同じ回数だけ比べても、遅い記憶を読む回数は木の分岐で決まる

— 要素 10^9 個で、二分探索の比較は **29.897353** 回、分岐 **100** の木も深さ **4.500000** で比較は **29.897353** 回と同じ — 分離子は、数えているのが比較の回数か、ページの読み出しの回数か — [論文759]

山岸 勇輝 (Yuuki YAMAGISHI)
Yamagishi Complexity Torus Lab, Japan｜山岸勇輝ラボ
2026年9月

— From thinking to syncing. —

要旨

二分探索も木の索引も、探す手間は要素の数の対数で増える。だから同じだと言われやすい。比べる回数と、遅い記憶を読む回数を分けて数える。新しい定理も法則も主張しない。 第一に、比べる回数は、木の分岐に依らず同じである——要素 10^9 個で、二分探索も分岐 100 の木も **29.897353** 回（第2節）。 第二に、遅い記憶を読む回数は、大きく違う——一ページ 100 個で、二分探索 **24.253497**、木 **4.500000**（第2節）。 第三に、これが本稿の芯である。読む回数の比は、対数の底の比に近づく——比は 10^9 個で **5.389666**、10^{18} 個で **6.016761** と増え、$\log_2 100 = $ **6.643856** を超えない（第3節）。 第四に、小さい表では差が小さい——10^3 個で比は **2.881285**（第3節）。 第五に、試行でも同じ向きに出る——要素 10^6 個・2000 回で、読むページは二分探索 **13.212000**、木 **3.000000**（第3節）。 第六に、分離子は、数えているのが比較の回数か、ページの読み出しの回数かである——「どちらも対数」は、何を数えるかを言うまで速さを語れない（第4節）。 キーワード：二分探索、B木、索引、対数の底、ページの読み出し、記憶の階層。

本稿の射程（射程注記）：新しい定理も法則も主張しない——二分探索の比較回数が $\log_2 n$ であること、分岐 B の木の深さが $\log_B n$ であること、外部記憶の索引に分岐の大きい木を使うことは、いずれも既知である（コルメンほか 2009、バイヤー＆マクレイト 1972）。実時間を測らない——読み出しの回数だけを数え、一回の読み出しにかかる時間や装置は扱わない。木の手入れを扱わない——挿入、削除、節の分割、空きの割合は扱わず、満ちた静的な木だけを見る。ページの大きさを一つに置く——一ページ 100 個は比べるための値である。数え方を分ける——試行の二分探索では、等しいかと大小の二回を比較として数えたので、比較の回数は閉じた式と並べない。読み出しの回数だけを並べる。既刊との関係：論文702 は、分けても速くならない場合があり、決めているのは分け方ではなくまとめる仕事であることを示した——その射程注記は「メモリ階層、キャッシュ、並列化の効きは入れない」と明記していた。そこでは漸近の指数どうしが競った。ここでは同じ $O(\log n)$ の中で、遅い記憶を読む回数が対数の底を効かせる。論文705 も引く——そこでは当たり率 0.99 でも、重い外れの一回が時間の半分以上を食った。ここでは、その重い一回をどれだけ減らせるかを木の分岐が決める。加えたのは、比較の回数が底に依らないことと、読む回数が底で変わることを同じ表に並べたこと、読む回数の比が要素とともに $\log_2 100 = 6.643856$ に近づくことを示したこと、試行で読み出しを数えたこと、分離子を「比較か、読み出しか」に置いたことである。

1. 記法と問い

設定：要素 n 個の整列した表を探す。二分探索の比較は $\log_2 n$ 回、分岐 B の木は深さ $\log_B n$ で、節の中を二分探索すれば比較は同じ $\log_2 n$ 回になる。遅い記憶のページに B 個ずつ入るとき、読み出しの回数は次になる。

$$C_{\text{bin}} = \log_2 n = \log_B n \cdot \log_2 B, \qquad P_{\text{bin}} = \log_2 \frac{n}{B} + 1, \qquad P_{\text{tree}} = \log_B n, \qquad \frac{P_{\text{bin}}}{P_{\text{tree}}} \to \log_2 B$$

$B = 100$ と置く。

問い：同じ対数の手間なら、二分探索と木の索引は同じ速さなのか。

2. 比べる回数は同じ、読む回数は違う

計算 2.1：要素 10^9 個で、比較の回数と読み出しの回数を出した。

量	二分探索	分岐 100 の木
比較の回数	**29.897353**	**29.897353**
遅い記憶を読む回数	**24.253497**	**4.500000**

> **観察 2.2**
>
> 比べる回数は、木の分岐に依らず同じである。木の比較は $\log_2 n$ と差 3.553×10^{-15} で一致し、線形の平均 5×10^{8} 回の 1.672389×10^{7} 分の 1 である——比較で数えれば、二つの手間に差は無い。

> **観察 2.3**
>
> 遅い記憶を読む回数は、大きく違う。二分探索は 24.253497 回、木は 4.500000 回で、比は 5.389666 である——二分探索は、最後の一ページに入るまで毎回ちがうページを読む。

3. 読む回数の比は、底の比に近づく

計算 3.1：要素の数を変えて、読む回数の比を出した。

要素	読む回数の比（二分探索／木）
10^{3}	2.881285
10^{6}	4.762571
10^{9}	5.389666
10^{12}	5.703213
10^{18}	6.016761
行き着く先 $\log_2 100$	6.643856

> **観察 3.2**
>
> これが本稿の芯である。読む回数の比は、要素とともに増えて対数の底の比に近づく。10^{3} 個の 2.881285 から 10^{18} 個の 6.016761 まで増え、6.643856 を超えない——同じ $O(\log n)$ でも、読み出しで数えると底の違いが定数倍として残る。

計算 3.3：要素 10^{6} 個の表で、無作為に選んだ 2000 個を探し、二分探索が触れたページの数と、木が読んだ節の数を数えた。

量	平均
二分探索が読んだページ	**13.212000**
木が読んだ節	**3.000000**

> **観察 3.4**
>
> 試行でも、木の読み出しは二分探索のページより少ない。木は毎回ちょうど **3** 節を読み、二分探索は平均 **13.212000** ページに触れた——比較の回数は数え方で違って見えるので並べず、読み出しだけを並べた。

4. 分離子

> **観察 4.1**
>
> 分離子は、数えているのが比較の回数か、ページの読み出しの回数かである。比較なら底は消え、読み出しなら底の比 **6.643856** に向かう差が残る——「どちらも対数」は、何を数えるかを言うまで速さを語れない。

> **観察 4.2**
>
> 論文**702** と並べる。**702** は、分割統治の速さを漸近の指数の比べ合いで決め、メモリ階層は入れないと明記した——ここは指数が同じ場合で、入れなかった機械の事情が底を効かせる。

> **観察 4.3**
>
> 論文**705** とも並べる。**705** は、当たり率 **0.99** でも重い外れが時間の半分以上を食うことを示した——遅い記憶を読む一回がその重い外れで、木の分岐はその回数を減らす。

> **観察 4.4**
>
> 言わないこと。実時間も、木の手入れも、キャッシュの細部も扱わない——示したのは、比べる回数と読む回数が別の数で、底が効くのは後者だけだということである。

5. 点呼と結び

1. 比較 **29.897353** 回（計算 2.1）：二分探索と木で同じ。
2. 二分探索の読み出し **24.253497**（計算 2.1）：毎回ちがうページ。
3. 木の読み出し **4.500000**（計算 2.1）：深さだけ。
4. 比 **5.389666**（観察 2.3）：10^9 個で。
5. 10^3 個で **2.881285**（計算 3.1）：小さい表。
6. 10^{18} 個で **6.016761**（計算 3.1）：増える。
7. 行き着く先 **6.643856**（観察 3.2）：本稿の芯。
8. 試行で **13.212000** ページ（計算 3.3）：二分探索。
9. **3.000000** 節（計算 3.3）：木。
10. 比較か、読み出しか（観察 4.1）：分離子・論文702・705。

結び：同じ回数だけ比べても、遅い記憶を読む回数は木の分岐で決まる。要素 10^9 個で、二分探索も分岐 100 の木も比較は 29.897353 回で同じだが、一ページ 100 個として読み出しは 24.253497 回と 4.500000 回で、比は 5.389666 になる——二分探索は最後の一ページに入るまで毎回ちがうページを読む。読む回数の比は要素とともに 2.881285（10^3 個）から 6.016761（10^{18} 個）まで増え、$\log_2 100 = 6.643856$ に近づく——同じ $O(\log n)$ でも、読み出しで数えると底の違いが定数倍として残る。要素 10^6 個の試行でも、読んだページは二分探索 13.212000、木 3.000000 だった。分離子は、数えているのが比較の回数か、ページの読み出しの回数かである——「どちらも対数」は、何を数えるかを言うまで速さを語れない。既刊との位置——論文702 は指数を比べ、メモリ階層を入れないと明記した。ここは指数が同じ場合に、その機械の事情が底を効かせることを数えた。論文705 の重い外れを減らすのが木の分岐である。正直に言えば——実時間も木の手入れもキャッシュの細部も扱っていない。

参考文献

1. Bayer, R., & McCreight, E. (1972). Organization and maintenance of large ordered indexes. *Acta Informatica*.
2. Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms* (3rd ed.). MIT Press.

3. 山岸勇輝 (2026).『分けても速くならない場合がある——決めているのは、分け方ではなく、まとめる仕事である』[論文702]. Zenodo. 10.5281/zenodo.23014909

4. 山岸勇輝 (2026).『当たり率 **0.99** でも、時間の半分以上は外れが使っている』[論文705]. Zenodo. 10.5281/zenodo.23014923

— From thinking to syncing. —

山岸 勇輝 (Yuuki YAMAGISHI)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)