PROOF OF (THE INDUCTION STEP OF) THE SAUER'S LEMMA

Hsuan-Tien Lin

Setup

1. A set of N-bit vectors of $+/-$ is called a N-long set. For instance, the following matrix represents a 4-long set, where each row is an element of the set:

$$\begin{array}{cccc} + & + & + & + \\ + & + & - & + \\ + & - & - & + \\ - & + & + & - \end{array}$$

2. An N-long set is said to be **complete** if it includes all possible 2^N distinct vectors. Otherwise it is said to be **incomplete**. For instance, the following matrix represents an incomplete set:

$$\begin{array}{cc} + & + \\ + & - \\ - & - \end{array}$$

3. An N-long set is said to be M-**incomplete** if projecting the vectors to any M dimensions results in an incomplete set. For instance, by projecting the 4-long set above to any of the two columns (dimensions), we see that the set is 2-incomplete:

1	2	1	3	1	4	2	3	2	4	3	4
+	+	+	+	+	+	+	+	+	+	+	+
+	-	+	-	-	-	+	-	-	+	-	+
-	+	-	+			-	-	+	-	+	-

4. $B(N, M)$ is defined to be the maximum number of unique elements in an N-long and M-incomplete set. For instance, the following matrix represents a set that achieves $B(4, 2)$:

$$\begin{array}{cccc} + & + & + & + \\ + & + & - & + \\ + & - & - & + \\ - & + & + & - \\ + & + & + & - \end{array}$$

The Main Lemma

Lemma 1 *(The induction step)*

$$B(N, M) \leq B(N-1, M) + B(N-1, M-1).$$

Proof. Consider the set S that achieves $B(N, M)$. We first project the vectors in S into the first $N-1$ dimensions to get $V = \{v_i\}$, where v_i's are unique. For instance, consider the set that achieves $B(4, 2)$ above, after projecting we get:

$$\begin{array}{c|ccc} v_1 & + & + & + \\ v_2 & + & + & - \\ v_3 & + & - & - \\ v_4 & - & + & + \end{array}$$

We can then separate V to three disjoint subsets:

- A_1: there is only $(v_i, +)$ in S, but no $(v_i, -)$. For instance, $\{v_2, v_3\}$.
- A_2: there is only $(v_i, -)$ in S, but no $(v_i, +)$. For instance, $\{v_4\}$.
- A_3: both $(v_i, +)$ and $(v_i, -)$ are in S. For instance, $\{v_1\}$.

By reorganizing the rows, we get

$$S = \begin{bmatrix} A_1 & + \\ A_2 & - \\ A_3 & + \\ A_3 & - \end{bmatrix} \quad ; \quad V = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \quad .$$

Now let $a = |A_1| + |A_2|$, and $b = |A_3|$. We see that[1]

$$|V| = a + b \quad ; \quad B(N, M) = |S| = a + 2b. \tag{1}$$

1. We first look at V:

$$V = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \quad .$$

If V is not M-incomplete, obviously S is also not M-incomplete—**a contradition**!. Thus, V must be M-incomplete, and

$$|V| \le B(N-1, M). \tag{2}$$

2. We now look at the subset (submatrix)

$$S_3 = \begin{bmatrix} A_3 & + \\ A_3 & - \end{bmatrix}$$

If A_3 is not $(M-1)$-incomplete, then S_3 (and hence S) is not M-incomplete—**a contradition**! Thus, A_3 must be $(M-1)$-incomplete, and

$$b = |A_3| \le B(N-1, M-1). \tag{3}$$

By combining (1), (2), and (3), we get the desired result.

Lemma 2 *(Sauer's Lemma)*

$$B(N, M) \le \sum_{m=0}^{M-1} C(N, m) \le N^{M-1} + 1.$$

Proof. The first inequality can be proved using mathematical induction (with Lemma 1). The second inequality can be proved using mathematical induction, too.

[1] Here $|\cdot|$ means the size of the set, or (equivalently) the number of rows in the representing matrix.