Open Problems Column

(ADDED LATER TO THIS E-VERSION:

The problem I talk about in this problem as open was already solved.

For information on this goto my blog `http://blog.computationalcomplexity.org/` and search for posts that have the phrase *David Harris*.

)

Edited by William Gasarch

This Issues Column! This issue's Open Problem Column is by William Gasarch and is *A Problem on linked sequences Inspired by an Oliver Roeder Column*

Request for Columns! I invite any reader who has knowledge of some area to contact me and arrange to write a column about open problems in that area. That area can be (1) broad or narrow or anywhere inbetween, and (2) really important or really unimportant or anywhere inbetween.

1 Oliver Roeder's Column

Nate Silver is a pollster who runs a website [4] which is mostly about politics and sports (sometimes its hard to tell the difference). He also has on it a problems column called *The Riddler*, edited by Oliver Roeder.

In the July 28, 2017 Riddler Column [3], the following question was posed (I paraphrase)

Def 1.1 Let $n \in \mathsf{N}$. A sequence is n-linked if

1. Every element in the sequence is in $\{1, \ldots, n\}$.
2. No element appears more than once in the sequence.
3. Every element is either a factor or multiple of the previous element (except the first element which has no previous element).

Example: The following is a 100-linked of length 17.

17, 1, 47, 94, 2, 4, 20, 5, 10, 90, 3, 21, 7, 42, 6, 30, 15

Problem: Find the longest 100-linked sequence.

2 My Solution

After much work and doodling I found a 100-linked sequence of length 42. I present it as a sequence of sequences that are either increasing (each number is a multiple of the prior one) or decreasing (each number is a factor of the prior one):

88 44 22 11
33 66
6 84
7 14 70
35 5
15 30 90
10 60
20 40 80
2 28
4 56
8 16 32 96
3 9 18 36
12 72
24 48
16 1
13 26 52

Is 42 the best you can do? Fans of *The Hitchhiker's Guide to the Galaxy* would say *of course!*. Alas, it is not the answer. The next section reveals what *The Riddler* said.

3 The Riddler's Answer

In the August 4 2017 Riddler Column [2] the answer was revealed! There is a 100-linked sequence of length 77:

93 31 62 1
87 29
58 2
92 46 23
69 3
57 19
38 76
4 68
34 17
85 5
35 70
10 100
50 25
75 15
45 90
30 60
20 40 80

16 64
32 96
48 24 12 6
78 26
52 13
91 7
49 98
14 56
28 84
42 21
63 9
81 27
54 18
36 72
8 88
44 22
66 33
99 11
55

They also showed another sequence of length 77. How were they found? Are they optimal? They were found by a computer program. The submitters claimed that the program proved it was optimal. Since two different people got it we are inclined to believe them; however, that is not a human-readable proof.

Open Problem 3.1 Find a human-readable proof that there is no 100-linked sequence of length 78.

4 Open Problems

Open Problem 4.1 Let $k(n)$ be the length of the longest linked n-sequence. Get upper and lower bounds on $k(n)$ asymptotically.

Convention 4.2 We will assume $k(n) = \Theta(n^a)$ for some $0 < a \leq 1$ while asking our later open questions. If this is incorrect we will need to modify some of the open questions.

We now ask some complexity questions.

LINKED-SEQ: Given n find a longest n-linked sequence.

We give a naive algorithm for LINKED-SEQ.

1. Input(n)
2. maxlength = 0, maxseq = λ (the empty string)
3. (SEQ_n is the set of all sequences of length $\leq n$ from $\{1, \ldots, n\}$.)
   For $s \in \text{SEQ}_n$
   If s is linked and $|s| >$ maxlength then maxlength $\leftarrow |s|$ and maxseq $\leftarrow s$.

4. Output maxlength and maxseq.

This algorithm takes time roughly $O(|S|)$ which is $O(n \times n!)$.

Open Problem 4.3 Is there an algorithm for LINKED-SEQ that runs in time polynomial in n? (Note that we said polynomial in n, not in $\lg n$ since we are assuming that the sequence is of length $\Theta(n^a)$, hence just outputing it takes a long time.)

We now consider how to phrase the complexity in terms of NP-completeness. Let

$$\text{LINKED} = \{1^{n,k} : \text{ there is an } n\text{-linked sequence of length } k\}$$

(We use $1^{n,k}$ since we are assuming that the interesting cases are when $k = \Theta(n^a)$ and hence the witness is of size $\Theta(n^a \log n)$.)

Def 4.4 A is a *tally set* if $A \subseteq 1^*$.

LINKED is clearly in NP. However, since it is a tally set it is not NP-complete unless P=NP (we reconsider this issue in Section 5). So we look at:

$$\text{LINKEDEXT} = \{(1^{n,k}, x_1, \ldots, x_m) : x_1, \ldots, x_k \text{ can be extended to an } n\text{-linked sequence of length } k\ \}$$

Open Problem 4.5 Is LINKEDEXT NP-complete? Either prove it or show that if it is then the polynomial hierarchy collapses, or some other unlikely consequence.

Since the function LINKEDSEQ outputs the sequence itself, which we are assuming is length $\Theta(n^a)$, the complexity has to be measured as a function of n, possibly polynomial. If all we want is the *length* of the longest n-linked sequence the complexity can be measured as a function of $\lg n$, possibly polynomial.

LINKED-LENGTH:

Given n find the length of the longest n-linked sequence.

Open Problem 4.6 Is there an algorithm for LINKED-LENGTH that runs in time polynomial in $\lg n$? Polynomial in n?

LINKED-COUNT problem: Given n find the number of n-linked sequences of max length.

Open Problem 4.7 Is there an algorithm for LINKED-COUNT that runs in time polynomial in $\lg n$? Polynomial in n? The $\lg n$ question only makes sense if LINKED-COUNT is not that large.

We now generalize the problem further.

Def 4.8 Let $A \subseteq \mathsf{N}$. A sequence is A-linked if

1. Every element in the sequence is in A.

2. No element appears more than once in the sequence.

3. Every element is either a factor or multiple of the previous element (except the first element which has no previous element).

All of the above problems can be asked with input a set A of naturals and asking about A-linked sequences. We restate one of them in full to make a point.

$$\text{LINKEDA} = \{(A, 1^k) : \text{ there is a } A\text{-linked sequence of length } k \}$$

(A is coded by a set of numbers in unary.)

LINKED is clearly in NP. Since A can be anything we find it quite plausible that LINKEDA is NP-complete.

Open Problem 4.9 Is LINKEDA NP-complete?

5 The Complexity of Tally Sets

Recall that we briefly considered the complexity of:

$$\text{LINKED} = \{1^{n,k} : \text{ there is an } n\text{-linked sequence of length } k\}.$$

Our thought was that to show that LINKEDSEQ is probably not in P we should show that LINKED is NP-complete. But, alas, LINKED is a tally set and hence can't be NP-complete unless P=NP.

Def 5.1

1. TALLY is the set of all tally sets.
2. A is NP $\cap$TALLY-complete if $A \in$ NP, $A \in$ TALLY and, for all $B \in$ NP $\cap$TALLY, $B \le_m^p A$.

Let $M_1, M_2, \ldots$ be an enumeration of NP-machines. Buhrman et al. [1] observed that the following set is NP $\cap$TALLY-complete:

$$\text{COMP} = \{1^{i,n,t} : M_i(1^n) \text{ accepts on some path in } t \text{ steps}\}.$$

One way to show that a set in $A \in$ NP $\cap$TALLY is probably difficult is to show A is NP $\cap$TALLY-complete. One way to do that is to show that COMP $\le_m^p A$. However:

1. COMP is the only NP $\cap$TALLY-complete set that we know of.
2. A contrast: (1) the thousands of NP-complete problems have been worked on for a long time, without any poly time algorithm, so we believe P $\neq$ NP, but (2) there has been very little study of NP $\cap$TALLY-completeness. So COMP might not even be hard.

A similar notion, NP $\cap$SPARSE-completeness, has been studied by Buhrman et al. [1].

6 Acknowledgments

I would like to thank Clyde Kruskal for proofreading and discussion. I would like to thank Oliver Roeder for his column that inspired this column.

References

[1] H. Buhrman, S. Fenner, L. Fortnow, and D. van Melkebeek. Optimal proof systems and sparse sets. In *Seventeenth International Symposium on Theoretical Aspects of Computer Science: Proceedings of STACS 2000,* Lille, France, Lecture Notes in Computer Science, New York, Heidelberg, Berlin, 2000. Springer-Verlag.

[2] O. Roeder. Is this bathroom occupied. *Fivethirtyeight.com*, 2017. `https://fivethirtyeight.com/features/is-this-bathroom-occupied/`.

[3] O. Roeder. Pick a number any number. *Fivethirtyeight.com*, 2017. `https://fivethirtyeight.com/features/pick-a-number-any-number/`.

[4] N. Silver. Fivethirtyeight.com, 2008-. `https://fivethirtyeight.com/`.