ON SELECTING THE t = 2 BEST OF n ITEMS

٤

USING BINARY COMPARISONS

Ъy

Milton Sobel[§] Technical Report No. 114 University of Minnesota, Minneapolis, Minnesota and Imperial College, London, England

Supported partially by a National Institutes of Health Special Fellowship while at Imperial College, London and partially by National Science Foundation Grant Number Gp-9018 through the University of Minnesota.

1. Introduction.

- **-**

In a recent paper [5] the author has considered the problem of ranking the t = 2 best (i.e., the largest two) of n unequal numbers when only binary errorless comparisons are made; this paper considers the analogous problem of selecting the t = 2 best without ordering them. We are interested in two criteria: one is to minimize the expected number of comparisons required (called the E-criterion) and the other is to minimize the maximum number of comparisons required (called the M-criterion). Unlike the ranking problem which was considered by different authors, the selection problem appears not to have been previously considered; hence all the procedures discussed are new. The ideas behind some of the procedures and one method for obtaining an E-lower bound are similar to those used in [5]. The E- and M-efficiences of our procedures are numerically investigated.

In order to evaluate efficiency or prove optimality we need to develop an attainable lower bound over all possible procedures. The best M-lower bound is obtained; the E-lower bound obtained is only over a certain class of procedures. Our results (see table in Section 2) are optimal for $n \leq 5$. With the help of the above M-lower bound, one of the procedures R_{M} is shown to be M-optimal for all n.

This formulation is directly applicable to tournament problems and we use the associated terminology, i.e., the best player corresponds to the largest number, etc. The better player always wins and, since no two players have the same ability, a draw cannot occur.

2. Procedures for the Selection Problem for t = 2.

Six procedures for the selection problem with t = 2 are defined. Three of them use the concepts of pairing and one-step or higher-step expected

- 1 -

T. LEGGIN ... OF

The absorbingtor [p] the subject his construct the problem of vaniang the absorbingtor [p] the subjective) of a unequal number, when only bluely exactlices expansions are made; this paper considers the analogous problem of selecting the t = 4 best without optaring there. We are interested in two criteries and is continuated the exterior of comparisons required (called the M-subjection) and the other is to similar the restinue unoder of comparisons required (called the M-exiterion). Unlike the remining problem which was considered by sinfferent authors, the collection problem appears not to have seen recylously considered; hence all the procedures for obtaining an K-lower out out are similar to fine procedures and are similar for obtaining an K-lower out out are similar to fine procedures and the H- even for obtaining an K-lower out procedures needs on the procedures of for the field of the field

In order to evaluate structurer or prove ordered. The bard M-lower is obtained for the bound of the transformed of the formation of the formatio

This formulation is directly applicable to four density oblication and we use the associated terminology, i.e., the best slaper corresponds to the largeon number, sid. The better player slape wine and, since so the flater the same stillity, a draw carbot opens.

7: Proceduret for the Salection Repolem for S = 2:

Sin puradduren for the nelection problem with the star daring .

Three of the not the concepts of putring and one-step or higher-step expected

entropy (the adjective expected is later deleted) and one is related to a procedure suggested by Picard [2] for the corresponding ordering problem. After some preliminary definitions, we briefly describe the procedures and give a table comparing the numerical results obtained for n = 1(1)10.

A state of nature (or case) is any one of the n! possible ordered arrangements of the n players. There are, of course, $\binom{n}{2}$ possible decisions. At any stage of the procedure, we are concerned with the number of cases that are consistent with the results of comparisons already made for each of the $\binom{n}{2}$ decisions; these $J = \binom{n}{2}$ integers are proportional to the conditional probabilities that each of the $\binom{n}{2}$ decisions is the correct one given the results of the comparisons already made. Hence these J integers describe the state of our system, say S_{α} . Let the integers be $n_{i}^{(\alpha)}$ with sum $N^{(\alpha)}$ and let $p_{i}^{(\alpha)} = n_{i}^{(\alpha)}/N^{(\alpha)}$ (i = 1, 2,..., J) be the conditional probabilities given the system state S_{α} . The entropy (or uncertainty) associated with S_{α} is given by

(2.1)
$$\sum_{\alpha} (s_{\alpha}) = -\sum_{i=1}^{J} p_{i}^{(\alpha)} \log p_{i}^{(\alpha)}$$

v ₹ 1

all logs in this paper are to the base 2 unless stated otherwise. If we start from S_{α} and a comparison $C = C(a \lor b)$ (where \lor means versus) leads to states S_1 (resp., S_2) with probabilities $q_1^{(\alpha)}$ (resp., $q_2^{(\alpha)} = 1 - q_1^{(\alpha)}$), then the expected one-step reduction in entropy due to the comparison C, applied to the system state S_{α} , is given by

(2.2)
$$E(\Delta \mathcal{E}|c, s_{\alpha}) = \mathcal{E}(s_{\alpha}) - \{q_1^{(\alpha)}\mathcal{E}(s_1) + q_2^{(\alpha)}\mathcal{E}(s_2)\}.$$

If we look s steps ahead then the expected reduction is again given by $\mathcal{E}(S_{\alpha})$ minus the appropriate average of (at most) 2^S uncertainties. Our basic idea is to fix an s and find the comparison that maximizes the expected

- 2 -

entropy (the adjective expected is later delated) and the fourth is related to a probably subjected by Piceric [2] for the entropoid η_{ij} ordering problem Mitter some preliminary definitions, we objectly desorry the form probably desired for n = 1(1).

A state of matter (or cased) is any one of the mit possible ordered arrangements of the mit players. There are of course, $\binom{n}{2}$, possible decision. There are of course, $\binom{n}{2}$, possible decision. There are conserved with the matter, if exactions are consistent with the procedure, we are conserved with the matter, if exacting a consistence with the procedure, we are conserved with the matter, if exacting a consistence with the procedure, we are conserved with the matter, if exacting a consistence with the procedure, we are conserved with the matter, if exacting a consistence with the procedure, we are conserved with the matter, if exacting a constraint the procedure, we are conserved with the matter of the constraint of the comparisons already with, fuere $(\frac{n}{2})$ integers are proportional to the locational proportion of the comparisons already with, fuere $(1 - \binom{n}{2})$ integers describe the sector of $\binom{n}{2}$ and $\binom{n}{2}$ and $\binom{n}{2}$ integers are proportional to the indication of the off the constraint of the constraint of the indication of the constraint of the constraint of the last $\binom{n}{2}$ into an $\binom{n}{2}$ integers is indicated by the sector of the system state is the order of the system state $\binom{n}{2}$ into an $\binom{n}{2}$ is the system of the system state $\binom{n}{2}$ for a state of $\binom{n}{2}$ is integers of the system state $\binom{n}{2}$ into an arbor. The state $\binom{n}{2}$ is integers by the system state $\binom{n}{2}$ into a state $\binom{n}{2}$ is integers by the system of the system of the system by the system of the system of the system state $\binom{n}{2}$ is integers of the system of the system by the system of the system by the system of the system by the system state $\binom{n}{2}$ is the system by the system state $\binom{n}{2}$ integers of the system of the system of the system by the system

will lows in this payer are to the base 2 values stated of states. It we start from S_{j} and a comparison $C_{j} = \tilde{C}(a, \gamma)$ (where γ means verses) leads to states S_{j} (resp., S_{j}) with proceedingless $\binom{G}{1}$ (resp., $\binom{G}{2} = 1 - \binom{G}{1}$) then the spected one-step rollation on entropy due to the comparison G_{j} splited to the system state S_{j} is given by

$$(\mathbf{s} \cdot \mathbf{s}) \qquad \ast (\mathbf{s} \cdot \mathbf{c}, \mathbf{c}) \stackrel{\boldsymbol{s}}{=} \mathbb{S}(\mathbf{s}^{2}) - \mathbf{s} \stackrel{\boldsymbol{s}}{\leftarrow} (\mathbf{s}^{2}) + \stackrel{\boldsymbol{s}}{\leftarrow} \mathbb{S}(\mathbf{s}^{2}).$$

Filea (a) tourtise, an (a which find the Obstation that no clarges the e cleater

s-step reduction in entropy at each stage. When the one-step plan (our procedure R_E) does not give optimal results, we investigate the improvement of a two-step plan by allowing the use of two-step reduction in a non-systematic manner (see the procedure R_E * below). It is conjectured that a systematic two-step plan would do at least as well, but this has not been proved.

The use of the expected reduction in entropy as a tool for search problems was used by Sobel and Groll [6] for group-testing, by F. Dubail [1] for other search problems and also by the present author in [5].

Another point of interest is the distinction between cycle pairing and complete pairing. For any n, let the binary structure of n be

(2.3)
$$n = 2^{r_1} + 2^{r_2} + \ldots + 2^{r_s} (r_1 > r_2 > \ldots > r_s \ge 0),$$

. ۲

> so that s is the number of ones in the binary notation for n. Let p be the highest power of 2 that factors into n!. Then it is easy to prove (see [5]) that p = n - s.

For $n = 2^{r}$ a knock-out tournament for finding the best one consists of r rounds where the number of contenders is halved at each round. Under complete pairing we start a procedure by randomly breaking up n into subsets of size $2^{r_{i}}$ as in (2.3) and doing a knock-out tournament within each of these subsets. After this, we use the comparison that maximizes the one=step expected reduction in entropy.

Under cycle pairing we start with a knock-out tournament only for one subset of size 2^{s} (usually $s = r_{1}$ defined in (2.3)) and then continue with the exp§cted reduction in entropy. The procedure R_{E}^{*} uses cycle pairing for n - 1 i.e., it uses cycle pairing with $s = r_{1}$ for $n \neq 2^{r}$ and for $n = 2^{r}$ we take s = r-1.

- 3 -

s.stepredoction in entropy at each stage. Since the one-step plan (onversion of the optimal results, we investigate the optimal results, we investigate the optimal scales of threaded of a two-step plan by allow. the epsilos of the bits the filles reduction in a non-opsicate to manner (see the procedure, R.*, below). It is the investigate to optimal test of the process of the bits the interval of the process of the bits the interval of the process of the proc

The use of the apported reduction in entropy as a tool for notion problems was used by Sobel and Croll (5] for properturbance, by F. Onbarl 1, Sor other search problems and clao (7 [1], prepent rethor in [5].

Anotras polition interest is the distinction between cycle and nomplete milete pairs of in the dominant structure of a be

 $(\mathbb{S}^*\overline{\mathbb{S}}) \quad \forall = \mathbb{S}_{\mathbb{T}^{\overline{1}}} + \mathbb{S}_{\mathbb{T}^{\overline{N}}} + \cdots + \mathbb{S}_{\mathbb{T}^{\overline{N}}} \quad (\mathbb{T}^{\overline{1}} > \mathbb{T}^{\overline{N}} > \cdots = \mathbb{S}^{\overline{n}} > 0)^{+}$

so that s is the number of ones in the dimary hotalion for u. Let i we the unimest power of 2 that factors set up. Then it is can' to prove (see [3]) to at p = n - s.

For $n = 2^{n}$ where the number of contendary is helfed at each road. Under of r rounds where the number of contendary is helfed at each road. Under complete pairing we start a probability dual p realized in into subjects of size 2^{n} as is $(2, \cdot)$ and do ng a knock-out to realize within each of these subjects. After this, we use the comparison that mathematic the integral expected reduction in a crowr.

Under cycle parr m) we start write a knock-out formanant only for one subset of size 2^5 (usually $s = c_1$ defined in (2.3)) and then contractly the orggright reduction in antropy. The properties 2, * use cycle value for for n - 1 i.e. it fiss cycle value vite $i = r_1^2$ for $n = c_1^2$ and for $n = 2^7$ we take $s = y_{-1}$. <u>Procedure R</u>_A: Let $n = 2^{r} + c$ with $0 < c \le 2^{r}$. This procedure uses cycle pairing with 2^{s} players with some $s \le r$ (actually s = r-1 or r, the particular value to be determined later). After this we have one large connected set of size 2^{s} for the graph (or tree) and the remaining $n - 2^{s}$ players (called newcomers below) are unconnected, i.e., have not yet played. The connected set has a best player x_1 with i_1 inferiors and among these a contender for second best x_2 with the largest number of inferiors, say i_2 among the contenders for second best. Then $i_2 \le i_1 - 1 \le n - 2$ and our goal is to make $i_2 = n - 2$, which implies that $i_1 = n - 1$.

3 1

> Each newcomer except for the last one (and possibly the one before that as explained in the ending E_1 below), comes up in turn (we assume they are in order) and plays x_2 . If he loses he retires; if he wins he plays again, this time against x_1 . If he loses to x_1 he takes over as the new x_2 ; if he wins against x_1 he becomes the new x_1 and the old x_1 becomes the new x_2 .

Three different endings are used with this procedure, say E_1 , E_2 and E_3 , according to whether the original x_1 (after cycle pairing) is beaten by a newcomer not among the lasttwo, he is beaten for the first time by the next-to-last newcomer n_1 , or he is better than all the newcomers except possibly the last one n_0 .

Under E_1 there is only 1 contender for second best (in the connected subset), namely x_2 , when n_1 is ready to play. Again n_1 plays x_2 and retires if he loses, but if he wins he 'sits out' a game letting n_0 play x_2 . Then if n_0 loses, he(i.e., n_0) retires and we are through. If n_0 wins, then x_2 retires and we need exactly 2 more games to find the best 2 of the 3 players, x_1 , n_0 and n_1 .

- 4 -

Probedies R : Let $n = 2^{+} + c$ with $0 + c - 2^{+}$. This procedure is solved, the initial set $n = 2^{+} + c$ with $0 + c - 2^{+}$. This procedure is a cycle, particular to be determined later). After this we have do not a connected set of size 2^{3} for the staps (of the) and the remining $n = 2^{2}$ players (balled naveomers below) are momented into $n = 2^{2}$ players connected set has a best player z_{1} with n inderious and money these a contender for second peak z_{2} with n = 1 and one data is to make onther z_{2} which implies that $1_{2} = 1_{1} + 1_{2} + 1_{2} + 1_{3}$ is to make $1_{2} = n + 2_{1}$ which implies that $1_{2} = n + 1$.

Back news not module for the lest one (and possibly the ordebed that as amplained on the ondone π_1 below) codes of the the (we assume they the in order) and they π_2 . If he boses is set lest in order of a matrix this time against π_1 , if he loses to π_1 the codes over so the metric π_1 . It is the against π_1 , if he loses to π_1 the codes over so the metric π_1 .

Three different and higs are sed with this procedure only z_1 , $\overline{z_1}$, and $\overline{z_1}$, according to whather the original $\overline{z_2}$ (after cycle partial) is reacted by a newcomer not among the histotroches beater for the first time of the added to fast nonconer $\overline{z_1}$, or he is fact the all the newconers chapt possioly the last one $\overline{z_1}$.

Under u_{1} there is only 1 contended for some base (in the connected subact), numery u_{2} when u_{1} is ready to day. dain u_{1} plays u_{2} and retries it he loses, but if he clus he side out is name latting u_{1} day u_{2} . u_{2} . Then if u_{0} . $U((-, u_{3}))$ retries and we are thrown. If u_{0} when let u_{2} we need we could be date to find the bast 2 of the players u_{1} , u_{0} and u_{1} .

- ji.

Under E_2 there is only 1 contender for second best (in the connected subset), namely x_2 , when n_0 is ready to play (here n_1 plays as usual). Then n_0 plays x_2 and the procedure terminates.

V

Under E_3 there are s contenders (including x_2) for second best (in the connected subset) when n_0 is ready to play. Again n_0 plays x_2 . If he loses, he retires and we need exactly s - 1 more games to find the 2 best players. If he wins, he continues to play the other s - 1 contenders, eliminating one at each step, until he loses (at which point he retires) or he wins over all of them. If he loses and there are still c contenders then we need exactly c - 1 more games to complete the procedure.

To determine the value for s for the initial cycle pairing we use the exact formula for the expectation derived in Section 3 for procedure R_A and use the s that gives the smaller expectation. Thus for $5 \le n \le 10$ it can be verified that s = 2, or cyle pairing with $2^S = 4$ units, is best and for n = 11 we start cycle pairing with 8 units (see equation (3.20) for the asymptotic equivalent).

<u>Procedure R_E^* </u>: Start with cycle pairing for n - 1 and then maximize the 1 step (or 2 step) expected reduction in entropy. Two-step reductions are sometimes used but not systematically.

<u>Procedure R_E </u>: This consists of a 'pure' strategy for maximizing the one-step expected reduction in entropy. Here two-step reductions are used only when different comparisons give the same entropy reduction in 1 step.

<u>Procedure R_{CP} </u>: Start with complete pairing (for n) and then continue with the strategy of maximizing the one-step expected reduction in entropy.

- 5 -

<u>Procedure R_M </u>: For this procedure we use 'complete pairing for n - 1' and for $n \ge 3$ make use of the binary expansion of n - 1.

(2.4)
$$n - 1 = 2^{r_1} + 2^{r_2} + \ldots + 2^{r_s} \quad (r_1 > r_2 > \ldots > r_s \ge 0)$$

to explain the different steps in the procedure, after putting one unit (or player) aside until the very last comparison.

1. Find the best one separately in each of the subsets for which $r_i > 0$.

2. Play the best one of the smallest subset of size 2^{r_s} against the best of the second smallest, the best of these two against the best of the third smallest subset, etc., until the best one of n - 1 is determined. Let c denote the number of contenders obtained for second best. 3. Use any knock-out tournament with exactly c - 1 games to determine the second best of n - 1.

4. Play the second best against the one set aside to complete the procedure.

<u>Procedure R</u>: Let the players in random order be denoted by 1, 2,..., n; we describe the scheme in three steps:

1. Play 1 v 2 and assume 1 loses to 2.

2. Play 3 v 1. If 3 loses then he is removed from contention; if 3 v 1 wins then 1 is removed and we play 3 v 2 to reestablish an ordering between the top two contenders.

3. Repeat this procedure, except that if the last player n wins then the extra game to reestablish the order is not played.

Although the procedure R_p appears to be inadmissible in the sense that one of the other procedures is at least as good or better for every n (for

- 6 -

Procedine R.: For this procedure we use [complete proving for a - 1" and for a > 1 main use of the binary argument of a - 1.

 $(2, 3) \qquad n = 2^{\frac{1}{2}} + 2^{2} + 2^{2} + \dots + 2^{\frac{1}{2}} \quad (x_{\frac{1}{2}} + x_{\frac{1}{2}} + x_{\frac{1}{2}} = 0)$

to explain the difference steps in the procedure, after posting one as t (or player) astie and i the wery last comparison.

I. Find the work one separately interval of the subsets for the charge $r_{i} > 0$.

2. Flay the Dest one of the soullest orbeet of size 2"s against the part of the second smallest, the bast of these the hast of

the third spallest jubict; dts., mith the base one of m - 1 is detarmined. Let d denote the ranker of contenders for deapart bestond best.

). Use any modulotic tormanant with sumptime 2 - 1 games to determine the success base of 1 , 1 ,

4. The record best clarest the one set state to complete the procedure.

Frocadore M.: Let the playare in realise order is denoged by 1, 2.... ut na desories the robers in these orders:

1. Play 1 V 2 and assays 1 Losds to 2.

. bucween che con the concenters.

2. Vig. 5.v.5. 37 % iosas than he is tasovad iros contention; "i v whus them i by resoved and we play the C its trastablish an order wi

A. Napeat this mroseduri, encopt that 'd the lest flager in which the test flager in which the the encoded the the order is not the which

one of the other procedures is at least as not better for when (for

both the E-problem and the M-problem), the simplicity of the procedure enables us to get specific formulas, which throws light on the asymptotic properties of the other procedures.

Numerical Results for Six Procedures for Selecting the t = 2 Best Players									
Using Only Binary Errorless Comparisons									
Bounds and Procedures	Expected Values E{T R}								
	n= 2	n= 3	n=4	n= 5	n= 6	n =7	n= 8	n= 9	n=10
LBE [#]		1.918	2.918	4.522	5.174	5.775	6.095	8.837	9.292
CLB ^{##}			3.386	4.685	5.899	7.066	8.197	9.307	10.401
RA		2	3 <u>4</u>	50	6 <u>8</u>	7309	8144	9121	11 <u>38</u>
R _E ∗		2	3 <u>4</u>	50	64	7 <u>315</u>	8 <u>149</u>	N.C.	N.C.
R _E		2	3 <u>5</u>	52	610	7 <u>550</u>	N.C.	N.C.	N.C.
R _{CP}		2	40	50	6 <u>10</u>	7 <u>550</u>	<u>90</u>	100	11 <u>336</u>
RM		2	3 <u>4</u>	5 <u>0</u>	6 <u>18</u>	7420	8180	10 ⁰	11 <u>840</u>
R _P		2	3 4	5-1	6 <u>17</u>	7 <u>567</u>	9 <u>39</u>	10 <u>61</u>	11829
D ^{§§}			6	6	30	630	210	140	1260
		Maximum Length M{T R}							
LBM [§]		2	4	5	7	8	9	10	12
RA		2	4	5	7	9	11	13	15
R _E ∗		2	4	5	7/	9	11	N.C.	N.C.
R _E		2	4	5	7	9	N.C.	N.C.	N.C.
R _{CP}		2	4	5	7	9	9	10	12
R _M		2	4	5	7	8	9	10	12
Rp		2	4	6	8	10	12	14	16

The LBE is the lower bound (3.14) for all procedures that use 'cycle pairing for n-1' defined in the text.

The CLB is the conjectured lower bound (3.22) for all procedures. Since the LBE-values are smaller, they are also conjectured to hold for all procedures.

§ The LBM is the best lower bound for the maximum branch length given on the right sides of (3.23) and (3.25).

 $\S\S$ Each D is the common denominator for all the underlined numerators above it.

3. Properties and Bounds.

· · ·

Since procedure R_p gives easy results, we consider it first. Let $f_p(n) = E\{T | R_p\}$ denote the expected number of comparisons and let $\overline{f}_p(n)$ denote the maximum length under R_p . It is easy to see that the expected number of games for the jth player $(3 \le j \le n - 1)$ is $1 + \frac{2}{j}$, while for j = n, only one game is played. It follows that for $n \ge 3$

(3.1)
$$f_p(n) = n - 1 + \sum_{j=3}^{n-1} 2/j = n - 4 + 2 \sum_{j=1}^{n-1} 1/j \approx n + 2 \ln n.$$

Clearly if players 3, 4,..., n-1 all win we obtain the maximum length; hence for $n \ge 3$

(3.2)
$$\overline{f}_{p}(n) = 2n - 4.$$

Although R_p has a better expectation than R_E for small powers of 2 (see n = 4), the maximum length grows very rapidly compared to that of R_E .

Let $f_M(n) = E\{T | R_M\}$ denote the expected number of comparisons under R_M and let $\overline{f}_M(n)$ denote the maximum number under R_M . Using the result for R_M in [5] for the ordering problem with n replaced by n - 1, and adding 1, gives

(3.3)
$$\overline{f}_{M}(n) = n - 1 + [\log (n - 2)]$$

and

(3.4)
$$f_{M}(n) = n - 2 + \frac{1}{n-1} \sum_{j=1}^{s} (r_{j} + j - \delta_{js})^{2^{r_{j}}}$$

where the r_j are now defined by (2.4) and $\delta_{js} = 1$ if j = s and = 0 otherwise. The asymptotic analysis in [5] is also appropriate here. It is shown below that (3.3) is the smallest possible maximum branch length and hence the procedure R_M is M-optimal.

- 8 -

Since proceedings \hat{x}_{p} , three eacy require, to construct it littles. Let $\hat{x}_{p}(n) = \mathbb{E}[\hat{x}_{p}] \hat{x}_{p}$, denote the superior of number of connertence and let $\tilde{x}_{p}(n)$ for the neutrinom largely under N_{p} . It is easy to see that the angle of each index of neutrinom largely index N_{p} . It is easy to see that the angle of other index of neutrinom for the $j^{(n)}$ player $(3 \leq j \leq n-1)$ for $1 + \frac{1}{j}$, while for j = n, only one same if played. It follows that for $n \geq 3$

Closely if ployers 3, 4,000, s-1 cll win we colean the reminim length; Sonce for the 2 3

$$(z,z) = (z,z) + i$$

Although P_{Γ} less a wither remodention there λ_{B} for the Lipbour of f (see in = 1), the mariner length press vary repitely compared to that of R_{σ} .

Let $\mathcal{X}_{H}(n) = \mathbb{E}\{\widehat{\gamma}_{H}(n)$, denote the expected wither of concertsons under \mathcal{X}_{H} and let $\widehat{\mathcal{Z}}_{H}(n)$, denote the continue number width \mathcal{X}_{n} . Using the result for $\widehat{\mathcal{Z}}_{H}$ in [5] for the predict, where with a solution of the $\widehat{\mathcal{Y}}_{n}$, the loss $\widehat{\mathcal{Y}}_{n}$ and define $\mathbb{E}_{\mathcal{X}}$ from

$$(z,z) = u + z + [zoz (u - z)]$$

in a start in the start in the

$$(x,y) = \mu - 2 + \frac{\mu - 1}{2} + (2 + 1 - 2^{2})$$

where, the τ_j function defined by (\cdot, ϕ) put $\phi_{j,g} = 1$ if j = n and = 0otherwise. The service indicates in [5] is also actuated here. It is shown balow that (3, 3) is the archiver notable maximum branch lungth and

gladae dha axoochare (g^{al}) dal gebruare (

Since we have a special interest in procedures that use cycle pairing for n - 1 we let R_{c} denote any such procedure and study its properties.

For $n = 2^{r} + 1$ the cycle-pairing and complete pairing are the same and after n - 2 comparisons we have 1 player (say n) that has never played and for the remaining 2^{r} we have a best player (say $n - 1 = 2^{r}$) and exactly r contenders for 2nd best. If player n does not play in the next comparison then two of the r contenders play and exactly one is eliminated in each such game. If all of these contenders but 1 is removed before player n plays then the remaining one plays against player n for the last comparison; this takes exactly r games after the pairings. If player n comes in earlier under R_{E} (actually for small r he plays first under R_{E}) then it still takes exactly r comparisons to complete the procedure. If player n wins then he continues to play each of the contenders for 2nd best until he loses. Since 1 contender is removed as a result of each comparison, it again takes exactly r comparisons after the pairings. Thus for $n = 2^{r} + 1$ and any procedure R_{c} that uses 'cycle pairing for n - 1' we have

(3.5)
$$E\{T|R_C\} = 2^r + r - 1 = M\{T|R_C\}.$$

¥

This result is analogous to (4.3) in [5]; it is proved to be M-optimal in (3.23) below, but the table above shows that it is not E-optimal.

We could also have gone through a cycle-pairing procedure for $n = 2^{r}$ and obtained the result 2^{r} + r - 2 as in (4.3) of [5] but, as our table shows, this result can be improved upon.

Clearly any of the attained values for the t = 2 ordering problem in [5] can be used as an upper bound for the optimal procedure in the t = 2selection problem.

- 9 -

and any proceeding light uses creic parting for a - 1 we have and in calles exactly x comparisons after the puttinuges. Thus for $n = 2^{+} + 1$ he loses. Since I builted as is readed as a readit of each conversion, it n wins thoshie contrarts to make each of the contraction for Sud Mest mill. dt still takes acestly i somperisons to complete the procedure. If player carlier under A. (actually for small y he plays direct under P.) that this takes eractly if games after the patrings. If player in cones in n plays then the constructions plays quarks player in for the last comparisout lin each such game. If will of these contenders but 1 is removed beince playm nart comparison them the of the s contranders play and exactly one is eliminated exactly i contenders for Sud best. The slares under not glay in the and for the relation $n_{\mathcal{S}}$. We have a bost player (sur $n-1=2^{\circ}$) and affer n - 2 confirtions de nave 1 player (say n) that has never played For a = S"+ 1. the cycler pair ing and complete pair ing are the date and Tor in - 1 we let an ever through the surface the start for the second start the structure and Since we bere a special interest in procedure. Chet uso wolk rein al

$(\cdots) \qquad \mathbb{E}[\mathbb{T}|\mathbb{R}_{0}] = \mathbb{E}^{1} + \mathbb{P}[\mathbb{L} = \mathbb{M}\mathbb{R}^{n} : [\mathbb{T}|\mathbb{R}_{0}].$

This result is graiogous to (4, 2) in []: it is proved to be Heopenial in ((4.2)) below, but the table prove shows that it is not d-optiment. (e doubt also have gone through a optimenting procedure for $n = 2^{\frac{1}{2}}$ and obtained the result $2^{\frac{1}{2}} + \pi - 2$ as in (h, i) of [] but, a our table shows, this result can be improved dom.

Showly why of the detained values for the t = 2 ordering problem duty] can be used as apper bound for the optimal procedure in the t = 2 scheetion problem.

- 3.4

We define a comparison $C(a \ v \ b)$ to be of level j(j = 1, 2..., [log n])and denote it by C_j if the two players a and b each have exactly 2^{j-1} -1 inferiors, the two sets of inferiors are disjoint and neither a or b has any proven superiors. We want to prove a result about the reduction in entropy for any comparison C_j of level j.

<u>Lemma</u>: If two players a, b, with no proven superiors, are best in disjoint subsets S_a, S_b respectively, with common size I for each, then the expected reduction in entropy $E\{\Delta\}$ due to the comparison a v b is given by

(3.6)
$$E{\Delta} = \frac{2I(2n-1-3I)}{n(n-1)}$$
,

i i i

regardless of any knowledge previously obtained that affects only the relative ordering of the remaining n - 2I players.

<u>Proof</u>: Suppose that the two players a and b each have exactly I - 1inferiors with no overlap (in our application $I = 2^{j-1}$) and, to begin with, there is no other previous information.

Those states of nature which can still be correct are partitioned into subsets so that each subset corresponds to one of the possible true decisions. The decisions D(x, y) corresponding to these subsets will be grouped into a convenient table, and the number of cases given for each, before we play a v b. Let $S_{\alpha}(\alpha = 1, 2, ..., m)$ denote the seconds (or immediate inferiors) of a in the connected set of size I and let M_{α} denote the number of ways of linearizing this set of size I with S_{α} in second position (consistent with all known order relations); let $M = M_1 + M_2 + ... + M_m$. Then $S'_{\beta}(\beta = 1, 2, ..., m')$, M'_{β} and M' are defined similarly for player b. Let p (or p_i) denote any one of the n - 2I players not in these two connected subsets.

- 10 -

We define a comparison $G(a \vee b)$ to be of level $\beta(f = 1, f..., [10, 4])$ and denote it if G_{f} if the two players a = c.d is each have encoding infertions, the two bets of infertern and disjoint and notified a or to be approved superfors. We sent to prove a result about the reduction in entropy for any comparison G_{f} of lowel it

Lenne: If two players of the n fives players and host in fojoint subsets J_np5₀ respectively, uith conner size I for each, then the unpoted reduction in entropy [10.4] can to the gamericon of a la diven by

$$(3.6) \quad \forall (5.5) = \frac{\nabla I(82^{-1} - 3I)}{H(1 - 1)},$$

referdiese of the indeficus providently obtained thet affects only the relative ordering effects are the relative ordering effects are the second of the provident of the provi

There also also a first each guines conversioner for each of the possible true into subsets so that each guines covrespondents to one of the possible true decisions. The decketone $D(\pi, \pi)$ corrected for a close subset will be grouped into a conventent coller, and the number of order piver for each, he fore we play only by the $J_{\rm eff}$ is the terms of also find the econd. (or invading inferious) of a in the converted set of the I and let $M_{\rm eff}$ denote the number of each of the still cover relations); is essent position (converted with all cover relations); is a set with the observe of the still of the relations of the subset of the still be

<u>_</u> T	ype of Decision	Number of Each Type	Number of Cases per Decision
1.	D(a, b)	1	2MM' $\frac{(n-2)!}{(I-1)!(I-1)!}$
2.	$D(a, S_{\alpha})$	$(\alpha = 1, 2,, m)$	$M_{\alpha}M' \frac{(n-2)!}{1!(1-2)!}$
3.	D(a, p)	n - 21	2MM' $\frac{(n-2)!}{1!(1-1)!}$
(3.7) 4.	D(b, S _β)	$(\beta = 1, 2,, m')$	$MM'_{\beta} \frac{(n-2)!}{1!(1-2)!}$
5.	D(b, p)	n - 21	2MM' $\frac{(n-2)!}{1!(1-1)!}$
6.	D(p ₁ , p ₂)	(ⁿ - 2I)	2MM' (n-2)! I! I!

The uncertainty before playing a v b can now be obtained directly from the above table. The total number of cases is easily checked to be n! $MM'/(I!)^2$ and with this as denominator and the number of cases above as numerator, we have the probability for each decision type.

We now play a v b and suppose a > b. Then types 2 and 3 in the table remain in their entirety and we also have half the number of cases in type 1 and half for each decision of type 6. Since types 3 and 5 have the same number of cases and the sum over α for type 2 equals the sum over β for type 4, it follows that the comparison a v b partitions the entire set of cases exactly in half. Hence we use a simple average for finding the expected uncertainty after playing a v b.

The uncertainty U_1 after finding that a > b is given by

$$(3.8) \qquad U_{1} = \frac{2I(I-1)}{n(n-1)} \sum_{\alpha=1}^{m} \frac{M_{\alpha}}{M} \log \frac{M n(n-1)}{2M_{\alpha}I(I-1)} + (n-2I) \frac{4I}{n(n-1)} \log \frac{n(n-1)}{4I} \\ + \frac{2I^{2}}{n(n-1)} \log \frac{n(n-1)}{2I^{2}} + {n-2I \choose 2} \frac{2}{n(n-1)} \log \frac{n(n-1)}{2}$$

and we omit the corresponding U_2 after finding that b > a. The orginal uncertainty U_0 from the above table is

- 11 -

	E(p* L)	사망하다. 1997년 - 전문 전 전화 (1997년 1997년) 1997년 - 전문 전 전화 (1997년 1997년)	······································
3- , -	ມ(;; ; ;)	$(\mathbf{p}_{i}) = \mathbf{p}_{i} \cdot \mathbf{p}_{i} \cdot \mathbf{p}_{i} \cdot \mathbf{p}_{i}$	$\frac{\mathbf{I} \cdot (\mathbf{I} - \mathbf{i}_{1})}{(\mathbf{I} - \mathbf{i}_{1})}$
	1)(a., 1))		∰77. <u>I:(I-I):</u> (#):
	π(υ. 3.)	$\left(\mathbf{c}_{i} = \left\{ \mathbf{c}_{i}^{*}, \mathbf{c}_{i}^{*}, \mathbf{c}_{i}^{*}, \mathbf{c}_{i}^{*} \right\}$	$\sum_{i=1}^{N} M_{i} = \frac{1}{(1-i)}$
	n(c* =)		UTEN: (I-I);(I-I); (I-I);
		<u>। भूगलेल्डा दर्श उत्तव्य होन्द्र</u>	Manbar al Graes you Beatrain

The hypericative terms first fighting the can new be obtained discosing through the first fighting the case is called a discound to be using the case is called to be using (1,). and the chies as dependence and the proper of near shore commendate, we have the properties in the decision type.

The non-plethor of a subfracty and to how of No. The type, the dec solic reading in their entirent and we size have that the measure is exact in type 1 and bold lot each decision type, or size type is and y have the same Weyber of cance and the subject of the type of the type is the sume even for type is, it follow that the constitution is the custom

seboef carge anactly digheld. Hence we use a shale average for finding off composited underesting affect (lefting a

The mocrueiner U sister Christing the set is riven by

 $(3,5) \quad I_{1}^{(1)} = \frac{n(n-1)}{2\mathbb{E}(1-1)} \stackrel{(3=1)}{=} \frac{1}{2} \frac{1}$

 $+ \frac{u(z-z)}{zz_{z}} + \frac{u(z-z)}{u(z-z)} + \frac{u(z-z)}{z} + \frac{u(z-z$

rud no child the qointerpositing. U. siden finding, that the cryinel uncantering U. Siden field for the standard in

$$(3.9) \qquad U_{o} = \frac{I(I-1)}{n(n-1)} \sum_{\alpha=1}^{m} \frac{M_{\alpha}}{M} \log \frac{M n(n-1)}{M_{\alpha} I(I-1)} + \frac{I(I-1)}{n(n-1)} \sum_{\beta=1}^{m'} \frac{M_{\alpha}'}{M'} \log \frac{M' n(n-1)}{M_{\beta}' I(I-1)} + 2(n-2I) \frac{2I}{n(n-1)} \log \frac{n(n-1)}{2I} + \frac{2I^{2}}{n(n-1)} \log \frac{n(n-1)}{2I^{2}} + (\frac{n-2I}{2}) \frac{2}{n(n-1)} \log \frac{n(n-1)}{2} .$$

To obtain the expected reduction in uncertainty we first remove log $\frac{1}{2}$ = -1 from each term in U_1 and U_2 and we then easily obtain from (3.8) and (3.9)

(3.10)
$$U_0 - (\frac{U_1 + U_2}{2}) = 1 - \frac{2I^2}{n(n-1)} - \frac{(n-2I)(n-1-2I)}{n(n-1)}$$

= $\frac{2I(2n-1-3I)}{n(n-1)}$.

If there is present previous knowledge about some of the remaining n - 2^{j} players (say, among r of them) then we first consider each possible fixed relative ordering among these r players separately. Each of the numbers in the third column above is affected (we actually divide by r!) but this constant r! also divedes the total. Hence the probabilities and the uncertainty is unchanged. Since the previous knowledge about these r players can be written as a union of such fixed relative orders among them, we now average our result over all the relative orders of these r players consistent with our previous knowledge and obtain the same result; this proves the lemma.

For the application of this lemma that is needed, we have the <u>Corollary</u>: The expected reduction in entropy $E\{\Delta_j\}$ due to any jth level comparison $C_j(a v b)$ for $1 \le j \le \lfloor \log n \rfloor$ is given by

(3.11)
$$E\{\Delta_j\} = \frac{2^{j}(2n-1-3\cdot 2^{j-1})}{n(n-1)}$$

regardless of any knowledge previously obtained that affects only the relative ordering of the remaining $n - 2^j$ players. - 12 -

and the second se $\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} = \frac{1}{2} - \frac{1}{2} -$ $= \frac{1}{2(1-1)} = \frac{1}{2} \frac{1$

or carve of the Received and a part of the

We now use the lemma above to obtain lower bounds for all procedures that start with any given pairing scheme. We consider only the 'cyclepairing for n-1' schemes since they give the better results, but the method can also be applied to complete-pairing schemes, cycle-pairing schemes (for n), ordinary-pairing schemes, etc. For $n = 2^r + c$ ($0 < c \le 2^r$) any cycle-pairing procedure R_c has at least 2^{r-1} comparisons of level i(i = 1, 2, ..., r-1) and we assume it has exactly that many among the first $2^r - 1$ comparisons. Then the (expected) reduction in entropy due to these comparisons is, by (3.6),

(3.12)
$$Q = \sum_{j=1}^{r} \frac{2^{r-j} 2^{j} (2n-1-3 \cdot 2^{j-1})}{n(n-1)} = \frac{2^{r} \{(2n-1)r-3(2^{r}-1)\}}{n(n-1)}$$

Let the total number of comparisons T be partitioned into T_1 and T_2 where $T_1 = 2^r - 1$ are the pairings and $T_2 = T - T_1$ are the remaining comparisons under procedure R_c . Since the total uncertainty at the outset is log $\binom{n}{2}$ and 1 is an upper bound for the reduction in entropy in any one step, it follows that

$$(3.13) \qquad Q + 1 \cdot E\{T_2\} \ge \log {n \choose 2}.$$

Hence, using (3.13), we obtain the desired lower bound for the expectation under any cycle-pairing procedure R_{c} with $n > 2^{r}$

$$(3.14) \quad E\{T|R_{C}\} = 2^{r} - 1 + E\{T_{2}\} \ge 2^{r} - 1 + \log \binom{n}{2} - \frac{2^{r}\{(2n-1)r - 3(2^{r}-1)\}}{n(n-1)}$$

that there will note then thinking schools. We consider only fits legaleretring for n-1 -cohomes since they have the follow perils, but the retring for n-1 -cohomes since they have the follow releates, but the reheated can also be applied to complete think releates, or de-pairing reheated (for n), endinerg-pairing schemes, not. For n = 1 + c ($0 < c \leq 1$) sup of the maining pronodume T_c for at least 2^{n-1} containing on a first i(1 = 1, 2, 4, ..., n-1) and we as more for the coordination is enough the first $\frac{1}{n} - 1$ comparisons. This the (schemested) relation is enough the steril the comparisons is, by (3.6);

No now use the liture share it often lover bernet for all recondence

$(z \cdot z \cdot z) = \frac{z}{z} \frac{z(z-1)}{z(z-1)} = \frac{z(z-1)}{z(z-1)^{2} - z(z-1)^{2}} \cdot \frac{z(z-1)}{z(z-1)^{2} - z(z-1)^{2} - z(z-1)^{2}} \cdot \frac{z(z-1)}{z(z-1)^{2} - z(z-1)^{2}} \cdot \frac{z(z-1)}{z(z-1)^{2} - z(z-1)^{2}} \cdot \frac{z(z-1)}{z(z-1)^{2} - z(z-1)^{2} -$

Let the bols fourism of comparisons if sempticities is a $\mathbb{Z}_{\hat{1}}$ con $\mathbb{Z}_{\hat{2}}$, where $\mathbb{Z}_{\hat{1}} = \mathbb{Z}^{1} - 1$. Che ind particular and $\mathbb{Z}_{\hat{2}} = \mathbb{Z}^{1} - 1$, che ind particular and $\mathbb{Z}_{\hat{2}} = \mathbb{Z}^{1} - 1_{\hat{1}}$, and the reaction $\mathbb{Z}_{\hat{2}}$, where $\mathbb{Z}_{\hat{2}} = \mathbb{Z}^{1} - 1_{\hat{1}}$, and the reaction $\mathbb{Z}_{\hat{2}}$, where $\mathbb{Z}_{\hat{2}} = \mathbb{Z}^{1} - 1_{\hat{1}}$, and the reaction $\mathbb{Z}_{\hat{2}}$, where $\mathbb{Z}_{\hat{2}} = \mathbb{Z}^{1} - 1_{\hat{1}}$, and the reaction $\mathbb{Z}_{\hat{2}}$, the first of $\mathbb{Z}_{\hat{2}}$ of the order $\mathbb{Z}_{\hat{2}}$. Since the conduction definition of the conduction $\mathbb{Z}_{\hat{2}}$ and the conduction $\mathbb{Z}_{\hat{2}}$ and $\mathbb{Z}_{\hat{2}}$ and $\mathbb{Z}_{\hat{2}}$ and $\mathbb{Z}_{\hat{2}}$ of $\mathbb{Z}_{\hat{2}}$ and $\mathbb{Z}_{\hat{2}}$ and $\mathbb{Z}_{\hat{2}}$.

Merice, meint (3.12), we belief in the declined lower bound for the emperication muter only endiaterizing presentate 2. with a >

 $(2,1^{*}) = (2^{*})^{*} = 2^{*} - 1 + 2(2^{*}) \ge 2^{*} - 1 + 1e_{2} \left(\frac{n}{2}\right) = \frac{n(n-1)^{*}}{n(n-1)^{*}}$

For $n = 2^{r} + 1$ we obtain an improvement over (3.14) by using (3.3). For $n = 2^{r}$ we can use the same result (3.14) provided we replace r by r - 1. The values of (3.14) are given as the LBE in the table in Section 2. It follows from the above construction that the LBE is strictly increasing in n. Asymptotically $(n \rightarrow \infty)$, the value of (3.14) is between $(n/2) + \log n$ (the limit for $n = 2^{r}$ as $r \rightarrow \infty$) and $n + (\log n)/n$ (the limit for $n = 2^{r} + 1$ as $r \rightarrow \infty$).

÷.

We now derive an exact formula for the expectation under procedure R_A . Let 2^s denote the number of players involved in the initial cycle-pairing. If $n = 2^r + c$ (with $0 \le c < 2^r$) then s will usually be r - 1 or r; the value of n where it changes from r - 1 to r is close to $3 \cdot 2^{r-1}$.

The probability that the q^{th} newcomer is the first one to beat the original x_1 is easily shown to be

(3.15)
$$\frac{1}{2^{s}+q} \prod_{\alpha=0}^{q-2} \left(\frac{2^{s}+\alpha}{2^{s}+\alpha+1} \right) = \frac{2^{s}}{(2^{s}+q-1)(2^{s}+q)} (q = 1, 2, ..., n-2^{s})$$

and the sum of these is $1 - 2^{8}/n$. It follows that the probability that x_{1} is beaten by the last newcomer or by no one is $2^{8}/(n-1)$. Conditional on the event that the q^{th} newcomer is the first to beat x_{1} , we utilize one extra comparison (beyond the basic one needed for each newcomer) under endings E_{1} and E_{2} for earlier newcomers n_{α} $(1 \le \alpha \le q-1)$ with probability $(2^{8-1}+\alpha)^{-1}$ and for later newcomers n_{β} , not including the last two, (i.e., for $q + 1 \le \beta \le n - 2 - 2^{8}$) with probability $2(2^{8}+\beta)^{-1}$. In addition the q^{th} newcomer also uses 1 extra comparison to beat x_{1} . For $\beta = n - 1 - 2^{8}$ (and $n \ge 2^{8} + 3$) we add 2 extra comparisons only if n_{0} beats x_{2} , which then has probability 3/n. Given ending E_{3} , which, as stated above, has probability $2^{8}/(n-1)$, we have to play s - 1 extra games after n_{0} plays x_{2} and each of the previous newcomers n_{α} $(1 \le \alpha \le n-1-2^{8})$

- 14 -

For $n = 2^{n}$ we obtain an approvement over ((.15) of veloci (...), For $n = 2^{n}$ we can use the annamesult (with) why delive velocity to by n = 1. The values of ((.15) are fiven as the JBD in the fable in puttion 2. It follows from the source construction that the LDN is structly include in in ... Asymptotically (n - ...), the value of ((.15) is between (n/2) + 100 n (the fimit for $n = 2^{n}$ as $n \to 0$) and n + (100 n)/n (the list for $n = 2^{n} + 1$ as $n \to 0$).

We now derive in endot forwill for the expectation multiproduct in . Let 2^{3} denote the number of players involved in the initial dysid-pair n_{0} . Fi $n = 2^{2} + 2$ (with $0 \le 0$ 2^{3}) from a will contribute r = 1 dy risting value of n where it changes from r = 1 to r is close to 2^{n-1} . The probability that the q^{3} near q^{3} have is the first one to part the

art frail and so so so so be

$$(\cdot, \tilde{\mathbf{r}}) \qquad \frac{1}{2^2 + q} \quad \frac{2^2 + q}{2^2 + q} \quad (\frac{2^2 + q}{2^2 + q}) = \frac{2^2}{(2^2 + q - \tilde{\mathbf{r}})(2^2 + q)} \quad (q = 1, 2, \dots, 2^{-2^2})$$

and the sum of them is $1 - n^2/n$. To follow that is probability that n_1 is parten by the init members on by moveds is $n^2/(n-1)$. Conditional on the event that the numbers of the direct to heat n_1 , we still a one entry condition (beyond the basic one manded for each headplas) under antings n_1 and n_2 for carlier nervouses n_1 (1 $\leq -1-1$) with probantings n_1 and n_2 for carlier nervouses n_1 , not including the last two, $(1, 2, ..., for <math>q + 1, ..., 5, ..., 1 - ..., 2^3)$ with probability $2(2^3 + ...)^{-1}$. In addition the $q^{(2)}$ nervower also uses 1 entry comparison to beat n_1 , for $i = n - 1 - 2^3$ (and $n \geq 2^3 + ...)$ we add 2 extra comparison to beat n_1 . To be ats n_2^2 , which then has probability n_1 . Given and n_2^2 is follow as n_2^2 and $n_2^3 + ...)$ is a down of the direction $n_1 + 2^3$ ($n_1 + ... + 2^3$) we have to obtain n_1 . The $n_1 + 2^3$ (and $n \geq 2^3 + ...)$ we add 2 extra comparison n_1 which has n_2^2 beats n_2^2 , which then has probability n_1 . Given and n_2^2 is followed n_1 . plays an extra game with probability $(2^{s-1} + \alpha)^{-1}$. If we add these extra games to the set of $n - 2^s$ basic games, one for each newcomer, and the 2^s-1 comparisons used in the initial pairing, then we obtain

(3.16)
$$E(T|R_{A}) = (2^{s}-1) + (n-2^{s}) + \frac{2^{s}}{n-1} \left[s-1 + \frac{2^{s}}{2} - \frac{1}{2^{s-1}+\alpha} \right] + \frac{n-1-2^{s}}{2^{s-1}+\alpha} + \frac{n-2-2^{s}}{2^{s}-1} - \frac{1}{2^{s-1}+\alpha} + \frac{n-2-2^{s}}{2^{s}+\beta} - \frac{2^{s}}{2^{s}+\beta} + \frac{2^{s}}{2^{s}+q} - \frac{2^{s}}{2$$

In (3.16) we make extended use of the elementary identity

(3.17)
$$\sum_{\substack{i=a+1 \\ i=a+1}}^{n-b} \frac{1}{(c+i+1)(c+i)} = (\frac{1}{c+a} - \frac{1}{c+n-b})h_n(a+b+1)$$

where $h_n(x) = 1$ for $n \ge x$ and = 0 otherwise; in particular each of the double sums can be summed or simplified by (3.17). After some straightforward algebra and simplification we obtain

(3.18)
$$E\{T|R_A\} = n - 1 + 2^{s}(\frac{s-1}{n-1}) + 2\sum_{\alpha=1}^{n-2-2^{s}} \frac{1}{\alpha+2^{s-1}} + h_n(2+2^{s})\{\frac{n2^{s}}{(n-1)(n-2)} + \frac{12(n-2-2^{s})}{n(n-1)(n-2)} + \frac{2^{s}}{(n-1)(n-1-2^{s-1})} - 1\}.$$

Under 'cycle-pairing for n-1' with $n = 2^r + 1$ we set s = r and we easily obtain n + r - 2 in agreement with (3.5).

Asymptotically $(n \rightarrow \infty)$ we obtain the maximum and minimum of (3.18) by searching for the value of n for which g(r-1, n) = g(r, n) where g(s, n)is the right side of (3.18). In fact we obtain from (3.18) the quadratic equation

$$(3.19) n2 ln 2 - n 2s-2(s + 2 ln 2) + (s + 1)22s-3 \approx 0$$

plays an extra game with probability $(2^{3-3}+..)^{-1}$. If no and thus extra games to the set of $n - 2^{3}$ basic games one for each numbers, and the $2^{3}-1$ comparisons used in the initial pairing, then we obtain

$$(1,1) \qquad \mathbb{S}(\mathbb{T}[\mathbb{R}]) = (2^3 - 1) + (2 - 2^3) + \frac{2^3}{2^{3-1}} (3 - 1 + \frac{2^{3-2}}{2^{3-1}} (3 - 1 + \frac{2^{3-2}}{2^{3-1}})$$

$$\begin{array}{c} \overset{ij}{\leftarrow} \overset{ij$$

In (.. 1 .) we make submided the of the chemertary (dent ty

where $h(x) = h^{2}$ for $h \ge x^{2}$ and = 0, otherwheet in purcharlar and of the double sums can be simmed of simplified $\gamma(x, \lambda)$. After none structure forward cheapterent succession we obtain

$$(S, I, U) \quad \mathbb{E}\left[T : \mathbb{R}_{A} \right] = n - 1 + 2^{3} \left(\frac{n-1}{2+1}\right) + 2^{3} \frac{1}{1-2}$$

$$+ p^{n}(3+2) \frac{(n-1)(n-2)}{(n-1)(n-2)} + \frac{n(n-1)(n-2)}{(n-1)(n-2)} + \frac{(n-1)(n-1-2)}{(n-1-2)(n-1-2)} - 1$$

Uniter the particulation $\pi - 1$, with $\eta = 2^{\frac{1}{2}} + 1$ we set $s = \pi$ and we conside obtain $\pi + \pi - 2$ in spreakent with (-1).

Asymptopically $(n \rightarrow w)$ we obtain the autimum and summan of (...15) by searching for the value of n for this (v, 1) = f(v, 2) where f(v, 1)is the value of (...15). In fact we obtain from (...15) the duckate equation

where $\ln x$ is the natural logarithm of x. Since we are looking for a final change for s as n increases, we take the largest root for n in (3.19) and obtain

(3.20)
$$n \approx 2^{s-2} \left(\frac{s}{\ln 2} + \frac{\ln 2}{s}\right).$$

The limiting value of (3.18) for this value of n (which is the same for both s = r and s = r - 1 and hence is the desired limiting value) is by (3.20)

$$(3.21) \quad E\{T | R_{A}\} = n + 2 \ln s + \mathcal{O}(1) = n + 2 \ln \ln n + \mathcal{O}(1).$$

Based on the form of (3.21) we conjecture that a lower bound for all procedures for $n \ge 3$ is given by

$$(3.22) \qquad \text{CLB} = n - 2 + (2 \ln 2) \log \log n.$$

A lower bound for the minimax problem can be obtained using a modification of a method due to Slupecki [4]. Our result is stated in two parts; the first gives some easily derived inequalities that any lower bound has to satisfy and the second shows that what we got in the first part is an attainable lower bound for the maximum length of any branch of the procedure. Lemma: For $n \ge 3$ players a lower bound (LBM) for the maximum branch length of any procedure for the selection (t = 2) problem must satisfy the inequalities

$$(3.23) \qquad LBM \leq \begin{cases} n-2+[\log (n-1)] & \text{if } n \text{ is of the form } 2^r+1\\ n-1+[\log (n-1)] & \text{otherwise.} \end{cases}$$

<u>Proof</u>: The second part of (3.23) is the same as the Schreier [3] and Slupecki [4] result for the t = 2 ordering problem. Since the selection problem cannot require more than the ordering problem, this part is clear. For the first part we need only put one unit aside until the very last comparison (see procedure R_M above) and use the second part again for n - 1 units.

where Instricts the flattered logardithe of st. Since we are looting for a

finel change for 3 23 a indresses, we take the largest root for a in (:.ly) and obtain

$$(\cdot,2) \qquad n \leq 2^{5-2} \left(\frac{1}{16} + \frac{1}{2}\right),$$

The limiting value of (.11) for this value of a (which is the same for both s = r and s = r - 1 and hence is the desired limiting raise) is by (3.20)

(3.21)
$$\mathbb{E}[T_{N_{1}}] = a + 2$$
 in s $\pm 0(1) = a + 2$ in in $n + 0(1)$.

Based on the former of (3.21, we crack that a fourt bound for all y

A lower bound for the minimizer robbies can be obtained using a solution of a rethol ine to Simpole. We, we can be soluted a two educes the first gives some essify lettred inside that any lower bound has to satisfy and the second chore that what we get the first part is an attainable lower bound for the eminimum length of any branch of the provedure. Jumma: For $n \ge players a$ lower bound for the selection (LBM) for the second branch length of any procedure for the selection (t = 2) product must subtain the inside laws

(3.2.), IBM
$$\begin{bmatrix} n + 2 + [105 (n-1)] & \text{if } n + 8 & \text{of tBurform } 2^{+} \\ n + 1 + [103 (n-1)] & \text{otherwise.} \end{bmatrix}$$

Front: The shound part of $\binom{3}{3}$, 21) is the same as the following if each in each $\binom{3}{3}$ solution for the same as the school of $\binom{3}{3}$ solution problem. Thus the school of problem of the school of problem of the school of the sch

Since this actually gives us an ordering of the t = 2 best in n - 1, it follows that we need only one more comparison (the unit set aside versus the second best of n - 1). This gives a total of

$$(3.24) \qquad n - 2 + [\log (n - 2)] + 1 = n - 2 + [\log (n - 1)],$$

which proves the lemma.

For the second part we follow the method of Slupecki and consider a class of procedures (or a system) S that is characterized by the fact that defeated players do not enter into the first n - 1 comparisons. We show below (3.25) that the maximum branch length of any procedure in S is equal to or greater than the right side of (3.23). Since the early inclusion of defeated players can only lengthen our procedure, it follows that our results holds a fortiori for all procedures. The two results taken together then prove the

<u>Theorem</u>: For $n \ge 3$ players the best lower bound (LB_0) for the maximum branch length in any procedure is given by the right side of (3.23) or for all $n \ge 3$ by

$$(3.25) \qquad LB_0 = n - 1 + [log (n - 2)].$$

<u>Proof</u>: It is easily checked that the theorem holds for n = 3 and 4; we use these as starting values for an induction proof. Let k be any integer > 4 and suppose the theorem holds for every integer i such that $3 \le i < k$. Let s (approximately equal to k/2) denote the number of games in the first round in which the k players take part. In the games that follow the first round we use the induction hypothesis on k - s players (the winners of the first round plus any players that did not play in the first round). The first round gives us no information for the M-problem for selecting (or ordering) these k - s players. After getting the two best of these k - s

- 17 -

Jince this actually gives is an ordering of the 't = 2 bast in n - 1, it follows that we need only one move comparison (the unit set as is versus the second bast of 'n - 1). This gives a total of

 $(1.2+) \quad n-2+(10g(n-2)]+1=n-2+(10g(n-1)),$

which proves the loans.

For the second part we follow the method of simpaci, and consider a class of procedures (or a system) 3 that is characterized by the fact that defeated players do not enter into the first n - 1 comparisons. A shart to show the the tender length of any properime in is equal to or rester than the roght side of (.23). Since the early inclusion of defeated players can only lengthen our procedury. I incluse the results holds a forther all plocedures. The from results to be forther all plocedures. The from the forther the formal for all plocedures.

Theorem: For $n \ge 3$ pluyers the base lower bound (LD₀) for the mathematical branch langed in the procedure is given by the tilthe side of (2.23) or for ell $n \ge -by$

(3.25) $3.50 = 2 - 2 + (10) (2 - 2)^{-1}$

Proof: It is easily diveded that the theorem bolds for n = 3 and k_3 we be there as starting values for an intervent proof. Let is be approve the theorem holds for every intervent is such that is the theorem holds for every intervent is such that is the theorem holds for every intervent of fames in the first is capability equal to k/k_3 denote the number of fames in the first of the first is the first of the

players, we then take into account the two people (at most) that they defeated in the first round.

We first note that the two expressions on the right side of (3.23)are equivalent to the single expression on the right side of (3.25). To see this we need only consider the cases $n = 2^{r} + 1$ and $n \neq 2^{r} + 1$ separately. Hence we can use either of these expressions; we prefer to use (3.23) for the induction hypothesis on k - s players.

For k = 4m we take s = k/2 and by the induction hypothesis the k - s players require $(k - s - 1) + [\log (k - s - 1)] = k - s - 2 + [\log (k-2)]$ games. By the Schreier-Slupecki result we can assume that not only have we selected the two best of k - s - 1 but also that they are ordered. Hence we need only 1 extra game for completion (the contender for second best against the one defeated in the first round by the best player). This gives the total number of binary comparisons

 $(3.26) \qquad s + (k - s - 2) + [\log (k - 2)] + 1 = k - 1 + [\log (k - 2)],$

which is the desired result for k = 4m.

्र**म** ्री

For k = 2(4m - 1) we take s = k/2 and the steps are similar to those above giving (3.26).

For k = 2(4m + 1) we take s = (k/2) - 1 and again we obtain the right side of (3.26).

For k = 4m - 1 we take s = (k - 1)/2, so that k - s = 2m is even and the steps are again the same.

For k = 4m + 1 we take s = (k + 1)/2, so that k - s = 2m is even, and we obtain for the total

$$(3.27) \qquad s + (k - s - 1) + [\log(\frac{k-3}{2})] + 1 = k - 1 + [\log(k - 2)],$$

which completes the induction.

plarors we then take into account the two people (at most) that they defeated in the first round.

We direct mode that the two emphassions on the right side of (3.25) are aquivalant to the single expression on the right side of (3.21). To see this we need only consider the cases $n = 2^{4} + 1$ and $n = 2^{4} + 1$ such the Canada de close of the superstelly. Made we can use either of these expressions; we prefer to use (3.23) for the induction hypothes side k + s players.

For, k = 2n - 2n take s = k/2 and by the induction hypothesis the k - s players fequire (k - s - 1) + (log (k - s - 1)) = k - s - 2 + (log (k - 2)) gaves. (Bytche Schrim Flupeck vestic we can assume that not only have represed the two bast of k - s - 1 but also that they are ordered. Hence we want only light for some lattice (the contendar for sound bast

ac con our restances densities come to the concernes for the second second

(3.26) $s \div (1 - s - 2) \div 108 (1 - 2) + 1 \div (1 - 2) + 108 (1 - 2) + 1$

which is the desired result for the burn

For k = 2(4k - 1) we take s = k/2 and the step, are similar to thus above given (.26).

For $\mathbb{R} \Rightarrow 2^{(3m)} \oplus 2^{(3m)}$ we fake s = (3/2) - 1 and such we obtain the

AT & PE (3:59)

For T = 400 - 1 we take s = (t - 1)/2 so that k - s = 2. 's weak and the steps are again the bane.

For $\alpha = 4n_{12} + 1$ we take s = (k + 1)/2, so that $\alpha - s = 2n_{12}$ is even, and we obtain for the total

$$(1,2?)$$
 $s \div (\underline{s} - s) + (1o_{\mathbb{S}} (\frac{\overline{s} - s}{2})) \div 1 = 1 - 1 \div 1o_{\mathbb{S}} (1 - 2)),$

Wolath completes the induction.

Since the procedure R_{M} has maximum length equal to the right side of (3.25) it follows that R_{M} is M-optimal for the t = 2 selection problem.

4. Acknowledgement.

ŝ

The author wishes to thank Dr. Abdollah Hadian for many interesting conversations on this research. Thanks are also due to the referee who suggested that the proof of the lemma in section 3 could be generalized from the original form, which is now written as a corollary.

References

- [1] Dubail, F. (1967). Algorithmes de questionnaires realisables optimaux au sens de différents critères. Thesis à la Faculte des Sciences de l'Université de Lyons.
- [2] Picard, C. (1965). Theorie des Questionnaires. Gauthiers-Villars, Paris.
- [3] Schreier, J. (1932). On tournament elimination systems (in Polish). <u>Mathesis Polska</u> 7 154-160.
- [4] Slupecki, J. (1949-51). On the system S of tournaments. <u>Colloq</u>. <u>Math</u>. <u>II</u> 286-290.
- [5] Sobel, M. (1968). On ordering the t = 2 best of n items using binary comparisons. Submitted for publication.
- [6] Sobel, M. and Groll, P. A. (1959). Group testing to eliminate efficiently all defectives in a binomial sample. Bell System Tech. J. <u>38</u> 1179-1252.

<u> andre a</u>

- [1] Dubail, F. (1907). Algèrichase de quéstionnaires réclisables operands an cens de différence outeures. Thesis à le Foculte des Cohmess de 1⁴ Université de Lyoné.
- [0] Fienre, C. (1965) & Treoric cos muceriorschien. Generatore-Villan, Faris.
- [3] Sohuter, J. (1.30). On connector classication systems (in Belink).

MALIANTE SPIRES INF-100:

[4] Sluppett, J. (195-51). On the system S of territorent. Colled. Main.

II \$90-500

- [5] Sobel, M. (1963): (main optimal scarch for the forest using only instbinary erroless comparisons :I. Schaifled to Collequium Fublications (empt) is optimated in
- (CNRS) on Questionnaire Theory, edited by C. Picard.
- [5] Sobel, H. and Grolf, P. .. (1.79). Group feature to eliminate efficiently
- cli defective inad binómial servic. Jell System Techi J. 32

11.0-1555