A SUBWORD VERSION OF D'OCAGNE'S FORMULA

Helmut Prodinger

The base two version of d'Ocagne's formula reads as follows: Let $B(n)$ be the sum of digits of $(n)_2$, where $(n)_2$ means the binary representation of n. If $x = 2^{e_1} + 2^{e_2} + \ldots + 2^{e_r}$, where $e_1 > e_2 > \ldots > e_r \geq 0$, then

$$(1) \qquad \sum_{1 \leq n < x} B(n) = \sum_{1 \leq i \leq r} (e_i + 2i - 2)\, 2^{e_i - 1} .$$

For this result we refer to [4]; that paper contains an extensive list of references on digital sums and related topics.

Let us recall that $B(n)$ is just the number of 1's in $(n)_2$. In this note we are counting subwords rather than the occurrences of the digit 1. To be more precise, let $B_s(n)$ be the number of subword occurrences of 11...1 (s consecutive 1's) in $(n)_2$. Compare [1,2,3] for some recent results on that subject.

If x is represented as above, we prove the following subword version of d'Ocagne's formula:

THEOREM.

$$(2) \qquad \sum_{0 \leq n < x} B_s(n) = \sum_{1 \leq i \leq r} 2^{e_i} \left[\frac{e_i \dot{-} (s-1)}{2^s} + \sum_{1 \leq k \leq i-s} \delta_k^{(s)} \right]$$

where for $s \geq 2$ $\delta_k^{(s)} = 1$ *if* $e_j = e_{j+1} + 1$ *for* $k \leq j < k+s-1$ *and* $\delta_k^{(s)} = 0$ *otherwise;* $\delta_k^{(1)} = 1$ *for* $1 \leq k \leq r$. *The empty sum is to be interpreted as* 0 *and* $a \dot{-} b$ *means that* $a-b$ *is to be replaced by* 0 *if* $a-b<0$.

To prove this Theorem we first need a lemma.

UTILITAS MATHEMATICA, Vol. 24 (1983), pp. 125-129

LEMMA.

(3) $$\chi := \sum_{0\le n<2^m} B_s(n) = 2^{m-s}[m\dot{-}(s-1)] .$$

Proof. This result appears implicitly in [1,2,3] as a special case of a more general result. However, it may be interesting to give a more direct proof in this case. Assume $m \ge s$. Let $[x^n]f$ denote the coefficient of x^n in the (formal) power series f. We need the following formula:

(4) $$\sum_{j\ge 1} \left(j\dot{-}(s-1)\right) x^j = \frac{x^s}{(1-x)^2} .$$

The numbers n in the sum (3) are, written in the binary representation, just the words with the letters 0 and 1 of length m. The set of these words is now partitioned according to the blocks of 0's and 1's. We may write $(n)_2 = 0^{i_0}1^{j_1}0^{i_1}1^{j_2}\ldots 1^{j_t}0^{i_t}$ with $t \ge 1$, $i_0+j_1+i_1+\ldots+j_t+i_t = m$; $i_0, i_t \ge 0$, $i_k \ge 1$ $(1\le k<t)$, $j_k \ge 1$ $(1\le k\le t)$. ($n=0$ is not obtained in this way, but $B_s(0)$ gives the contribution 0.) Hence we may write

(5) $$\chi = \sum_{t\ge 1} \sum_{*} \left[(j_1 \dot{-} (s-1)) + \ldots + (j_t \dot{-} (s-1))\right] = \sum_{t\ge 1} t\cdot\xi .$$

Here, "*" is an abbreviation for "$i_0+j_1+i_1+\ldots+j_t+i_t = m$; i_0, $i_t \ge 0$, $i_k \ge 1$ $(1\le k<t)$, $j_k \ge 1$ $(1\le k\le t)$" and

$$\xi = \sum_{*} (j_1 \dot{-} (s-1))$$

$$= [x^m] \sum_{i_0\ge 0} x^{i_0} \cdot \sum_{j_1\ge 1} (j_1 \dot{-} (s-1))x^{j_1} \cdot \sum_{i_1\ge 1} x^{i_1} \ldots \sum_{j_t\ge 1} x^{j_t} \cdot \sum_{i_t\ge 0} x^{i_t}$$

(6) $$= [x^m] \frac{1}{1-x} \cdot \frac{x^s}{(1-x)^2} \cdot \frac{x}{1-x} \cdot\ldots\cdot \frac{x}{1-x} \cdot \frac{1}{1-x}$$

$$= [x^m] \frac{1}{(1-x)^2} \frac{x^s}{(1-x)^2} \left(\frac{x}{1-x}\right)^{2t-2} = [x^m] \frac{x^{2t-2+s}}{(1-x)^{2t+2}}$$

$$= [x^{m+2-s-2t}] (1-x)^{-2t-2} = \binom{m+3-s}{m+2-s-2t} = \binom{m+3-s}{2t+1} .$$

We have used that $(1-x)^{-\alpha} = \sum_{k\geq 0} \binom{\alpha+k-1}{k} x^k$.

Now

$$\eta = \sum_{t\geq 1} t\binom{M+2}{2t+1} = \sum_{t\geq 1} \left(t+\frac{1}{2}\right)\binom{M+2}{2t+1} - \frac{1}{2}\sum_{t\geq 1}\binom{M+2}{2t+1}$$

(7)

$$= \frac{1}{2}(M+2)\sum_{t\geq 1}\binom{M+1}{2t} - \frac{1}{2}\sum_{t\geq 1}\binom{M+2}{2t+1} .$$

Since

(8) $$\sum_{t\geq 1}\binom{M+1}{2t} = \sum_{t\geq 1}\left[\binom{M}{2t} + \binom{M}{2t-1}\right] = \sum_{t\geq 1}\binom{M}{t} = 2^M - 1$$

and

(9) $$\sum_{t\geq 1}\binom{M+2}{2t+1} = \sum_{t\geq 1}\left[\binom{M+1}{2t+1} + \binom{M+1}{2t}\right] = \sum_{t\geq 2}\binom{M+1}{t} = 2^{M+1} - M - 2,$$

we find

(10) $$\eta = \frac{1}{2}(M+2)\left(2^M - 1\right) - \frac{1}{2}\left(2^{M+1} - M - 2\right) = M.2^{M-1} .$$

Using (10) with $M = m - (s-1)$,

$$\chi = \sum_{t\geq 1} t\binom{m+3-s}{2t+1} = (m-(s-1))\, 2^{m-s} . \qquad \square$$

Now we are ready for the proof of the Theorem:

$$\sum_{0\leq n<x} B_s(n) = \sum_{0\leq n<2^{e_1}} B_2(n) + \sum_{2^{e_1}\leq n<x} B_s(n)$$

(11)

$$= (e_1 \dot{-} (s-1))2^{e_1-s} + \sum_{0\leq n<x-2^{e_1}} B_s(n) + \delta_1^{(s)}\left(x-2^{e_1}-\ldots 2^{e_s}\right) .$$

Iterating (11) we have

(12) $$\sum_{0\leq n<x} B_s(n) = \sum_{1\leq i\leq r} (e_i \dot{-} (s-1))2^{e_i-s} + \sum_{1\leq i\leq r-s} \delta_i^{(s)}\left(x-2^{e_1}-\ldots-2^{e_{s-1+i}}\right).$$

The second sum in (12) is

$$(13) \qquad \sum_{1\le i\le r-s} \delta_i^{(s)} \left(2^{e_{s+i}} + \ldots + 2^{e_r}\right) = \sum_{s<j\le r} \left(\delta_1^{(s)} + \ldots + \delta_{j-s}^{(s)}\right) 2^{e_j} ;$$

thus the proof is finished. □

It should be possible to derive similar formulas for more general subwords and more general number systems. This will be done in another paper.

REFERENCES

[1] P. Kirschenhofer, *Subblock occurrences in the q-ary representation of* n, SIAM J. Algebraic and Discrete Methods, to appear (1982).

[2] H. Prodinger, *Generalizing the sum of digits function,* SIAM J. Algebraic and Discrete Methods *3* (1982), 35-42.

[3] H. Prodinger, *Subblock occurrences in representations of integers,* preprint, TU Wien (1981).

[4] K. B. Stolarsky, *Power and exponential sums of digital sums related to binomial coefficient parity,* SIAM J. Applied Math. *32* (1977) 717-730.

Institut für Algebra und Diskrete Mathematik
Technische Universität Wien
Gusshausstrasse 27-29, A-1040
Austria

Received April 17, 1982; Revised September 21, 1982.