

山梨での研究活動

カップル ドミニク
大学院総合研究部工学域
dkppl@yamanashi.ac.jp

1 はじめに

昨年の9月に山梨大学に着任した。それ以前は、東京医科歯科大学で勤務していた。甲府市は壮大な山々に囲まれた山間部に位置しており、果樹栽培が盛んなことで知られ、自然環境に恵まれた雰囲気がある。特に東京のような大都会とは対照的だ。その大きな違いに従って、研究環境も大きく変わった。ここでは、私の職場環境の変化について説明したいと思う。

1.1 生活の変化

甲府での生活について紹介したい。まずは山梨大学周辺に部屋を借りる場合、家賃を低く抑えることができる。駅周辺には飲食店が多いので、困ることがない。甲府は自然に囲まれているので、ハイキング・山登りをして、展望台から街の眺めを楽しむことができる。

職場について言えば、研究室は十分な広さがある。よく整備された会議室を予約してディスカッションを行うことができる。天気が良ければ、様々な場所から富士山の素晴らしい眺めが楽しめる(図1)。構内に二つの食堂がある。食堂のテラスは木陰にあり、屋外での食事はリフレッシュできる。食堂の代わりに、キャンパスの近くには山梨名物のほうとう屋さんなどもある。昼食時に東京のように待ち時間がほとんどなく、混雑することはないので快適である。

1.2 山梨へのお誘い 🍷

私の所属する山梨大学甲府キャンパスは、甲府駅からバスで約10分の距離にある。徒歩でも20分ほどの場所にあり、アクセスが良い。そして、研究環境が整っているので、一緒にここで議論する機会を設け、共同研究活動をともに進めたいと思う。研究者の皆様の訪問を歓迎する。是非山梨に訪れてみてほしい。

また、研究以外でも山梨を十分に満喫することができる。甲府駅周辺には有名な温泉地である石和温泉や湯村温泉があり、ほうとうや鳥もつ煮などの名物料理も楽しむことができる。さらに、ワイナリーでテイスティングも可能である。先日、複数の研究者は(共同研究のために)山梨大学で2日間の打ち合わせを行った。



図 1: 山梨県内の観光スポット。上左：昇仙峡，上右：新倉山浅間公園，下：河口湖

2 本研究

次に、最近の研究活動について報告したいと思う。文字列問題の ASP による符号化についての内容である。2 年前から AFSA B01 班の公募研究を通じて、様々な B01 班の打ち合わせで、同じ班に所属している名古屋大学の番原先生と親しくさせていただき、共同研究を始めた。番原先生は命題論理の充足可能性判定問題 (SAT 問題) のソルバーを研究している。その専門の中で、特に解集合プログラミング (ASP) に興味があり、ASP のソルバーを実際に提供している。もちろん、問題を ASP で符号化することにも詳しい。一方で、私は文字列理論を専門としているため、二人は全く別の研究分野である。

今まで、私の研究で困難な問題が発生したとき、問題に取り組む代わりに困難を避けてきた。困難な問題を引き出しの中にしまうように見ないふりをして、研究を進めてきたが、どの程度まで困難な問題を実際に解けるかは考えていなかった。番原先生に出会って、この引き出しを開けてみようという気持ちになった。そして、共同研究でどのような問題に取り組んだかを以下の例を用いて説明する。その例とは文字列アトラクタである。表 1 は例を示す。

文字列アトラクタは以下のように定義される。ただし、整数 n は入力文字列 $T[1..n]$ の長さとする。

定義 2.1. $\Gamma \subseteq [1..n]$ は整数区間 1 から n までの任意の部分集合を示す。 T の任意の部分文字列 $T[i..j]$ について、 $p \in \text{cover}(T[i..j])$ になる位置 $p \in \Gamma$ が存在すると、 Γ を**アトラクタ**と呼ぶ。ただし、 $\text{cover}(T[i..j]) := \bigcup_{p \in \text{occ}(T[i..j])} [p..p+j-i]$ を T 中における $T[i..j]$ の含まれる位置集合と、 $\text{occ}(T[i..j]) := \{k \mid T[k..k+(j-i)] = T[i..j]\}$ を $T[i..j]$ の T 中の出現位置集合とする。

素朴な例として、 $[1..n]$ は T のアトラクタである。理由は、任意 $i, j \in [1..n]$ に対して $\text{cover}(T[i..j]) \subseteq [1..n]$ である。しかし、探したいものは、一番要素が少ないアトラクタ Γ である。 Γ のサイズは γ とし、以下のことが知られている。まず、 γ の計算は NP 困難である [5]。さらに、 γ の計算は APX 困難であるが、対数的な近似が存在する [5]。最近、様々な特別な文字列の γ が特定された。例えば、ドゥブリュエイン文字列 [7]、無限の文字列 [8]、シュー・モールス文字列 [2]、フィボナッチ文字列の変種 [3]。一般のおよび厳密な計算については、[4] が指数的なアルゴリズム、[1] が MAX-SAT の符号化、[10] が ZDD を基にした列挙アルゴリズムを提案した。

もちろん、アトラクタは文字列の世界の中で唯一の NP 困難な問題ではない。様々な困難な問題について、[6] と 2023 年冬の LA シンポジウムで [9] がいくつかの文字列問題の ASP で符号化を紹介した。現段階で、複数の研究者がこの研究に注目し、グループ研究活動を山梨大学で行った。これからも、積極的にこの研究を協力して進めたいと思う。

表 1: 文字列 $T = \text{ananas}$ の中に出現する部分文字列の cover . 各部分文字列は行で、ある部分文字列の cover の要素は $\circ$ で表現される。例えば、 $\text{cover}(n) = \{2, 4\}$. 最小のアトラクタの例は $\Gamma = \{1, 4, 6\}$.

	1	2	3	4	5	6
$T =$	a	n	a	n	a	s
a	$\circ$		$\circ$		$\circ$	
an	$\circ$	$\circ$	$\circ$	$\circ$		
ana	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$	
anan	$\circ$	$\circ$	$\circ$	$\circ$		
anana	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$	
ananas	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$
n		$\circ$		$\circ$		
na		$\circ$	$\circ$	$\circ$	$\circ$	
nan		$\circ$	$\circ$	$\circ$		
nana		$\circ$	$\circ$	$\circ$	$\circ$	
nanas		$\circ$	$\circ$	$\circ$	$\circ$	$\circ$
anas			$\circ$	$\circ$	$\circ$	$\circ$
nas				$\circ$	$\circ$	$\circ$

参考文献

- [1] Hideo Bannai, Keisuke Goto, Masakazu Ishihata, Shunsuke Kanda, Dominik Köppl, and Takaaki Nishimoto. “Computing NP-hard Repetitiveness Measures via MAX-SAT”. In: *Proc. ESA*. Vol. 244. LIPIcs. 2022, 12:1–12:16. doi: 10.4230/LIPIcs.ESA.2022.12.
- [2] Francesco Dolce. “String Attractors for Factors of the Thue–Morse Word”. In: *Proc. WORDS*. Vol. 13899. LNCS. 2023, pp. 117–129. doi: 10.1007/978-3-031-33180-0_9.
- [3] France Gheeraert, Giuseppe Romana, and Manon Stipulanti. “String Attractors of Fixed Points of k -Bonacci-Like Morphisms”. In: *Proc. WORDS*. Vol. 13899. LNCS. 2023, pp. 192–205. doi: 10.1007/978-3-031-33180-0_15.
- [4] Dominik Kempa, Alberto Policriti, Nicola Prezza, and Eva Rotenberg. “String Attractors: Verification and Optimization”. In: *Proc. ESA*. Vol. 112. LIPIcs. 2018, 52:1–52:13. doi: 10.4230/LIPIcs.ESA.2018.52.
- [5] Dominik Kempa and Nicola Prezza. “At the roots of dictionary compression: string attractors”. In: *Proc. STOC*. 2018, pp. 827–840. doi: 10.1145/3188745.3188814.
- [6] Dominik Köppl. “Encoding Hard String Problems with Answer Set Programming”. In: *Proc. CPM*. Vol. 259. LIPIcs. 2023, 17:1–17:21. doi: 10.4230/LIPIcs.CPM.2023.17.
- [7] Sabrina Mantaci, Antonio Restivo, Giuseppe Romana, Giovanna Rosone, and Marinella Sciortino. “String Attractors and Combinatorics on Words”. In: *Proc. ICTCS*. Vol. 2504. CEUR Workshop Proceedings. 2019, pp. 57–71.
- [8] Antonio Restivo, Giuseppe Romana, and Marinella Sciortino. “String Attractors and Infinite Words”. In: *Proc. LATIN*. Vol. 13568. LNCS. 2022, pp. 426–442. doi: 10.1007/978-3-031-20624-5_26.
- [9] クップル ドミニク and 番原 睦則. *Answer Set Programming を用いた圧縮指標の計算*. Tech. rep. Local Proceedings of the LA Symposium Winter 2023, 2024.
- [10] 藤岡 祐太, 斎藤 寿樹, and クップル ドミニク. *ZDD を用いた最小文字列アトラクタの列挙*. 日本オペレーションズ・リサーチ学会 九州支部 九州地区における OR 若手研究交流会. 2023.