A Combinatorial Proof of the Reed-Dawson Identity DRAFT

Gara Pruesse

Department of Computing Science

Malaspina University-College, Nanaimo, BC V9R 5S5 Canada

email: pruesseg@mala.bc.ca

January 23, 2007

First we give a combinatorial proof of the following:

Lemma 1 *The number of words of length n over the alphabet $\Gamma = \{a, b, c, d\}$ where the number of a's is equal to the number of b's is $\binom{2n}{n}$.*

Proof Encode the symbols of the alphabet as follows: $a = 00, b = 11, c = 01, d = 10$. Observe that a bitstring encoding of a word with the same number of a's as b's will have the same number of 0's as 1's. Indeed, it is easy to see that there is a 1-1 correspondence between the bitstrings of length $2n$ that have the same number of 0's and 1's, and the words over Γ that have the same number of a's as b's. The number of such bitstrings is $\binom{2n}{n}$. □

Corollary 1
$$\sum_{k\geq 0} \binom{n}{2k}\binom{2k}{k} 2^{n-2k} = \binom{2n}{n}.$$

Proof Counting the words of Lemma 1 by summing over k, the number of a's (and therefore the number of b's) gives the left hand side of the equation. □

Lemma 2
$$\sum_{k\geq 0} \binom{n}{k}\binom{2k}{k} (-2)^{n-k} = \begin{cases} 0 & \text{if } n \text{ odd} \\ \binom{n}{n/2} & \text{if } n \text{ even.} \end{cases}$$

Proof. Consider the words over the alphabet $\{a, b, c, d, c', d'\}$ that have the same number of a's as b's; let L be the language of such words. A word in L is considered

even if it has an even number of letters from $\{a, b, c, d\}$, and *odd* otherwise. The characteristic of being even or odd is the *parity* of a word. Let d be the number of even words minus the number of odd words.

One way to compute d is to sum, for each k, the number of words with k symbols from $\{a, b, c, d\}$, and multiply by -1 if k is odd. For a given k, we choose from the n places the k places to be filled from $\{a, b, c, d\}$. From Lemma 1, there are $\binom{2k}{k}$ ways to fill those k places so that a and b appear the same number of times. Then we select a subset of the $n - k$ remaining places to be filled by c'; the places yet unfilled after this are filled by d'. Hence

$$d = \sum_{k \geq 0} \binom{n}{k} \binom{2k}{k} 2^{n-k} \times (-1)^k.$$

We find a partial matching ϕ as follows: given a word w we scan it from the left until we encounter the first letter from $\{c, d, c', d'\}$. If the letter is c, change it to c', if the letter is d change it to d'; if the letter is c', change it to c, if the letter is d' change it to d. The resulting word is $\phi(w)$. Clearly w and $\phi(w)$ have different parity, except when w is a fixpoint of ϕ – that is, when w contains no symbols from $\{c, d, c', d'\}$.

Hence to compute d we need only consider the fixpoints of ϕ, i.e., the strings over $\{a, b\}$ of length n with the same number of a's as b's. If n is even, that number is $\binom{n}{n/2}$. If n is odd, there are no such strings. □