A Sample Article for the LIPIcs series*

Oliver D. L. Grant[1] and J. Ian Munro[2]

1 **University of Waterloo, Waterloo, Canada**
`odlgrant@uwaterloo.ca`

2 **University of Waterloo, Waterloo, Canada**
`imunro@uwaterloo.ca`

Abstract

We examine optimal and near optimal solutions to the classic binary search tree problem of Knuth. We are given a set of n keys $B_1, B_2, ..., B_n$ and $2n+1$ frequencies. $p_1, p_2, ..., p_n$ represent the probabilities of searching for each given key, and $q_0, q_1, ..., q_n$ represent the probabilities of searching in the gaps between and outside of these keys. We have that $\sum\limits_{i=0}^{n} q_i + \sum\limits_{i=1}^{n} p_i = 1$. First, we re-examine an approximate solution of Güttler, Melhorn and Schneider which was shown to have a worst case bound of $c \cdot H + 2$ where $c \geq \frac{1}{H(\frac{1}{3}, \frac{2}{3})} \approx 1.08$, and $H = \sum_{i=1}^{n} p_i \cdot \lg(\frac{1}{p_i}) + \sum_{j=0}^{n} q_i \cdot \lg(\frac{1}{q_j})$ is the entropy of the distribution. We give an improved worst case bound on the heuristic of $H+4$. Next, we examine the optimum binary search tree problem under a model of external memory. We use the Hierarchical Memory Model of Aggarwal et al. The model has an unlimited number of registers, $R_1, R_2, ...$ each with its own location in memory (a positive integer). We have a set of memory sizes $m_1, m_2, ..., m_l$ which are monotonically increasing. Each memory level has a finite size except m_l which we assume has infinite size. Each memory level has an associated cost of access $c_1, c_2, ..., c_l$. We assume that $c_1 < c_2 < ... < c_l$. We propose an approximate solutions which run in $O(n)$ time where n is the number of words in our data set. We also examine the related problem of binary trees on multisets of probabilities where keys are unordered and we do not differentiate between which probabilities must be leaves, and which must be internal nodes. We provide a simple $O(n \lg(n))$ algorithm that is within an additive $\frac{n+1}{2n}$ of optimal on a multiset of n keys.

1998 ACM Subject Classification Dummy classification – please refer to `http://www.acm.org/about/class/ccs98-html`

Keywords and phrases Dummy keyword – please provide 1–5 keywords

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

1 Introduction and Background

In this chapter we provide an introduction to binary search trees and the optimum binary search tree problem. We also give motivation for studying binary search trees and give an overview of the work presented in this thesis. We note that for the entirety of this work, we will use lg to represent log_2.

Knuth first proposed the optimum binary search tree problem in 1971 [**?**]. We are given a set of n keys (originally known as words), $B_1, B_2, ..., B_n$ and $2n+1$ frequencies. $p_1, p_2, ..., p_n$ represent the probabilities of searching for each given key, and $q_0, q_1, ..., q_n$ represent the

* This work was partially supported by someone.

probabilities of searching in the gaps between and outside of these keys. We have that

$$\sum_{i=0}^{n} q_i + \sum_{i=1}^{n} p_i = 1$$

We also assume without loss of generality that $q_{i-1} + p_i + q_i \neq 0$ for any $i \in \{1, ..., n\}$. Otherwise, we could simply solve the problem with key p_i removed. The keys must make up the internal nodes of the tree while the gaps make up the leaves. Our goal is to construct a binary search tree such that expected cost of search is minimized. This expected cost of search is also sometimes referred to as the expected path length. It is formally defined as:

$$P = \sum_{i=1}^{n} p_i \cdot (d_T(B_i) + 1) + \sum_{j=0}^{n} q_j \cdot (d_T(B_j, B_{j+1})) \tag{1}$$

where p_i and q_j are the probabilities of searching for key B_i or gap (B_{i-1}, B_i) respectively and $d_T(B_i)$ and $d_T(B_j, B_{j+1})$ are the depths of the internal node for B_i and the leaf for (B_j, B_{j-1}) respectively in the tree T. Note that we assume that B_0 represents $-\infty$ and B_{n+1} represents ∞. Note that we charge 1 extra to search for a key at depth l than a leaf at depth l because it requires an extra operation to confirm an internal node, whereas we do not need this confirmation if the node is a leaf. The optimal solution of Knuth uses dynamic programming and requires $\Theta(n^2)$ time, and $\Theta(n^2)$ space [**?**]. This solution is both time and space intensive. We will later examine an approximate solution to this problem of Güttler, Mehlhorn and Schneider (the Modified Entropy Rule) which uses $O(n^2)$ time but $O(n)$ space [**?**]. We will improve its worst-case expected search cost bound. While all of the aforementioned algorithms examine the problem in the RAM model, we will also examine the problem in more realistic models of memory and look at approximate solutions under these settings.

While modern computers typically only support two-way branching, the optimum binary search tree (BST) problem proposed by Knuth uses the three-way branch model. This model allows a single comparison operation to transfer control to three different locations.

Examples of this can be seen in FORTRAN IV which describes the arithmetic IF [**?**]:

```
IF (EXPR) LABEL1, LABEL2, LABEL3
```

Control is transferred to LABEL1, LABEL2 or LABEL3 if $expr < 0$, $expr = 0$, or $expr > 0$ respectively using a single comparison command. While modern programming languages scarcely use this arithmetic IF, and compilers may simply encode such expressions using multiple logical if statements, many machines in in the FORTRAN IV era did. For example, the ARM instruction set would utilize condition codes based on comparison operations which could express negative, zero, or positive values. The condition codes would then be examined to determine control flow [**?**].

In 1971, C. Gotlieb and Walker gave an approximate solution to the optimum binary search tree problem [**?**]. Knuth shortly thereafter gave the first optimal solution [**?**]. Knuth's optimal solution requires $O(n^2)$ time and space which is too costly in many situations. Several others have since examined the approximate version of the problem. While unable to bound an approximate algorithm within a constant of the optimal solution, many authors have been able to bound the cost based on the entropy of the distribution of probabilities, H. Specifically,

$$H = \sum_{i=1}^{n} p_i \cdot \lg(\frac{1}{p_i}) + \sum_{j=0}^{n} q_i \cdot \lg(\frac{1}{q_j}).$$

In 1975, P. Bayer showed that

$$H - \lg H - (\lg e - 1) \leq C_{Opt} \leq C_{WB}, C_{MM} \leq H + 2$$

where C_{Opt}, C_{WB}, and C_{MM} are costs for the optimal solution, as well as weight-balanced method of Knuth [**?**] and min-max heuristic of P. Bayer [**?**]. Weight-balanced and min-max cost heuristics are greedy and require both $O(n)$ time and $O(n)$ space to run with the $O(n)$ implementations due to Fredman [**?**]. These greedy heuristics use a top-down approach where the tree root is selected from among the n keys, and we recurse in both the left and right subtrees. Let $P_L(B_i)$ and $P_R(B_i)$ represent the probabilities of searching for a key before or after key B_i respectively. The Weight-balanced approach, makes this greedy root selection by picking the root B_i such that $|P_L(B_i) - P_R(B_i)|$ is minimized. The min-max heuristic selects the root B_i such with minimum $\max(P_L(B_i), P_R(B_i))$. In 1980, Güttler, Mehlhorn and Schneider presented a new heuristic, the modified entropy rule [**?**] which built upon the ideas of Horibe [**?**]. The Entropy Rule greedily selects B_i as the root such that

$$H(P_L(B_i), p_i, P_R(B_i)) = P_L(B_i) \cdot \lg(\frac{1}{P_L(B_i)}) + p_i \cdot \lg(\frac{1}{p_i}) + P_R(B_i) \cdot \lg(\frac{1}{P_R(B_i)})$$

is maximized. As discussed in Chapter **??**, this was modified to improve its performance. Güttler, Mehlhorn and Schneider gave empirical evidence that the modified heuristic outperformed others [**?**]. While the heuristic took $O(n^2)$ time, it only required $O(n)$ space, a huge savings over the optimal solution. However, they were unable to prove that the cost of the modified entropy rule $C_{ME} \leq H + 2$ (unlike previous weight-balanced and min-max heuristics) and settled with $C_{ME} \leq c_1 \cdot H + 2$ where $c_1 = \frac{1}{H(\frac{1}{3}, \frac{2}{3})} \approx 1.08$. We re-examine this method and provide a new bound of $H + 4$ in Chapter **??**. In 1993, De Prisco and De Santis presented a new heuristic for constructing a near-optimum binary search tree [**?**]. The method is discussed in more detail in section **??** and has an upper bounded cost of at most $H + 1 - q_0 - q_n + q_{max}$ where q_{max} is the maximum weight leaf node. This method was later updated by Bose and Douïeb (and is also discussed in section **??**) to have a worst case cost of [**?**]

$$H + 1 - q_0 - q_n + q_{max} - \sum_{i=0}^{m'} pq_{\text{rank}[i]}.$$

Here, $m' = \max(2n - 3P, P) - 1 \geq \frac{n}{2} - 1$ where P is the number of increasing or decreasing sequences in a left-to-right read of the access probabilities of the leaves (gaps) and, $pq_{\text{rank}[i]}$ is the i^{th} smallest access probability among all keys and gaps except q_0 and q_n.

In Chapter **??**, we re-examine the modified entropy rule of Güttler, Mehlhorn and Schneider [**?**]. This is an $\Theta(n^2)$ time, $\Theta(n)$ space, algorithm for approximating the optimum binary search tree problem in the RAM model. The method works very well in practice, and the group had great experimental results, but unfortunately they could not bound the worst case expected cost as well as they would have hoped. While simpler solutions like the *Min-max* of P. Bayer [**?**] and *Weight Balanced* technique of Knuth [**?**] have worst case costs of at most $H + 2$, the trio's modified entropy technique was only shown to have a worst case expected search cost of at most $c \cdot H + 2$ where $c \approx 1.08$ [**?**, **?**]. We provide a new argument of the modified entropy rule's worst case expected search cost and show that it is within an additive factor of entropy: at worst $H + 4$. In Chapter **??**, we move on to external memory models, examining the optimum binary search tree problem under the Hierarchical Memory Model of Aggarwal et al. [**?**]. We provide two algorithms which run in $O(n)$ time and bound their worst case expected costs. We show that the solutions provided both give a direct improvement over a solution of Thite provided under the related HMM_2 model [?]. In Chapter ??, we consider a variant of the optimum binary search tree problem (in the RAM model) where the set of probabilities given are from an unordered multiset. We show that for a multiset with n probabilities, a simple greedy algorithm is within $\frac{n+1}{2n}$ of optimal.

2 An Improved Bound for the Modified Minimum Entropy Heuristic

We show that the Modified Minimum Entropy Heuristic of Güttler, Mehlhorn and Schneider [?] is within an additive factor of entropy: at worst $H + 4$. The previous bound was best upper bound was $c_1 \cdot H + 2$ where $c_1 = \frac{1}{H(\frac{1}{3}, \frac{2}{3})} \approx 1.08$.

Recall equation 1, $H = \sum_{i=1}^{n} p_i \cdot \lg(\frac{1}{p_i}) + \sum_{j=0}^{n} q_i \cdot \lg(\frac{1}{q_j})$. We also use

$$H(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i \cdot \lg(\frac{1}{x_i})$$

to describe the entropy of any probability distribution $(x_1, x_2, ..., x_n)$. For subtree t, we let

$$p_t = \sum_{i:B_i \in t} p_i + \sum_{i:(B_i, B_{i+1}) \in t} q_i$$

be its total probability (the sum of the probability of all nodes within the subtree). $P_L(B_i)$ and $P_R(B_i)$ are probabilities of searching lexicographically before or after (respectively) key B_i. $P_L(B_i, B_{i+1})$ and $P_R(B_i, B_{i+1})$ are probabilities of searching lexicographically before (or equal to) B_i and after (or equal to) B_{i+1} respectively. For a subtree t, $P_L^t(B_i)$ and $P_R^t(B_i)$ describe the normalized probabilities of searching for a key to the left or right of B_i within t, and $P_L^t(B_i, B_{i+1})$ and $P_R^t(B_i, B_{i+1})$ have analogous definitions. We let

$$E_t = H(P_L^t(B_i), \frac{p_i}{p_t}, P_R^t(B_i)) \tag{2}$$

be the local entropy of a subtree t rooted at key B_i.

The entropy rule originally by Horibe [?] for greedy root selection then explain how it was modified in [?]. For a subtree t with probability p_t, the entropy rule greedily chooses the key B_i as the root such that $H(P_L^t(B_i), \frac{p_i}{p_t}, P_R^t(B_i))$ is maximized. While this rule behaves quite well in practice, certain cases cause it to have poor performance (refer to Figure 3.1).

Figure 3.1 demonstrates the shortcomings of the entropy rule heuristic. Given the probability set $\{q_0 = \frac{1}{5}, p_1 = \frac{1}{5}, q_1 = 0, p_2 = 0, q_3 = \frac{3}{5}\}$ the entropy rule will mistakenly choose key B_1 as the root while selecting B_2 as the root produces a better tree. This mistake is remedied in the modified entropy rule of Güttler, Mehlhorn and Schneider [?]. The modified entropy heuristic chooses the root in one of the following three ways:

- If there exists key B_i such that $\frac{p_i}{p_t} > \max(P_L^t(B_i), P_R^t(B_i))$ we always select B_i as the root.
- If there exists a gap (B_i, B_{i+1}) such that $\frac{q_i}{p_t} > max(P_L^t(B_i, B_{i+1}), P_R^t(B_i, B_{i+1}))$ then we select the root from among B_i and B_{i+1}. B_i is chosen if $P_L^t(B_i, B_{i+1}) > P_R^t(B_i, B_{i+1})$ and B_{i+1} is chosen otherwise.
- Otherwise, B_i is selected such that $H(P_L^t(B_i), \frac{p_i}{p_t}, P_R^t(B_i))$ is maximized (as in the original entropy rule).

The approach proposed by Güttler, Mehlhorn and Schneider takes $O(n^2)$ time in the worst case and $O(n)$ space.

▶ **Lemma 1.** *If $x \le \frac{1}{2}$ then $H(x, 1-x) \ge 2x$.*

Proof. We refer the reader to Gallager's 1968 work [?]. ◀

Next, we describe a Lemma which breaks our choice of root in the greedy modified entropy heuristic into one of three cases (not to be confused with the three *rules* used in section **??**).

▶ **Lemma 2.** *When using the modified entropy rule chooses the root B_r of a subtree t with total probability p_t, one of the following three cases must occur:*

Case 1) $E_t \ge 1 - 2\dfrac{p_r}{p_t}$

Case 2) There exists gap(B_i, B_{i+1}) such that $\dfrac{q_i}{p_t} > max(P_L^t(B_i, B_{i+1}), P_R^t(B_i, B_{i+1}))$

Case 3) $\max(P_L^t(B_r), P_R^t(B_r)) < \dfrac{4}{5}$

Proof. At a high level, we first show that *Rule a)* from section **??** implies *Case 1*. We also show that if there exists B_i such that $P_L^t(B_i) \le \frac{1}{2}$ and $P_R^t(B_i) \le \frac{1}{2}$, but cannot apply *Rule a)*, then we still have *Case 1*. Assuming that neither of the two aforementioned conditions occur, we must have that there exists some gap (B_i, B_{i+1}) spanning the middle of the data set. Given this condition, we show that if *Case 2* does not occur (i.e. we cannot use *Rule b)* of section **??**) then *Case 3* must occur, completing the proof.
Rule a) $\implies$ ***Case 1***
First, suppose there exists some p_i such that $\frac{p_i}{p_t} > \max(P_L^t(B_i), P_R^t(B_i))$. By the *Rule a)* of section **??**, it must be selected as the root and thus $r = i$. Moreover, both $P_L^t(B_i)$ and $P_R^t(B_i))$ must be less than one half. Thus, using Lemma 1 we have:

$$\begin{aligned} E_t &\ge H(max(P_L^t(p_i), P_R^t(p_i)), 1 - max(P_L^t(p_i), P_R^t(p_i)) \\ &\ge 2 \cdot \max(P_L^t(p_i), P_R^t(p_i)) \\ &\ge 1 - \frac{p_i}{p_t} \\ &\ge 1 - 2\frac{p_i}{p_t} = 1 - 2\frac{p_r}{p_t} \end{aligned}$$

as required.

B_i ***spans middle*** $\implies$ ***Case 1***
If we do not have some p_i such that $\frac{p_i}{p_t} > \max(P_L^t(B_i), P_R^t(B_i))$ but do have some B_i such that $P_L^t(B_i) \le \frac{1}{2}$ and $P_R^t(B_i) \le \frac{1}{2}$ then we must use *Rule c)* of section **??**. We then must have that:

$$\begin{aligned} &E_t \ge H(P_L^t(B_i), \frac{B_i}{p_t}, P_R^t(B_i)) \text{ and} \\ &0 \le P_L^t(B_i) \le 0.5 \text{ and} \\ &0 \le \frac{p_i}{p_t} \le 0.5 \text{ and} \\ &0 \le P_R^t(B_i) \le 0.5 \end{aligned}$$

then we know that

$$H(P_L^t(p_i), \frac{p_i}{p_t}, P_R^t(p_i)) \ge H(1/2, 1/2) = 1$$

in this case. Thus, combining the above two cases, if have some B_i such that $P^t_L(B_i) \leq \frac{1}{2}$ and $P^t_R(B_i) \leq \frac{1}{2}$ then $E_t \geq 1 - 2p_r$ as required.

$\big($***NOT(****B_i* ***spans mid) AND NOT(Case 1) AND NOT(Case 2)***$\big) \implies$ ***Case 3***
Otherwise, we must have some gap (B_i, B_{i+1}) spanning the middle of the data set (i.e. $P^t_L(B_i, B_{i+1}) < \frac{1}{2}$ and $P^t_R(B_i, B_{i+1}) < \frac{1}{2}$). Suppose that *Case 2* does not occur (i.e. we cannot use *Rule b*): there does not exist a (B_i, B_{i+1}) such that $\frac{q_i}{p_t} > max(P^t_L(B_i, B_{i+1}), P^t_R(B_i, B_{i+1}))$. Then, for any root r of t we have that

$$\max(P^t_L(B_r), P^t_R(B_r)) \geq \min(P^t_L(B_r), P^t_R(B_r))$$
$$\max(P^t_L(B_r), P^t_R(B_r)) \geq q_i$$

Thus, for any root r:

$$\max(P^t_L(p_r), P^t_R(p_r)) \geq 1/3$$

and by our assumption

$$\max(P^t_L(p_r), P^t_R(p_r)) < \frac{1}{2}.$$

So, as in the proof of table 3 (5.3) in [?]

$$E_t \geq H(1/3, 2/3) \approx 0.92.$$

Either *Case 1* occurs, or we have that:

$$E_t < 1 - 2\frac{p_r}{p_t}$$
$$\implies \frac{p_r}{p_t} < \frac{1 - H(\frac{1}{3}, \frac{2}{3})}{2} \approx 0.04.$$

Suppose for contradiction that $\max(P^t_L(p_r), P^t_R(p_r)) \geq \frac{4}{5}p_t$ then we have:

$$E_t \leq H(\frac{4}{5}, \frac{1 - H(\frac{1}{3}, \frac{2}{3})}{2}, \frac{1}{5} - \frac{1 - H(\frac{1}{3}, \frac{2}{3})}{2}) \approx 0.87 < 0.92 \approx H(\frac{1}{3}, \frac{2}{3}) \leq E_t$$

which is a contradiction. Thus, if we do not have *Case 1* or *Case 2* we must have *Case 3* which completes the proof. ◀

Before we examine the main theorem we need a small claim (proven in the appendix TODO).

▶ Claim 1. $H(\frac{1}{2} - \frac{1}{2}x, \frac{1}{2} + \frac{1}{2}x) \geq 1 - \frac{4}{5}x^2$ when $0 < x < \frac{1}{2}$

▶ **Theorem 3.** *Let C_{ME} be the expected cost of search for a tree made by the modified entropy rule. Then $C_{ME} \leq H + 4$*

Proof. This uses a similar style to the proof of Theorem 4.4 in [?]. We bind each E_t for each subtree of our BST on a case by case basis using the cases of Lemma 2.
If *Case 1* occurs, we obviously have that

$$E_t \geq 1 - 2\frac{p_r}{p_t}. \tag{3}$$

Note that this can only happen once for each key (a key can only be root once).

As mentioned in Lemma 2 if some B_i spans in the middle of the data set, $P_L^t(B_i) \leq \frac{1}{2}$ and $P_R^t(B_i) \leq \frac{1}{2}$, we can still show that *Case 1* occurs. Suppose for the remainder of the proof that there is no such middle-spanning B_i.

Let (B_m, B_{m+1}) be the unique middle gap (i.e. $P_L^t(B_m, B_{m+1}) < \frac{1}{2}$ and $P_R^t(B_m, B_{m+1}) < \frac{1}{2}$) when *Case 2* or *Case 3* occurs. When *Case 2* occurs, we know that we could select a root from among the two keys outside of our middle spanning gap, and choose the one which is closer to the middle. Thus, we have that (using Lemma 1):

$$E_t \geq H(\frac{1}{2} - \frac{1}{2}\frac{q_m}{p_t}, \frac{1}{2} + \frac{1}{2}\frac{q_m}{p_t}) \geq 2(\frac{1}{2} - \frac{1}{2}\frac{q_m}{p_t}) = 1 - \frac{q_m}{p_t}. \tag{4}$$

Note that by the definition of the *Rule c)* of section **??**, when this occurs, (B_m, B_{m+1}) must be a leaf of depth at most 2. Thus, this condition can only happen twice for each (B_m, B_{m+1}) gap.

When neither *Case 1* nor *Case 2* occur (and we have a (B_m, B_{m+1}) spanning the middle) we must have *Case 3*. This gives us

$$E_t \geq H(\frac{1}{2} - \frac{1}{2}\frac{q_m}{p_t}, \frac{1}{2} + \frac{1}{2}\frac{q_m}{p_t}).$$

We again apply Claim 2 and get

$$E_t \geq 1 - \frac{4}{5}(\frac{q_m}{p_t})^2. \tag{5}$$

As in [**?**] we define a b_t for each subtree t as follows. We want to have a value for b_t such that $E_t \geq 1 - \frac{b_t}{p_t}$ in all cases. Using *Cases 1,2*, and *3* are their respective equations 3, 4, and 5 we do just that:
Let $b_t = 2 \cdot p_r$ when *Case 1* occurs. B_r is the root of b_t.
Let $b_t = 2 \cdot q_m$ when *Case 2* occurs. (B_m, B_{m+1}) is middle gap of b_t.
Let $b_t = \frac{q_m^2}{p_t}$ when *Case 3* occurs. (B_m, B_{m+1}) is middle gap of b_t.

Note that, in 1975 P. Bayer [**?**] showed that the cost C of our tree could be defined as (**Lemma 2.3**)

$$C = \sum_{t \in S_T} p_t$$

and the entropy could be calculated by

$$H = \sum_{t \in S_T} p_t \cdot E_t$$

where S_T is the set of all subtrees of our tree T.

Thus, by subbing in $E_t \geq 1 - \frac{b_t}{p_t}$ and rearranging we get:

$$H = \sum_{t \in S_T} p_t \cdot E_t \geq \sum_{t \in S_T} p_t - \sum_{t \in S_T} b_t = C - \sum_{t \in S_T} b_t$$
$$\implies C \leq H + \sum_{t \in S_T} b_t$$

As mentioned above, *Case 1* and *Case 2* can only occur once and twice respectively for any potential root B_r or gap (B_m, B_{m+1}). *Case 3* however, can occur many times for a gap (B_m, B_{m+1}). Each time it occurs though, $\frac{q_m}{p_t}$ must increase by a factor of at least $\frac{5}{4}$ since $\max(P_L^t(B_r), P_R^t(p_r)) < \frac{4}{5}$ for the root B_r of the subtree by *Case 3)* of Lemma 2. Moreover, if $\frac{q_m}{p_t} > \frac{1}{2}$ then we will have *Case 2*. Let S_m be the set of all subtrees t for which (B_m, B_{m+1}) is the middle gap and *Case 3* only applies. We have that

$$C \leq H + \sum_{t \in S_T} b_t = H + 2\sum_{r=1}^{n} p_r + 2\sum_{m=0}^{n} q_m + \sum_{m=0}^{n} \sum_{t \in S_m} \frac{4}{5} \frac{q_m^2}{p_t}$$

By factoring out q_m and examining only cases up to $\frac{q_m}{p_t} = \frac{1}{2}$ (since otherwise *Case 2* will occur) we get:

$$\begin{aligned}
C &\leq H + 2 + \sum_{m=0}^{n} (\frac{4}{5} \cdot q_m) \sum_{x=0}^{\infty} \frac{1}{2} \cdot (\frac{4}{5})^x \\
&= H + 2 + \sum_{m=0}^{n} \frac{4}{5} \cdot q_m \cdot \frac{1}{2} \cdot (\frac{1}{1 - \frac{4}{5}}) \text{ (geometric series)} \\
&= H + 2 + 2 \cdot \sum_{m=0}^{n} q_m \\
&\leq H + 4.
\end{aligned}$$

◀

3 APPENDIX - TODO AS REAL APPENDIX

▶ Claim 2. $H(\frac{1}{2} - \frac{1}{2}x, \frac{1}{2} + \frac{1}{2}x) \geq 1 - \frac{4}{5}x^2$ when $0 < x < \frac{1}{2}$

Proof. In order to prove the claim, we find the minimum of

$$H(\frac{1}{2} - \frac{1}{2}x, \frac{1}{2} + \frac{1}{2}x) - (1 - \frac{4}{5}x^2)$$

when $0 < x < \frac{1}{2}$. To do this, we define $F(x)$ and take the derivative with respect to x.

$$\begin{aligned}
F(x) &= H(\frac{1}{2} - \frac{1}{2}x, \frac{1}{2} + \frac{1}{2}x) - (1 - \frac{4}{5}(x)^2) \\
F(x) &= -(\frac{1}{2} - \frac{1}{2}x) \cdot \lg(\frac{1}{2} - \frac{1}{2}x) - (\frac{1}{2} + \frac{1}{2}x) \cdot \lg(\frac{1}{2} + \frac{1}{2}x) - (1 - \frac{4}{5}x^2) \\
\implies F'(x) &= \lg(\frac{1}{2} - \frac{1}{2}x) - \lg(\frac{1}{2} + \frac{1}{2}x) + \frac{8}{5}x \text{ (with some careful manipulation)}
\end{aligned}$$

The only root occurs when $x = 0$. Thus, we check when $x \to 0$ and $x \to \frac{1}{2}$. We note that:

$$F'(x) \xrightarrow{x \to 0} 0^+ \text{ and}$$

$$F'(x) \xrightarrow{x \to \frac{1}{2}} 0.0112781 > 0$$

Thus, $H(\frac{1}{2} - \frac{1}{2}x, \frac{1}{2} + \frac{1}{2}x) - (1 - \frac{4}{5}(x)^2) > 0$ for $0 < x < \frac{1}{2}$ which proves the claim. ◀

4 Typesetting instructions – please read carefully

Please comply with the following instructions when preparing your article for a LIPIcs proceedings volume.

- Use pdflatex and an up-to-date LaTeX system.
- Use further LaTeX packages only if required. Avoid usage of packages like `enumitem`, `enumerate`, `cleverref`. Keep it simple, i.e. use as few additional packages as possible.
- Add custom made macros carefully and only those which are needed in the article (i.e., do not simply add your convolute of macros collected over the years).
- Do not use a different main font. For example, the usage of the `times`-package is forbidden.
- Provide full author names (especially with regard to the first name) in the `\author` macro and in the `\Copyright` macro.
- Fill out the `\subjclass` and `\keywords` macros. For the `\subjclass`, please refer to the ACM classification at `http://www.acm.org/about/class/ccs98-html`.
- Take care of suitable linebreaks and pagebreaks. No overfull `\hboxes` should occur in the warnings log.
- Provide suitable graphics of at least 300dpi (preferrably in pdf format).
- Use the provided sectioning macros: `\section`, `\subsection`, `\subsection*`, `\paragraph`, `\subparagraph*`, ... “Self-made” sectioning commands (for example, `\noindent{\bf My subparagraph.}` will be removed and replaced by standard LIPIcs style sectioning commands.
- Do not alter the spacing of the `lipics-v2016.cls` style file. Such modifications will be removed.
- Do not use conditional structures to include/exclude content. Instead, please provide only the content that should be published – in one file – and nothing else.
- Remove all comments, especially avoid commenting large text blocks and using `\iffalse` ... `\fi` constructions.
- Keep the standard style (`plainurl`) for the bibliography as provided by the `lipics-v2016.cls` style file.
- Use BibTex and provide exactly one BibTex file for your article. The BibTex file should contain only entries that are referenced in the article. Please make sure that there are no errors and warnings with the referenced BibTex entries.
- Use a spellchecker to get rid of typos.
- A manual for the LIPIcs style is available at `http://drops.dagstuhl.de/styles/lipics/lipics-v2016-authors/lipics-v2016-manual.pdf`.

5 Lorem ipsum dolor sit amet

Lorem ipsum dolor sit amet, consectetur adipiscing elit [?]. Praesent convallis orci arcu, eu mollis dolor. Aliquam eleifend suscipit lacinia. Maecenas quam mi, porta ut lacinia sed, convallis ac dui. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Suspendisse potenti. Donec eget odio et magna ullamcorper vehicula ut vitae libero. Maecenas lectus nulla, auctor nec varius ac, ultricies et turpis. Pellentesque id ante erat. In hac habitasse platea dictumst. Curabitur a scelerisque odio. Pellentesque elit risus, posuere quis elementum at, pellentesque ut diam. Quisque aliquam libero id mi imperdiet quis convallis turpis eleifend.

▶ **Lemma 4** (Lorem ipsum). *Vestibulum sodales dolor et dui cursus iaculis. Nullam ullamcorper purus vel turpis lobortis eu tempus lorem semper. Proin facilisis gravida rutrum.*

Listing 1 Useless code

```
for i:=maxint to 0 do
begin
    j:=square(root(i));
end;
```

Etiam sed sollicitudin lorem. Proin pellentesque risus at elit hendrerit pharetra. Integer at turpis varius libero rhoncus fermentum vitae vitae metus.

Proof. Cras purus lorem, pulvinar et fermentum sagittis, suscipit quis magna. ◀

▶ **Theorem 5** (Curabitur pulvinar, [?]). *Nam liber tempor cum soluta nobis eleifend option congue nihil imperdiet doming id quod mazim placerat facer possim assum. Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.*

5.1 Curabitur dictum felis id sapien

Curabitur dictum felis id sapien mollis ut venenatis tortor feugiat. Curabitur sed velit diam. Integer aliquam, nunc ac egestas lacinia, nibh est vehicula nibh, ac auctor velit tellus non arcu. Vestibulum lacinia ipsum vitae nisi ultrices eget gravida turpis laoreet. Duis rutrum dapibus ornare. Nulla vehicula vulputate iaculis. Proin a consequat neque. Donec ut rutrum urna. Morbi scelerisque turpis sed elit sagittis eu scelerisque quam condimentum. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nec faucibus leo. Cras ut nisl odio, non tincidunt lorem. Integer purus ligula, venenatis et convallis lacinia, scelerisque at erat. Fusce risus libero, convallis at fermentum in, dignissim sed sem. Ut dapibus orci vitae nisl viverra nec adipiscing tortor condimentum [?]. Donec non suscipit lorem. Nam sit amet enim vitae nisl accumsan pretium.

5.2 Proin ac fermentum augue

Proin ac fermentum augue. Nullam bibendum enim sollicitudin tellus egestas lacinia euismod orci mollis. Nulla facilisi. Vivamus volutpat venenatis sapien, vitae feugiat arcu fringilla ac. Mauris sapien tortor, sagittis eget auctor at, vulputate pharetra magna. Sed congue, dui nec vulputate convallis, sem nunc adipiscing dui, vel venenatis mauris sem in dui. Praesent a pretium quam. Mauris non mauris sit amet eros rutrum aliquam id ut sapien. Nulla aliquet fringilla sagittis. Pellentesque eu metus posuere nunc tincidunt dignissim in tempor dolor. Nulla cursus aliquet enim. Cras sapien risus, accumsan eu cursus ut, commodo vel velit. Praesent aliquet consectetur ligula, vitae iaculis ligula interdum vel. Integer faucibus faucibus felis.

- Ut vitae diam augue.
- Integer lacus ante, pellentesque sed sollicitudin et, pulvinar adipiscing sem.
- Maecenas facilisis, leo quis tincidunt egestas, magna ipsum condimentum orci, vitae facilisis nibh turpis et elit.

6 Pellentesque quis tortor

Nec urna malesuada sollicitudin. Nulla facilisi. Vivamus aliquam tempus ligula eget ornare. Praesent eget magna ut turpis mattis cursus. Aliquam vel condimentum orci. Nunc congue, libero in gravida convallis [**?**], orci nibh sodales quam, id egestas felis mi nec nisi. Suspendisse tincidunt, est ac vestibulum posuere, justo odio bibendum urna, rutrum bibendum dolor sem nec tellus.

▶ **Lemma 6** (Quisque blandit tempus nunc)**.** *Sed interdum nisl pretium non. Mauris sodales consequat risus vel consectetur. Aliquam erat volutpat. Nunc sed sapien ligula. Proin faucibus sapien luctus nisl feugiat convallis faucibus elit cursus. Nunc vestibulum nunc ac massa pretium pharetra. Nulla facilisis turpis id augue venenatis blandit. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.*

Fusce eu leo nisi. Cras eget orci neque, eleifend dapibus felis. Duis et leo dui. Nam vulputate, velit et laoreet porttitor, quam arcu facilisis dui, sed malesuada risus massa sit amet neque.

Acknowledgements. I want to thank ...

A Morbi eros magna

Morbi eros magna, vestibulum non posuere non, porta eu quam. Maecenas vitae orci risus, eget imperdiet mauris. Donec massa mauris, pellentesque vel lobortis eu, molestie ac turpis. Sed condimentum convallis dolor, a dignissim est ultrices eu. Donec consectetur volutpat eros, et ornare dui ultricies id. Vivamus eu augue eget dolor euismod ultrices et sit amet nisi. Vivamus malesuada leo ac leo ullamcorper tempor. Donec justo mi, tempor vitae aliquet non, faucibus eu lacus. Donec dictum gravida neque, non porta turpis imperdiet eget. Curabitur quis euismod ligula.