都市（City）　解説

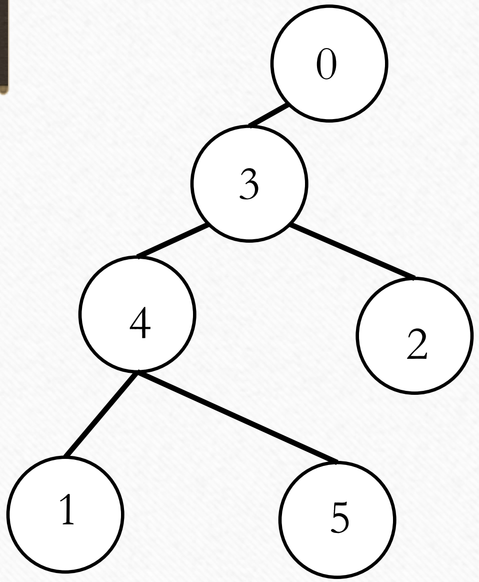
松崎照央

問題概要

- 頂点0が根の根付き木がもらえる(大きい)
- 木のある2つのノードの組について「その組は祖先-子孫か子孫-祖先かどちらでもないか」を調べる。
- **木の深さは18まで** (30点までは使いません)
- 頂点を符号化し、クエリの対象となる頂点の符号のみからクエリに答える。

例

- 各ノードを「自分の番号,頂点0が子孫かどうか,頂点1が子孫かどうか,頂点2が子孫かどうか,...」というように符号化してみる($N \leq 10$)



0: 0000 011111
1: 0001 000000
2: 0010 000000
3: 0011 011011
4: 0100 010001
5: 0101 000000

A 0101 000000
B 0011 011011
こたえ:0

本番で解く前に

- 符号化して復元するタイプの問題でした
- こういった形式の問題はなれてないとつらいかも・・・
- 過去問の似た形式の問題をやっておこう！

小課題1　8点

- $N \leq 10,$ 符号の大きさ $\leq 2^{60} - 1$　つまり60ビット使える。
- 頂点が少ないし、送れる情報もいっぱい
- なんでもできるわけではないが、手法は様々でやさしい。

小課題1　実装例

- 自分の番号(4bit)と辺を隣接行列(45bit)で送る(ありうる辺の種類は10C2通りなので)
- 自分の番号(4bit)と、各頂点が祖先かどうか(10bit)、子孫かどうか(10bit)をすべて保持する
- 祖先か子孫片方だけでもOK
- 祖先頂点の番号全列挙(36bit)するのもOK

ここまで解けた人

- 8点　　　5人
- 8点超え 7人

小課題2

- N ≤250000 (大変そう)
- 送る整数の最大値Lで得点が決まる。
- $L \geq 2^{38}$で0点(かなしい)
- $L \leq 2^{28}-1$で満点(うれしい)
- 点数はLの大きさで徐々に変わっていく。

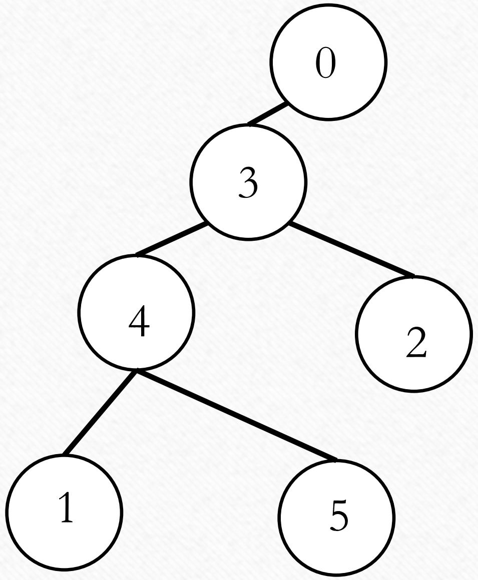
小課題2 (14点) or (10点)

- $L \leq 2^{36} - 1$　つまり36bit使える。
- 頂点も25万あるので、各頂点に対して何か持ったり全列挙などの解法は厳しそう
- 番号を1つ表すのに18bit必要なので、番号はたかだか2つだけしか表すことができない。(2つなら番号を持てる)

小課題2 Euler Tour

- Euler Tourで子孫を区間で扱えばよい。(頻出テク)
- Euler Tourとは？

Euler Tour



- 部分木を区間の左端と右端のindex2つで扱えるようになる(詳しくはググって)
- 0の部分木(0-5)
- 1の部分木(3-3)
- 2の部分木(5-5)
- 3の部分木(1-5)
- 4の部分木(2-4)
- 5の部分木(4-4)　　本来の番号はもはや不要

0	**3**	**4**	**1**	5	**2**

小課題2 Euler Tour

- Euler Tourをして、各頂点に対して(部分木の区間の左端、右端）または（部分木の区間の左端、長さ）などを持てばよい
- 普通にやれば14点、帰りがけでも加算をしてしまうと10点

0	3	4	1	5	2

14点ツアー

0	3	4	1	1	5	5	4	2	2	3	0

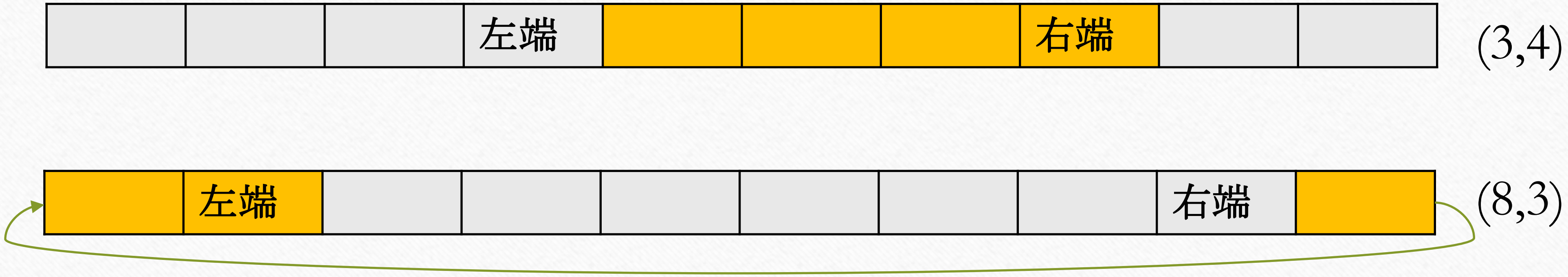
10点ツアー

小課題2 (22点)

- $L \leq 2^{35} - 1$ 1ビット節約で追加の8点を目指そう。
- 少し考察
- 区間の左の頂点は取りうる値が広いが、右は狭い
- 1番目の頂点　右端: $1 \sim N$ 長さ: $0 \sim (N-2)$
- N-1番目の頂点　右端: $(N-1)$ 長さ: 0
- 結構アンバランスで無駄がありそう

小課題2 (22点)

- 区間を円環(mod N)と見て、(区間の一端、端同士の差)を持つ(ただし差が短い方をもつ)



小課題2 (22点)

- 端同士の差は少なくとも一方は$2^{17}-1$以下なので1bit節約($2^{17}+2^{17}=2^{18}>250000$なので)
- おまけの8点ゲット
- しかし、この発想ではこれ以上の改善はできなさそう

- 他の方法もあるかも？

ここまでの得点分布

- 8点 5人
- 18点　0人
- 22点　4人
- 30点　3人
- 30点超 ???人

さらなる高みへ

- $L \leq 2^{28} - 1$(もしくは$2^{28} \leq L \leq 2^{34} - 1$)
- ここからが本番です
- まずは考察をしましょう

木の深さについて

- そういえば「木の深さは高々18以下」という制約がありました。
- これ何の意味があるんでしょう？

- 各頂点の祖先は高々18個（子孫はたくさんありうる）
- つまり祖先－子孫の関係は今回は高々18*N通り(O(N))
- ほかにも何かあるかも・・・（なくても解けます）

Euler Tourの方針は改善できるのか？

- まず左端は全通り存在する。
- (部分木の区間の左端、右端)という持ち方は厳しそう。(葉は全部右端になりうるので)
- (部分木の区間の左端、長さ)という持ち方はどうだろう・・・？

Euler Tourの区間の長さについて

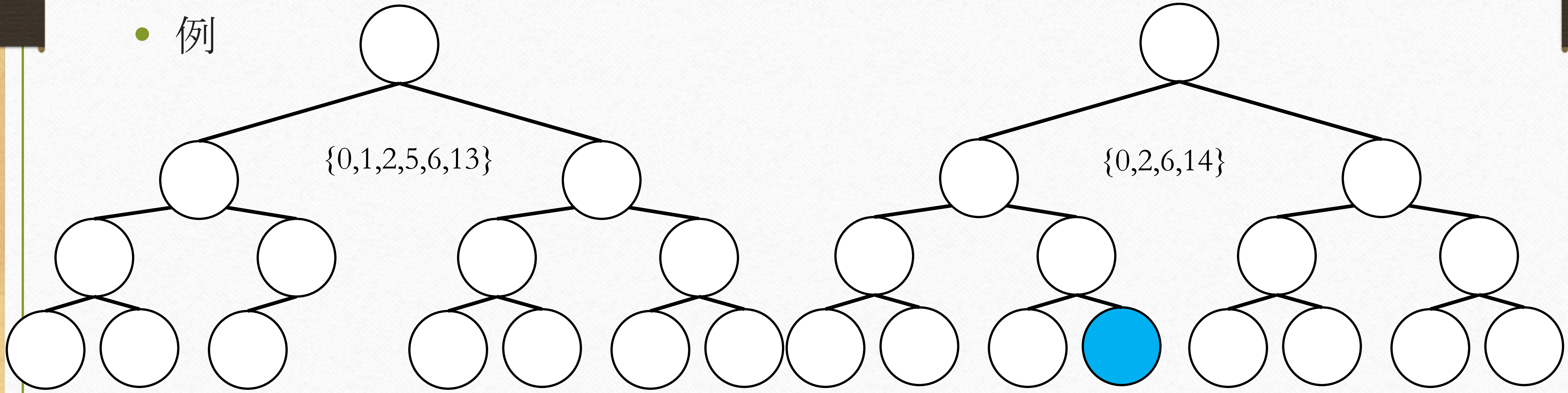
- Euler Tourの区間の長さは頂点の子孫の数である(とする)
- 祖先ー子孫の関係は高々18*N通り(O(N))　(再掲)
- →各頂点の子孫の数の総和は18*N以下になる。
- →各頂点の子孫の数は平均すると18以下になる。

Euler Tourの区間の長さについて

- 区間の長さはたいてい短い
- なのに18ビットも使うのはもったいなくない？
- でもどんな長さがあるかはわからないからなあ
- 区間の長さを調整できればなあ

ダミーの頂点

- ダミーの頂点を挿入する(区間に空白を挿入する)ことで、区間の長さを変えることができる(取りうる値を制限することができる。)
- 例



ちょっとした注意

- (ソースコードでは区間の長さを変えるだけなので安心して)
- 区間の左端の番号の最大値がNより大きくなることに注意(ダミーの頂点があるので)

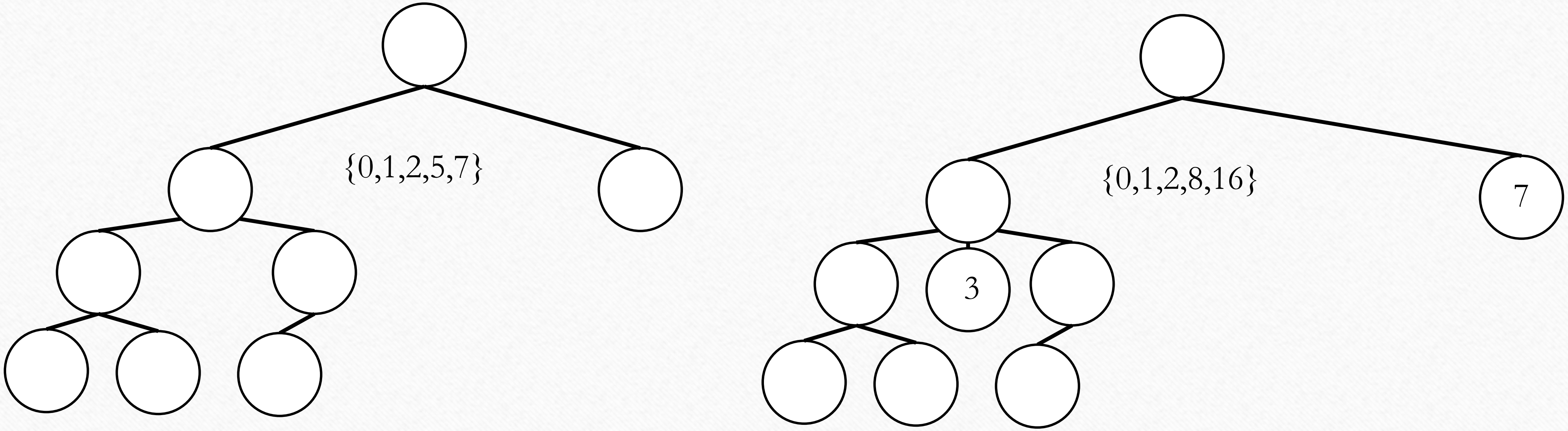
区間の長さを制限する

- 区間の長さとして取りうる値の数列A_nを考える
- 区間の長さが取る値は、小さいほど頻繁に現れる
- 値が小さいときほど密、大きいときほど疎な数列とかが良さそう
- 等比数列とかがいいのでは？（0も必要ですが）
- $\mathrm{A_n} = 0, \mathrm{r}^0, r^1, r^2, r^3, r^4, \ldots$(数列を整数に丸めればrは実数でもよい）
- 他の数列でもよいかもしれませんが、ここでは等比数列を考えます

区間の長さを制限する

- 試しに区間の長さが取りうる値を $r = 2$に制限してみる。(1,2,4,8,…)
- Lはどのくらいの大きさになるか？

ダミーの頂点の数は？



左端の番号の最大値は？

- 調整前の区間の長さsは $(A_n < s \leq A_{n+1} = \mathrm{A_n r})$である。
- → 1回の調整で、もとの長さのたかだかr倍にしかならない
- →部分木の深さがdである頂点の区間の長さは、たかだか実際の長さのr^d倍になる

左端の番号の最大値は？

- 左端の番号は最大Nr^dぐらいになる
- 区間の長さは$\log_r(Nr^d)$ 通りぐらい
- 全体でLはだいたい$Nr^d \log_r Nr^d = Nr^d (\log_r N + d)$

- r=2では区間の左端が取りうる値が大きくなりすぎてかなしい

満点へ

- 実験したりプログラム書いて計算するとr=1.05ぐらいで満点がとれることがわかります。(rの調整に失敗するとlogで計算される部分点になります)
- ここで注意、何も考えないと数列が1,1,1,1, ...となってしまう(通るかも)
- 数列を$a_1 = 1, a_{n+1} = \max\{a_n * r, a_{n+1} + 1\}$とすると解決。満点へ。
- 一見$a_{n+1} = a_n + 1$であるときにダミーの頂点の数の考察が成り立たない気がするが、a_{n+1}ならダミー頂点の子は1つもないので大丈夫

得点分布

- 8点 5人
- 18点　0人
- 22点　4人
- 30点　3人
- 30点超 0人