Some Run Problems in Hypercubes (Tic-tac-toe)

by

Eric M. Sobel and Milton Sobel*

Technical Report 142

University of Minnesota

Minneapolis, Minnesota 55455

July 1970

* Research supported by NSF Grant GP-11021

Some Run Problems in Hypercubes (Tic-tac-toe)

Eric M. Sobel and Milton Sobel*

Introduction

Several combinatorial questions are studied which arise in connection with runs in a n-dimensional hypercube $H^n(s)$ with $s > 1$ units along each edge. These questions are related to a 1-player version of the tic-tac-toe game in $H^n(s)$. A run is defined as any straight 'line' containing s cells. For example with $n = 3$ any one of the $3s^2$ row, column and vertical tubes, any one of the $6s$ diagonals parallel to a face (i.e. 2-diagonals) or any one of the 4 diagonals connecting 2 opposite vertices (i.e. 3-diagonals) are examples of runs. It follows that the total number of runs $R_1(n, s)$ for $n = 3$ and $s > 1$ is given by

$$(1) \qquad R_1(3, s) = 3s^2 + 6s + 4 = 3(s + 1)^2 + 1 = \tfrac{1}{2}[(s + 2)^3 - s^3].$$

One goal of this paper is to generalize this result and extend it in certain directions. A generalized run (or d-dimensional run with $d \leq n$) is a set of s^d cells lying in a d-dimensional flat in the hypercube $H^n(s)$, i.e. in the intersection of some d-dimensional hyperplane and the hypercube.

Let $Q_1, Q_2, \ldots$ denote the different questions posed as well as section numbers.

<u>Q_1</u>: What is the total number $R_d(n, s)$ of d-dimensional runs in the hypercube $H^n(s)$?

For each of the $\binom{n}{1}$ individual axes we can find s^{n-d} runs by varying only one coordinate. For each of the $\binom{n}{2}$ pairs of axes we can find $2s^{n-d-1}$ such runs by keeping n-d-1 coordinates fixed and for the two changing coordinates we can have them going in the same or opposite directions. For each of the $\binom{n}{3}$ sets of 3 axes we can find $2^2 s^{n-d-2}$ such runs. To explain the factor 2^2, note that if the first of these 3

*/Research supported by NSF Grant GP-11021, Statistics Dept., University of Minnesota.

coordinates is increasing then each of the other two can be the same or opposite. Thus for $1 \leq d < n$ and $s > 1$ we have the result

(2) $$R_d(n, s) = \binom{n}{1}s^{n-d} + 2\binom{n}{2}s^{n-d-1} + \cdots + 2^{n-d}\binom{n}{n-d+1}s^0$$

$$= \frac{1}{2s^{d-1}} \sum_{i=1}^{n-d+1} \binom{n}{i} 2^i s^{n-i}$$

$$= \frac{1}{2s^{d-1}} \left\{(s+2)^n \, I_{\frac{s}{s+2}}(d-1, n-d+1) - s^n\right\},$$

where $I_p(a, b)$ is the usual incomplete beta function and $I_p(0, b) = 1$ for all $b > 0$. Clearly $R_n(n, s) = 1$ for all s.

Corollary 1: For $d = 1$ and $s > 1$ we obtain the simple result

(3) $$R_1(n, s) = \frac{(s+2)^n - s^n}{2}.$$

The formula (3) was also obtained by A. L. Rubinoff, L. Moser and N. J. Fine in [1] where an elegant geometric proof due to L. Moser is given.

Corollary 2: For $d = 2$ and $s > 1$ we obtain for $n > 2$

(4) $$R_2(n, s) = \frac{1}{2s}\left\{(s+2)^n - s^n - 2^n\right\}$$

and for $n = 3$ this reduces to $R_2(3, s) = 3s + 6$, which is easily verified.

The following tables gives some numerical results

	d = 1					d = 2				
s \ n	1	2	3	4	5	1	2	3	4	5
2	1	6	28	120	496	0	1	12	56	240
3	1	8	49	272	1441	0	1	15	88	475
4	1	10	76	640	3776	0	1	18	128	840
5	1	12	109	1188	6841	0	1	21	176	1365

from (3) and (4) for the number of d-dimensional runs in the hypercube $H^n(s)$

for $d = 1, 2$, $n = 1(1)5$ and $s = 2(1)5$.

Q_2: How can we characterize an arbitrary cell in $H^n(s)$ and write an expression for the multiplicity of that cell, i.e., for the number of 1-dimensional runs it is contained in?

We first consider the case of s even and later the case of s odd. For s even let $s = 2m$ and assign to the s^n cells (starting near the center) the 'coordinates' $\pm 1, \pm 2, \ldots, \pm m$ so that the cells are designated by $(i_1, i_2, \ldots, i_n)$, where each i_α runs from $-m$ to $+m$, but is not zero. Disregarding the sign, let α_i denote the total number of coordinates equal to i or $-i$ $(i = 1, 2, \ldots, m)$ and let $\underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m)$, with $\alpha_1 + \alpha_2 + \ldots + \alpha_m = n$, denote the structure of the cell. Let $r(\alpha_1, \alpha_2, \ldots, \alpha_m) = r(\underset{\sim}{\alpha})$ denote the number of runs associated with a cell of structure $\underset{\sim}{\alpha}$.

The number of runs through any cell with structure $\underset{\sim}{\alpha}$ is

$$(5) \qquad r(\underset{\sim}{\alpha}) = \sum_{i=1}^{m} \sum_{j=1}^{\alpha_i} \binom{\alpha_i}{j} = \sum_{i=1}^{m} 2^{\alpha_i} - m.$$

This follows from the fact that we can 'run' on any combination of $j \geq 1$ of the α_i coordinates having value i.

For the case of s odd we add the center value 0 and let $s = 2m + 1$, β_0 is the number of zeros among the cell coordinates, $\underset{\sim}{\beta} = (\beta_0, \beta_{-1}, \beta_1, \ldots, \beta_{-m}, \beta_m)$ and $\beta_0 + \beta_{-1} + \ldots + \beta_m = n$. Zero coordinates contributed in a different manner to the total number of runs. For fixed $\beta_i (i > 0)$ the β_0 zeros contributed a number of runs equal to

$$(6) \qquad \sum_{j=1}^{\beta_0} \binom{\beta_0}{j} 2^{j-1} = \tfrac{1}{2}(3^{\beta_0} - 1).$$

The different powers of 2 in the left member of (6) arise from the symmetries that occur when $2 \leq j \leq \beta_0$; d-diagonals get multiplied by 2^{d-1}, which is the total number of d-diagonals through a cell which has d coordinates equal to zero.

Hence for odd $s = 2m + 1$ the total number of runs through any cell with structure $\underset{\sim}{\beta}$ is

$$(7) \qquad r(\underset{\sim}{\beta}) = \sum_{j=1}^{\beta_0} \binom{\beta_0}{j} 2^{j-1} + \sum_{i=1}^{m} \sum_{j=1}^{\alpha_i} \binom{\alpha_i}{j} = \tfrac{1}{2}\{3^{\beta_0} - 1 - 2m\} + \sum_{i=1}^{m} 2^{\alpha_i},$$

where $\alpha_i = \beta_i + \beta_{-i}$ $(i = 1, 2, \ldots, m)$.

Since we are only interested in the presence or absence of equalities among the coordinates and in the number of zeros, there are only a finite number of possible multiplicities for any cell if n and s are fixed.

$\underline{Q}_3$: For $n = 3$ what are the various possible multiplicities?

With no zeros present we can have all different, one pair or a triple. With one zero we can have the other two different or forming a pair. We can also have exactly two zeros or all three equal to zero. For the 3 cases with $\beta_0 = 0$ we obtain from (5) the results 3, 4 and 7, respectively. With one zero we obtain from (7) the results 3 and 4, respectively. For $\beta_0 = 2$ and 3 we obtain from (7) the results 5 and 13, respectively.

In fact we can now make some general statements about the distribution of these multiplicities in the cube $H^3(s)$ for odd and even s. Counting the number of cells for each of the above 7 types, we put the results in the forms of a table

Multiplicity	Odd s		Even s	
	Number of Cells	Product	Number of Cells	Product
3	(s-1)(s-2)(s-3)	3(s-1)(s-2)(s-3)	s(s-2)(s-4)	3s(s-2)(s-4)
4	6(s-1)(s-2)	24(s-1)(s-2)	6s(s-2)	24s(s-2)
5	3(s-1)	15(s-1)	0	0
7	4(s-1)	28(s-1)	4s	28s
13	1	13	0	0
Total	s^3	$s(3s^2+6s+4)$	s^3	$s(3s^2+6s+4)$

This also checks with (1) since the sum of the multiplicities should be s times the total number of runs.

More generally the total number of cells of the form $0^{\alpha_0}, 1^{\alpha_1}, \ldots, m^{\alpha_m}$ for even s and odd s, respectively, is given by

$$(8) \qquad 2^n \begin{bmatrix} n \\ \alpha_1, \alpha_2, \ldots, \alpha_m \end{bmatrix}, \quad \begin{bmatrix} n \\ \beta_0, \beta_{-1}, \beta_1, \ldots, \beta_{-m}, \beta_m \end{bmatrix},$$

where the square bracket denotes the multinomial coefficient (we write these as $[{n \atop \underset{\sim}{\alpha}}]$ and $[{n \atop \underset{\sim}{\beta}}]$, respectively, below) and the power of two arises from the $\pm$ sign on each coordinate. Another proof of (3) is now available since it can easily be shown (proof omitted) that for even s

$$(9) \qquad \frac{2^n}{s} \sum_{\underset{\sim}{\alpha}} [{n \atop \underset{\sim}{\alpha}}] \left\{ \sum_{i=1}^{m} 2^{\alpha_i} - m \right\} = \frac{(s+2)^n - s^n}{2},$$

where the sum is over all m-tuples $\underset{\sim}{\alpha}$ with $0 \leq \alpha_i \leq m$ and $\Sigma\alpha_i = n$. In (9) we added the runs in all cells and divided by s to get the total number of runs. A corresponding result with the same right hand side holds for odd s.

$\underline{Q_4}$: What is the expected value of the number of runs through a randomly selected cell?

If all the cells have equal probability $1/s^n$ of being chosen and the number of type $1^{\alpha_1}, 2^{\alpha_2}, \ldots, m^{\alpha_m}$ is given by (8), then our result for even s is available from (9) by replacing s in the denominator by s^n, i.e., for even $s > 0$

$$(10) \qquad E\{r(\underset{\sim}{\alpha})\} = \frac{(s+2)^n - s^n}{2s^{n-1}} \geq n$$

and the same result also holds for odd s. For $n = 3$, $s = 4$ we obtain $E\{r(\underset{\sim}{\alpha})\} = 4.75$.

For large n we can improve the inequality in (10) since a simple lower bound for (3) is

(11) $R_1(n, s) \geq n(s + 1)^{n-1}$.

This follows from the fact that (10) is a difference quotient of a convex function (for $s \geq 0$) and hence the derivative at the centerof the interval $(s, s+2)$ must lie beneath the curve for all $s \geq 0$. For $n = 3$, $s = 4$ the correct answer is 76 and the lower bound is 75. Hence we obtain for (10) the lower bound $n(s + 1)^{n-1}/s^{n-1}$.

Q_5: What is the variance of the number of runs through a randomly selected cell?

For even $s \geq 2$ we use the notation above to write for the second moment of $r(\underset{\sim}{\alpha})$

(12)
$$
\begin{aligned}
E\{r^2(\underset{\sim}{\alpha})\} &= \frac{2^n}{s^n} \sum_{\underset{\sim}{\alpha}} \left[{n \atop \underset{\sim}{\alpha}}\right] \left\{\sum_{i=1}^{m} 2^{\alpha_i} - m\right\}^2 \\
&= \left(\frac{1}{m}\right)^n \left\{m \sum_{\underset{\sim}{\alpha}} \left[{n \atop \underset{\sim}{\alpha}}\right] 4^{\alpha_1} + m(m-1) \sum_{\underset{\sim}{\alpha}} \left[{n \atop \underset{\sim}{\alpha}}\right] 2^{\alpha_1+\alpha_2} - 2m^2 \sum_{\underset{\sim}{\alpha}} \left[{n \atop \underset{\sim}{\alpha}}\right] 2^{\alpha_1} + m^{n+2}\right\} \\
&= \frac{(m+3)^n + (m-1)(m+2)^n - 2m(m+1)^n + m^{n+1}}{m^{n-1}} \\
&= \frac{2(s+6)^n + (s-2)(s+4)^n - 2s(s+2)^n + s^{n+1}}{4s^{n-1}} .
\end{aligned}
$$

Hence the variance $\sigma^2 = Er^2(\underset{\sim}{\alpha}) - [E\ r(\underset{\sim}{\alpha})]^2$ is given after algebraic simplification by

(13)
$$
\sigma^2 = \frac{2s^{n-1}(s+6)^n + (s-2)s^{n-1}(s+4)^n - (s+2)^{2n}}{4s^{2n-2}} ;
$$

this gives 0 for $s = 2$ and any fixed n which is correct since every cell is then a corner cell and has the same $r(\underset{\sim}{\alpha}) = 2^n - 1$.

For odd s a similar derivation gives for the second moment and variance, respectively,

(14) $$E\{r^2(\underset{\sim}{\beta})\} = \frac{1}{s^n}\sum_{\underset{\sim}{\beta}} [{}^n_{\underset{\sim}{\beta}}] \left(\sum_{i=1}^{m} 2^{\alpha_i} + \frac{3^{\beta_0} - 1 - 2m}{2}\right)^2$$

$$= \frac{1}{4s^n}\left\{(2m+9)^n + 4m(2m+7)^n + 4m^2(2m+5)^n + 2(2m+1)^2(2m+3)^n + (2m+1)^{n+2}\right\}$$

$$= \frac{(s+8)^n + 2(s-1)(s+6)^n + (s-1)^2(s+4)^n + 2s^2(s+2)^n + s^{n+2}}{4s^n} ,$$

(15) $$\sigma^2 = \frac{s^{n-2}(s+8)^n + 2s^{n-2}(s-1)(s+6)^n + s^{n-2}(s-1)^2(s+4)^n - (s+2)^{2n}}{4s^{2n-2}}$$

The latter is zero for $s = 1$ and any fixed n.

$\underset{\sim}{Q}_6$: What is the asymptotic $(s \to \infty)$ structure of the mean and variance of $r(\underset{\sim}{\alpha})$ and $r(\underset{\sim}{\beta})$?

From (10) for both odd and even s we have for $s \to \infty$

(16) $$E\{r(\alpha)\} \sim n.$$

In (13) and (15) the coefficients of s^{2n}, s^{2n-1} and s^{2n-2} all vanish and we obtain the same asymptotic result for both odd and even s, namely

(17) $$\sigma^2 \sim \frac{n(n-1)}{s} .$$

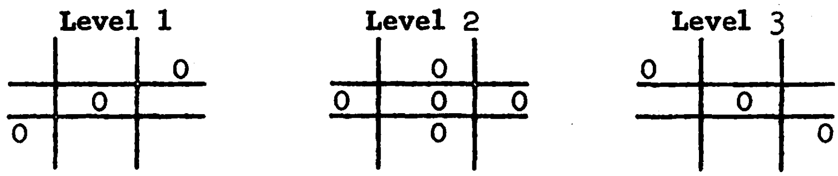
Thus we have not only confirmed the fact that $\sigma^2 \to 0$ as $s \to \infty$ for any fixed n but found the rate of approach.

$\underset{\sim}{Q}_7$: What is the maximum number of cells M that can be occupied without having a single completed run?

This question is equivalent to asking for the minimum number of 'holes' (or unoccupied cells) such that every run contains at least one hole. This in turn is equivalent (except that the answer is now $s^n - M$) to asking for the minimum number of cells that can be occupied such that every run contains at least one occupied cell.

For $n = 3$, $s = 3$ we have a simple construction

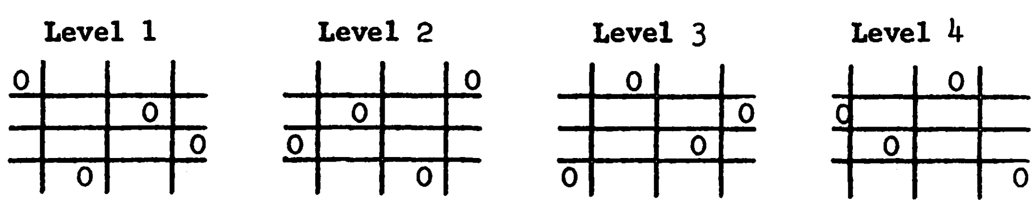
(18)



which requires 11 occupancies so that the original problem has 16 occupancies out of 27, with 6 on the first and third levels and 4 on the center level. To improve on this we would need 3 level designs with 6, 6 and 5 occupancies (since we can put at most 6 on each level). The only way to occupy 6 cells on any level is to leave a diagonal empty. If we leave the same (i.e., corresponding) diagonal empty on 2 levels then we have to leave at least 6 empty cells on the remaining level. If we put the two levels (with 6 occupancies each) adjacent to each other then the remaining level can contain at most 2 more occupancies; if they are separated as in (18) then we can have 4 more occupancies. This proves that 16 is optimal for $n = 3$, $s = 3$.

For $n = 3$, $s = 4$ we use the design

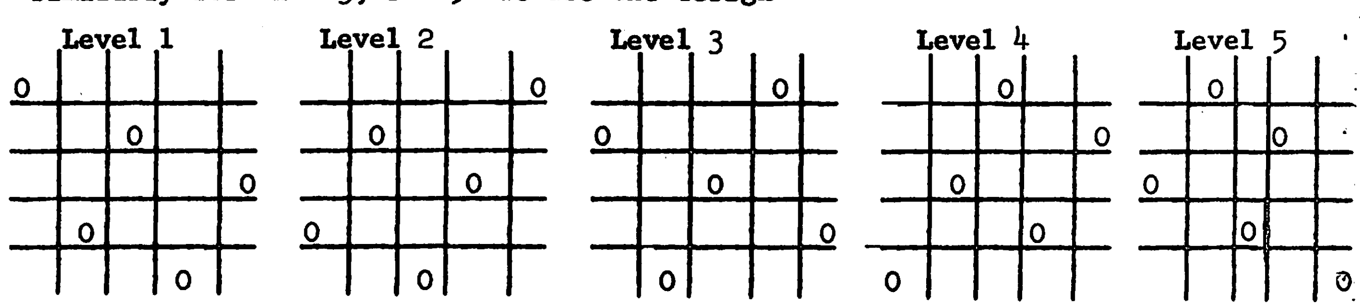
(19)



which has a hole in all runs except for three of the four 3-diagonals. Thus 19 holes will put at least one hole in each run and hence we can occupy $64 - 19 = 45$ cells without having a single completed run.

Similarly for $n = 3$, $s = 5$ we use the design

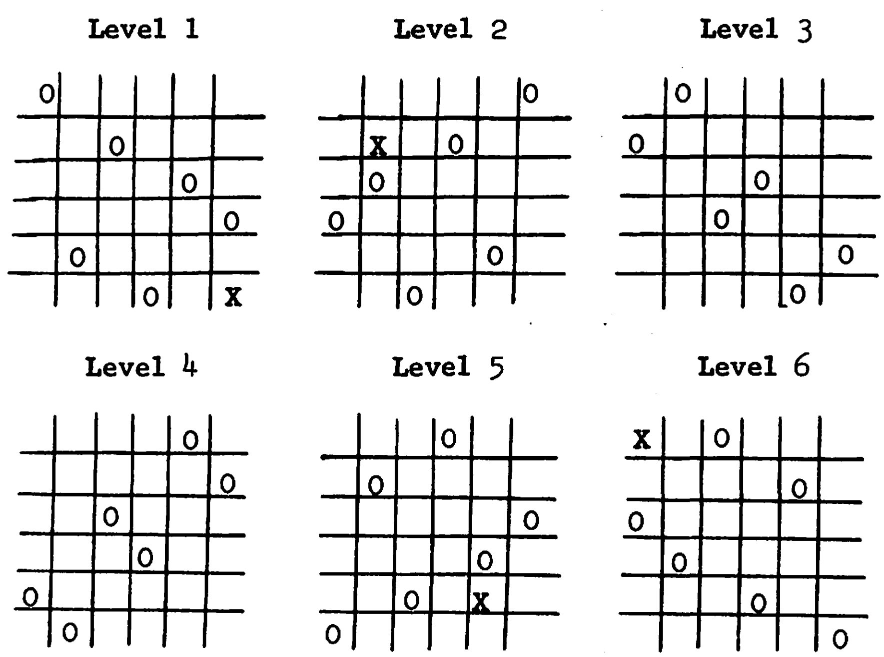
(20)



which has only 4 disjoint diagonals without holes. (They all start on the left in level 5 and end on the right in level 1.) Thus 29 holes will put at least one hole in each run and hence we can occupy 125 - 29 = 96 cells without having a single completed run.

We conjecture that these results for $n = 3$, $s = 4$ and 5 are optimal but have not proved this. It is somewhat curious that the results for $n = 3$, $s = 3$, 4 and 5 satisfy the formula $(s - 1)(s^2 - 1)$ but the correct answer for $s = 2$ is 1 and this formula gives 3. In addition for $s = 6$ the following scheme with 36 circles

(21)



has only eight 2-diagonals without holes. Since these diagonals overlap in pairs we need only put in 4 more circles in the four positions marked with an X in (21). Thus 40 holes is sufficient to put at least one hole in each run and hence we can occupy 216 - 40 = 176 cells without having a single completed run. For $s = 6$ the value of $(s - 1)(s^2 - 1) = 175$.

It should be pointed out that this problem Q_7 can be regarded as an integer programming problem; suppose for example we consider $n = 3$, $s = 4$.

For each cell (i, j, k) we set $X_{ijk} = 1$ if it is a hole and 0 if it is occupied. The condition that no runs are complete becomes a set of 76 inequalities

(22) $\Sigma X_{ijk} \geq 1$ (sum is over 4 cells forming a run).

Under these restrictions we want to

(23) Minimize ΣX_{ijk} (sum is over all cells).

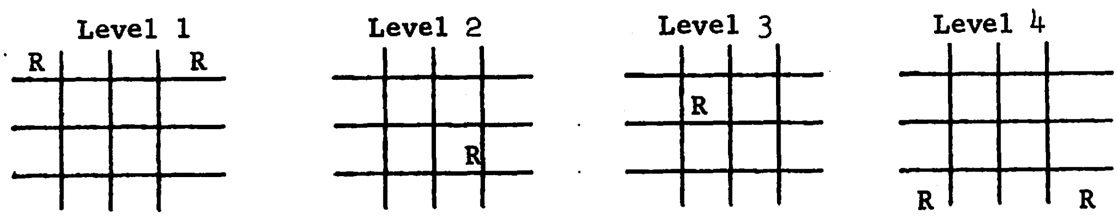
If we replace each X_{ijk} by $Y_{ijk} = 1 - X_{ijk}$ then we obtain the dual problem: find the smallest number of cells that can be occupied such that every run contains at least one occupied cell. We now add a few related questions which we had hoped to answer but have not yet done so.

$\underline{Q_8}$: If the cells of the cube $H^n(s)$ are occupied sequentially and independently with uniform probability over the unoccupied cells, what is the expected number of cells occupied when the first run appears? In particular, for $n = 3$ and $s = 4$ is this expectation smaller or larger than the average of the upper and lower bound, which the answer to, by $\underline{Q_7}$, is $(46 + 4)/2 = 25$?

$\underline{Q_9}$: As in $\underline{Q_8}$ what is the expected number of cells occupied when (for the first time) we have at least one occupancy in every run. In particular, for $s = 4$ is this expectation smaller or larger than the average of the upper and lower bound which, by the answer to $\underline{Q_7}$, is $(61 + 19)/2 = 40$?

$\underline{Q_{10}}$: If an occupant of any cell (which we shall call a Runner) moves, it goes only along runs of length s (almost like a Queen in the game of chess). The Runner is said to cover the cell he occupies and all the cells he can get to in one move. What is the minimal number of Runners required to cover the entire hypercube $H^n(s)$?

Clearly the result of $H^n(2)$ and $H^n(3)$ is 1 and for $H^2(s)$ the result appears to be $s - 2$ for all s (except for $s = 2$ and 4, in which case it is $s - 1$). For $H^3(4)$ the result appears to be 6 with one possible arrangement as



where R stands for Runner. It would be interesting to check the correctness of the following conjectured formula for the minimal number of Runners required to cover the cube $R^3(s)$, namely $1 + \{\frac{(s-3)(s+5)}{2}\}$, where $\{x\}$ is the smallest integer $\geq x$. The answers it gives for $s = 3, 4$, and 5, namely 1, 6 and 11, respectively appear to be correct. Even if it is not correct for all s, it would be interesting to check whether the leading term $s^2/2$ is correct asymptotically $(s \to \infty)$.

$\underline{Q}_{11}$: If we occupy cells at random as in $\underline{Q}_8$ what is the expected number of cells occupied by Runners when the hypercube is covered for the first time?

This last question $\underline{Q}_{11}$ is related to a rook covering problem treated by Kats and one of the authors [2].

<u>Acknowledgement</u>.

Thanks are due to Rammy Y. Uppuluri who interested the second author in the game of Qubic, corresponding to $n = 3$, $s = 4$. We would also like to thank Professor T. Ferguson of the Department of Mathematics, UCLA, for a stimulating conversation on this paper in which he pointed out that $\underline{Q}_7$ can be regarded as an integer programming problem. We also wish to thank Professor J. N. Srivastava of Colorado State for his communicated result that some of our answers for $\underline{Q}_7$ can be extended to cases where s is any power of a prime integer.

References

[1] Problem E 773 Proposed by A. L. Rubinoff and solved by L. Moser, Amer. Math. Monthly, Feb. 1948, p. 99.

[2] Katz, L. and Sobel, M. "Coverage by randomly placed rooks" (submitted to the VIth Berkeley Symposium).