; si enuncerà il teorema cinese dei residui, già citato in precedenza; poi se ne studieranno le applicazioni; in particolare, si studierà il modo di ottenere strutture aritmetiche particolarmente veloci e si accennerà ai settori ove tale rappresentazione può avere un effettivo interesse (poco per unità aritmetiche di calcolatori, molto in unità aritmetiche specializzate per elaborazione di segnali e immagini).

2. I residui e loro uso per rappresentare numeri interi positivi

Si consideri un numero intero positivo N e lo si divida per una base intera positiva b; si ottiene

un quoziente Q e un resto R tali per cui $N = Q\,b + R$ ove il resto R, compreso tra 0 e $b - 1$, viene chiamato "residuo del numero N rispetto alla base b" e viene indicato come $|N|_b$. Per esempio il residuo del numero 29 rispetto alla base 3 vale $|29|_3 = 2$ in quanto $29 = 9 \cdot 3 + 2$. Ciò definisce un singolo residuo del numero N rispetto a una sola base b. Si considerino ora tre basi $b_1 = 3$, $b_2 = 5$ e $b_3 = 7$; è allora possibile calcolare i tre residui del numero N rispetto alle tre basi. Si può verificare che, se N è inferiore a un certo limite (105 nell'esempio) e se le tre basi sono state scelte in modo opportuno (come si vedrà in seguito) allora esiste una relazione biunivoca tra il numero e la terna dei suoi residui. Si consideri, infatti, la tabella 1 dove sono riportate le terne di residui per le basi 3, 5, 7 e per numeri tra zero e 106; si verifica intanto che al numero 29 corrisponde la terna 2, 4, 1 e a nessun altro numero corrisponde la stessa terna (in realtà, per questioni di spazio, la tabella riporta solo alcuni casi; pregherei il lettore di credere all'affermazione precedente). Si verifica poi che la relazione biunivoca è presente solo per numeri inferiori a 105; infatti, a partire dal numero 105, la tabella si ripete fino al numero 208 ecc..

La condizione che deve essere verificata nella scelta delle basi è che queste siano "prime tra loro a coppie"; se ciò è verificato allora il limite di rappresentabilità del numero N corrisponde al prodotto delle basi; nell'esempio, $3 \cdot 5 \cdot 7 = 105$. La condizione è sicuramente verificata nell'esempio in quanto si sono scelte basi rappresentate da numeri primi; ma andrebbero bene anche basi come (8, 9, 11) oppure (6, 7, 11, 25) dove alcune delle basi non sono numeri primi ma ogni coppia di basi ha massimo comune divisore uguale ad uno.

Si esamineranno ora le proprietà aritmetiche dei residui, la conversione da binario a residuo e quella inversa da residuo a binario e le applicazioni alle unità aritmetiche veloci.

3. Proprietà aritmetiche dei residui

Le proprietà aritmetiche dei residui, per le operazioni di somma e di prodotto, si basano sul seguente teorema:

Si considerino due numeri X e Y rappresentati rispettivamente dai residui $x_1 \ldots x_n$ e $y_1 \ldots y_n$ nelle basi $b_1 \ldots b_n$; sia inoltre Z la somma (il prodotto) di X e Y e Z abbia i seguenti residui $z_1 \ldots z_n$; allora, per ogni base b_i, il risultato Z ha residui pari, al residuo della somma (del prodotto) dei residui degli addendi (fattori), ossia $z_i = |x_i + y_i|_{bi}$ ($z_i = |x_i \cdot y_i|_{bi}$). Come esempio si consideri il caso $3 + 4 = 7$ per la somma e $2 \cdot 4 = 8$ per il prodotto: esaminando la tabella si ha:

$$|0+1|_3 = 1 = |7|_3 ; |3+4|_5 = 2 = |7|_5 ;$$

$$|3+4|_7 = 0 = |7|_7$$

$$|2 \cdot 1|_3 = 2 = |8|_3 ; |2 \cdot 4|_5 = 3 = |8|_5 ;$$

$$|2 \cdot 4|_7 = 1 = |8|_7$$

Ne risulta, quindi, che, per fare la somma e il prodotto di due numeri, è sufficiente eseguire la somma e il prodotto, in modulo, dei residui omologhi. Attenzione: se somme e prodotti possono essere eseguiti in "modo facile", non esistono proprietà analoghe per eseguire il confronto tra due numeri, il rapporto tra due numeri oppure, cosa abbastanza complicata, il controllo se il risultato di un'operazione ha superato o meno i limiti di rappresentabilità. Questi sono ancora argomenti di ricerca e tuttora non si conoscono algoritmi efficienti per la loro implementabilità.

4. La conversione dalla rappresentazione binaria a quella a residui e quella contraria da residui a binario

Si suppone di partire da dati in ingresso a una unità aritmetica rappresentati in binario e si desidera

basi\N	0	1	2	3	4	5	6	7	8	9	..	28	29	30	..	103	104	105	106
3	0	1	2	0	1	2	0	1	2	0	..	1	2	0	..	1	2	0	1
5	0	1	2	3	4	0	1	2	3	4	..	3	4	0	..	3	4	0	1
7	0	1	2	3	4	5	6	0	1	2	..	0	1	2	..	5	6	0	1

TABELLA 1

Terne di residui per basi 3, 5, 7

che il risultato di un'operazione aritmetica sia ancora in formato binario. La conversione da binario a residuo con n basi può essere ottenuta mediante n circuiti divisori; si può notare però che in tali circuiti il divisore è un numero fisso e relativamente piccolo; la loro complessità è, quindi, inferiore a quella di un circuito divisore di tipo generale. Tali circuiti hanno complessità circuitale (area di silicio) e tempi di ritardo paragonabili a quelli di un normale addizionatore binario.
Più complessa è la conversione opposta, che costituisce il "teorema cinese sui residui", il contributo principale del matematico cinese Sun Tsu Suan-Ching.
Si fa l'esempio, per semplicità, delle tre basi $b_1 = 3$, $b_2 = 5$, $b_3 = 7$ (facile è l'estensione a basi in numero e valore qualsiasi). Si suppone di aver precalcolato tre numeri D_1, D_2, D_3 tali per cui:

$$|D_1 b_2 b_3|_{b1} = 1 \quad |b_1 D_2 b_3|_{b2} = 1 \quad |b_1 b_2 D_3|_{b3} = 1$$

Nel caso delle tre basi 3, 5, 7, per tentativi si trova $D_1 = 2$, $D_2 = 1$, $D_3 = 1$.
Si consideri ora il numero $\underline{A} = a_1 D_1 b_2 b_3 + a_2 b_1 D_2 b_3 + a_3 b_1 b_2 D_3$. Per verificare che tale numero rappresenti il risultato richiesto, se ne calcolano i tre residui onde controllare che questi siano uguali a quelli iniziali a_1, a_2, a_3 e che il numero calcolato sia compreso entro i limiti di rappresentabilità.
Si lascia, per brevità, al lettore la prima parte (calcolo dei tre residui) ricordando che il residuo della somma (prodotto) è il residuo della somma (prodotto) dei residui, che un termine proporzionale a una base b_i ha residuo nullo rispetto alla stessa base e che ogni residuo a_i è inferiore alla propria base b_i e quindi $|a_i|_{bi} = a_i$
Per verificare se $\underline{A}$ è compreso entro i limiti di rappresentabilità si considera l'esempio dei residui 2, 4, 1 nelle basi 3, 5, 7 che, dalla tabella, dovrebbe corrispondere al numero $A = 29$. Si ha:

$$\underline{A} = 2 \cdot 2 \cdot 5 \cdot 7 + 4 \cdot 1 \cdot 3 \cdot 7 + 1 \cdot 1 \cdot 3 \cdot 5 = 140 + 84 + 15 = 239,$$

numero che esce dai limiti di rappresentabilità; ma il numero 239 ha residuo 29 rispetto al prodotto M delle tre basi, pertanto:

$$A = |A|_M \text{ ove } M = b_1 \cdot b_2 \cdot b_3 = 105;$$

infatti, $239 = 2 \cdot 105 + 29$

5. Applicazione alle unità aritmetiche

Si desideri eseguire la somma (o il prodotto) di due numeri rappresentati in binario e si desideri che il risultato sia in forma binaria. Per eseguire tale operazione occorre:
1. Convertire i due numeri in residui: occorrono $2n$ circuiti di conversione per i due numeri e per le n basi (ogni circuito di conversione "lavora" indipendentemente dagli altri, quindi, i $2n$ circuiti operano contemporaneamente). L'operazione è *lenta* (tempi paragonabili a quelli di un addizionatore binario).
2. Eseguire l'operazione aritmetica: occorrono n sommatori (moltiplicatori), ciascuno che elabori due residui omologhi per produrre un residuo del risultato. Si noti che, dato che le operazioni avvengono "in modulo", ogni sommatore (moltiplicatore) è diverso dagli altri. Anche in questo caso, ogni circuito aritmetico opera indipendentemente dagli altri e, quindi, possono "lavorare" contemporaneamente (non si ha riporto da un circuito all'altro) e questo è il motivo principale di un eventuale vantaggio della rappresentazione a residui rispetto a quella tradizionale binaria. L'operazione aritmetica vera e propria è *velocissima* in quanto gli n circuiti aritmetici operano in parallelo su numeri costituiti da pochi bit.
3. Ottenuti i residui del risultato occorre riconvertire questo alla rappresentazione binaria e tale operazione richiede molte moltiplicazioni e somme, su numeri molto grandi, e il calcolo di un residuo rispetto al prodotto delle basi, normalmente un numero grande. Nel complesso la terza operazione è fortemente onerosa sia rispetto alla complessità (area di silicio), sia rispetto ai tempi, per cui la si puó indicare come *lentissima*.
Ma allora, se un'operazione con i residui è fortemente svantaggiosa sia come area di silicio che, e soprattutto, per la velocità, a cosa serve e dove può essere applicata?; si ha l'impressione di aver sparato a una mosca con un cannone (e forse senza essere sicuri di aver centrato la mosca!).
Si consideri l'unità aritmetica di un normale calcolatore: questa deve eseguire un'operazione singola per ogni istruzione di tipo aritmetico; in questo caso, l'unità aritmetica risulterebbe estremamente lenta rispetto a una soluzione tradizionale. Inoltre, tale unità

dovrebbe poter eseguire anche divisioni e confronti; nella rappresentazione mediante residui queste due operazioni richiederebbero una conversione da residui a binario per poterle eseguire in modo tradizionale. Il tutto sarebbe fortemente svantaggioso, e questa è la ragione per cui i residui non hanno avuto applicazione nelle unità aritmetiche dei calcolatori.

Si consideri ora un'unità aritmetica per un'applicazione "speciale", come, per esempio, un filtro digitale per segnali, una trasformata di Fourier per immagini bidimensionali ecc.: in questo caso, si hanno relativamente pochi dati in ingresso da convertire da binario a residui, relativamente pochi dati in uscita da convertire da residuo a binario, mentre si hanno molte operazioni di somme e prodotti da eseguire in modo veloce nel passo 2 precedentemente definito (in questo caso le *n* unità aritmetiche che operano contemporaneamente nella fase 2 corrisponderebbero non al calcolo di una sola somma o moltiplicazione, ma al calcolo completo dell'algoritmo di filtraggio o della trasformata di Fourier); inoltre, non si hanno operazioni di divisione e di confronto. In questo caso, la fase 2 risulta preponderante rispetto alle fasi 1 e 3; pertanto, si può avere un effettivo vantaggio, in termini di velocità, dall'uso della rappresentazione mediante residui. In questa ultima frase consiste l'effettivo interesse per la rappresentazione dei numeri mediante residui.

Bibliografia

Esistono diversi libri che trattano l'argomento generale dell'aritmetica; alcuni di questi hanno un capitolo dedicato all'aritmetica dei residui; si indica qui un testo che tratta l'argomento in modo succinto ma chiaro:

Koren I.: *Computer arithmetic algorithms*. Prentice Hall, 1993.

Esiste poi una diffusissima bibliografia specifica: cercando su internet con Google.com usando come frase chiave "Chinese Remainder Theorem" si trovano 12.900 siti; se la ricerca avviene sulla frase "Residue Number System" si trovano ben 464.000 siti! Il lettore interessato si armi di tanta pazienza!

Renato Stefanelli Professore ordinario di Calcolatori Elettronici presso la Facoltà di Ingegneria del Politecnico di Milano. È autore di circa 160 pubblicazioni principalmente sulla tolleranza ai guasti sia in unità aritmetiche che in sistemi regolari di multiprocesssori, sull'elaborazione di immagini, su circuiti aritmetici ad elevata velocità, su circuiti aritmetici basati sull'uso di residui. Ha diretto per diversi anni un Centro di Studio del CNR sull'Informatica presso il Politecnico di Milano.
stefanel@elet.polimi.it