The combinatorial equivalence of a computability theoretic question

Lu Liu

Email: g.jiayi.liu@gmail.com

Central South University School of Mathematics and Statistics

Delta 2024 Logic Workshop

September 4, 2025

Introduction

- ▶ We prove that a question of Miller and Solomon——whether every coloring $c: d^{<\omega} \to k$ admits a c-computable variable word infinite solution, is equivalent to a combinatorial question.
- ► The combinatorial question asked whether there is a sequence of positive integers so that each of its initial segment satisfies a Ramsey type property.
- Moreover, the negation of the combinatorial question is a generalization of Hales-Jewett theorem.

We thank Denis Hirschfeldt, Benoit Monin and Ludovic Patey for helpful discussion.

- A question of Miller and Solomon
- Related literature
- 3 The combinatorial equivalence
- **4** On $ENSH_k^d$ and Hales-Jewett Theorem
- Further discussion—Is this rare?

We adopt the problem-instance-solution framework.

Definition 1 (Variable word)

 \triangleright An *n-variable word over d* is a sequence v (finite or infinite) of $\{0, \dots, d-1\} \cup \{x_0, x_1, \dots\}$ where there are n many variables in v.

We adopt the problem-instance-solution framework.

Definition 1 (Variable word)

- An *n*-variable word over d is a sequence v (finite or infinite) of $\{0, \dots, d-1\} \cup \{x_0, x_1, \dots\}$ where there are n many variables in v.
- ▶ Given an $\vec{a} \in d^{\tilde{n}}$, an *n*-variable word v, suppose x_{m_0}, x_{m_1}, \cdots occur in v with $m_{\hat{n}-1} < m_{\hat{n}}$ for all \hat{n} . We write $v(\vec{a})$ for the $\{0, \cdots, d-1\}$ -string obtained by substitute $x_{m_{\hat{n}}}$ with $\vec{a}(\hat{n})$ in v for all $\hat{n} < \tilde{n}$ and then truncating the result just before the first occurrence of $x_{m_{\tilde{n}}}$.

We adopt the problem-instance-solution framework.

Definition 1 (Variable word)

- An *n*-variable word over d is a sequence v (finite or infinite) of $\{0, \dots, d-1\} \cup \{x_0, x_1, \dots\}$ where there are n many variables in v.
- Given an $\vec{a} \in d^{\tilde{n}}$, an *n*-variable word v, suppose x_{m_0}, x_{m_1}, \cdots occur in v with $m_{\hat{n}-1} < m_{\hat{n}}$ for all \hat{n} . We write $v(\vec{a})$ for the $\{0, \cdots, d-1\}$ -string obtained by substitute $x_{m_{\hat{n}}}$ with $\vec{a}(\hat{n})$ in v for all $\hat{n} < \tilde{n}$ and then truncating the result just before the first occurrence of $x_{m_{\tilde{n}}}$.
- We write $P_{x_m}(v)$ for the set of positions of x_m in v, namely $\{t: v(t) = x_m\}$; the first occurrence of a variable x_m in v refers to the integer min $P_{x_m}(v)$.

Example 2

Infinite variable word v on $\{0,1\}$:

$$011 x_0x_0 011 x_1 x_0x_0 x_1x_100 x_2x_2 (1.1)$$

$$\vec{a} = 10, v(\vec{a}) = 011 11 011 0 11 0000.$$

$$P_{x_0}(v) = \{3, 4, 9, 10, \dots\}.$$

Example 2

Infinite variable word v on $\{0,1\}$:

$$011 \quad \mathbf{x_0 x_0} \quad 011 \quad \mathbf{x_1} \quad \mathbf{x_0 x_0} \quad \mathbf{x_1 x_1 00} \quad \mathbf{x_2 x_2} \cdots \quad (1.1)$$

$$\vec{a} = \mathbf{10}, v(\vec{a}) = 011 \quad \mathbf{11} \quad 011 \quad 0000.$$

$$P_{x_0}(v) = \{3, 4, 9, 10, \cdots\}.$$

Definition 3

- ightharpoonup Problem: VWI(d, k).
- Instance: $c: d^{<\omega} \to k$.
- Solution: an ω -variable word v such that $\{v(\vec{a}) : \vec{a} \in d^{<\omega}\}$ is monochromatic.

VWI vs RCA

Joe Miller and Solomon proposed the following question in [?].

Question 4

Does every VWI(d, k)-instance c admit c-computable solution?

VWI vs RCA

Joe Miller and Solomon proposed the following question in [?].

Question 4

Does every VWI(d, k)-instance c admit c-computable solution?

Or in terms of reverse mathematics:

Question 5

Is VWI(d, k) provable in RCA?

Other versions of variable word problem

Definition 6 (VW, OVW)

If we require the occurrence of x_i being finite for all i then the problem is called VW.

If we require all the occurrence of x_i comes before any occurrence of x_{i+1} then it is called OVW (ordered variable word).

Other versions of variable word problem

Definition 6 (VW, OVW)

If we require the occurrence of x_i being finite for all i then the problem is called VW.

If we require all the occurrence of x_i comes before any occurrence of x_{i+1} then it is called OVW (ordered variable word).

The problem is proposed by [?] and studied in [?], [?]. Clearly,

Theorem 7

 $VWI(d, k) \le VW(d, k) \le OVW(d, k)$. $VWI(d, k) \Leftrightarrow VWI(d, k+1), VW(d, k) \Leftrightarrow VW(d, k+1), OVW(d, k) \Leftrightarrow$ $\mathsf{OVW}(d, k+1)$.

The complexity of OVW, VW

<u>Theorem 8 ([?])</u>

There exists a computable instance of OVW(2,2) that does not admit Δ_2^0 solution. Thus RCA₀ + WKL does not prove VW(2, 2).

The complexity of OVW, VW

Theorem 8 (?)

There exists a computable instance of OVW(2,2) that does not admit Δ_2^0 solution. Thus RCA₀ + WKL does not prove VW(2, 2).

The following result answers a question of [?] and [?].

Theorem 9 (Monin, Patey, L)

- For every computable OVW(2,2)-instance c, every \emptyset' -PA degree compute a solution to c.
- ightharpoonup There exists a computable $\mathsf{OVW}(2,2)$ -instance such that every solution is \emptyset' -DNC degree.

Corollary 10 (Monin, Patey, L)

ACA proves OVW(2, 2).

Question 11 ([?])

Does OVW(d, k) or VW(d, k) implies ACA_0 for some l?

A combinatorial equivalence of "VWI(2, 2) vs RCA"

Definition 12 $(ENSH_k^d)$

Let n_0, n_1, \dots, n_{r-1} be a sequence of positive integers, let $N_0 = \{0, \dots, n_0 - 1\}, \ N_1 = \{n_0, \dots, n_0 + n_1 - 1\}, \dots, N_{r-1} = \{n_0 + \dots + n_{r-2}, \dots, n_0 + \dots + n_{r-1} - 1\},$ and $N = \bigcup_{s \leq r-1} N_s$; let $f : d^N \to k$. We say $n_0 \cdots n_r$ is sectionally-homogeneous for f if there exists an $s \leq r-1$, an n_s -variable word v over d of length N such that the first occurrence of variables in v consist of N_s , i.e.,

$$\{\min P_{x_m}(v): m \in \omega\} = N_s,$$

and v is monochromatic for f.

▶ We write $ENSH_k^d(n_0 \cdots n_{r-1})$ iff there exists a coloring $f: d^N \to k$ such that $n_0 \cdots n_{r-1}$ is not sectionally-homogeneous for f. In that case we say f witnesses $ENSH_k^d(n_0 \cdots n_{r-1})$.

A combinatorial equivalence of "VWI(2,2) vs RCA"

Let $ENSH_{k}^{d}$ denote the set of infinite sequence of integers $n_{0}n_{1}\cdots$ such that $ENSH_{k}^{d}(n_{0}\cdots n_{r})$ holds for all $r \in \omega$.

Theorem 13 (?)

The following are equivalent:

- ightharpoonup There exists a VWI(d, k)-instance c that does not admit c-computable solution.
- ▶ There exists an $X \in ENSH_{\iota}^d$.

Intuition on $ENSH_k^d(n_0 \cdots n_{r-1})$

Proposition 14

If \vec{n} is a subsequence of $\vec{\hat{n}}$ or $\vec{n} \geq \vec{\hat{n}}$, then $ENSH_k^d(\vec{\hat{n}})$ implies $ENSH_k^d(\vec{n})$.

Intuition on $ENSH_k^d(n_0 \cdots n_{r-1})$

Proposition 15

 $ENSH_{2}^{2}(22)$, $ENSH_{2}^{2}(222)$ holds. $ENSH_{2}^{2}(n)$ holds for all n > 0.

Proof.

To see $ENSH_2^2(22)$, consider

$$f(\vec{a}) = \vec{a}(0) + \vec{a}(1) + \vec{a}(2) \mod 2.$$

To see $ENSH_2^2(222)$, consider

$$f(\vec{a}) = I(\vec{a}(0) + \vec{a}(1) > 0) + \vec{a}(2) + \vec{a}(3) + \vec{a}(4) \mod 2.$$

Where I() is the indication function. To see $ENSH_2^2(n)$, simply consider $f(\vec{a}) = \vec{a}(0) \mod 2.$

Intuition on $ENSH_{k}^{d}(n_{0}\cdots n_{r-1})$

Proposition 16

 $ENSH_2^2(2222)$ does not hold.

Proof.

We don't know the proof. Adam P. Goucher at Mathoverflow examined this using SAT solver (

https://mathoverflow.net/questions/293112/ramsey-type-theorem). It's easy to check that the following functions don't work:

$$f(\vec{a}) = I(\vec{a}(0) + \vec{a}(1) > 0) + \vec{a}(2) + \vec{a}(3) + \vec{a}(4) + \vec{a}(6) \mod 2;$$

$$f(\vec{a}) = I(\vec{a}(0) + \vec{a}(1) > 0) + I(\vec{a}(2) + \vec{a}(3) > 0) +$$

$$+ \vec{a}(4) + \vec{a}(5) + \vec{a}(6) \mod 2;$$
(3.1)



 (\Leftarrow)

 \triangleright A Turing functional Ψ^X computes a variable word if Ψ^X is an enumerable set (possibly finite) $\{v_0, v_1, \dots\}$ of finitely long variable words such that $v_0 \prec v_1 \prec \cdots$.

 (\Leftarrow)

- \triangleright A Turing functional Ψ^X computes a variable word if Ψ^X is an enumerable set (possibly finite) $\{v_0, v_1, \dots\}$ of finitely long variable words such that $v_0 \prec v_1 \prec \cdots$.
- Putting priority argument aside, assume each Turing functional is total. i.e.,
 - for each $r \in \omega$, let $v_r \in \Psi_r^X$ be such that v_r contains X(r) many variables whose first occurrence is after $|v_{r-1}|$.

(\Leftarrow)

- \triangleright A Turing functional Ψ^X computes a variable word if Ψ^X is an enumerable set (possibly finite) $\{v_0, v_1, \dots\}$ of finitely long variable words such that $v_0 \prec v_1 \prec \cdots$.
- ▶ Putting priority argument aside, assume each Turing functional is total. i.e., for each $r \in \omega$, let $v_r \in \Psi_r^X$ be such that v_r contains X(r) many variables whose first occurrence is after $|v_{r-1}|$.
- ▶ Suppose $(f_r: r \in \omega)$ witnesses $ENSH_k^d(X \upharpoonright r)$. We transform these f_r to a coloring c so that there is no $v \succeq v_r$ monochromatic for c.

▶ To define c on d^n , let r(n) be the maximal integer such that $|v_{r(n)}| \leq n$. We ensure that c on d^n "oppress" v_r for all $r \leq r(n)$.

- ▶ To define c on d^n , let r(n) be the maximal integer such that $|v_{r(n)}| \leq n$. We ensure that c on d^n "oppress" v_r for all $r \leq r(n)$.
- \triangleright Let P_r be the set of first occurrence of variables in v_r whose first occurrence is after $|v_{r-1}|$. W.l.o.g, suppose $|P_r| = X(r)$ for all $r \in \omega$.

- \triangleright To define c on d^n , let r(n) be the maximal integer such that $|v_{r(n)}| \leq n$. We ensure that c on d^n "oppress" v_r for all $r \leq r(n)$.
- \triangleright Let P_r be the set of first occurrence of variables in v_r whose first occurrence is after $|v_{r-1}|$. W.l.o.g, suppose $|P_r| = X(r)$ for all $r \in \omega$.
- ▶ Define $c(\vec{a}) = f_{r(n)+1}(\vec{a} \upharpoonright \bigcup_{r \le r(n)} P_r)$.



▶ Take advantage of some particular algorithms Φ_0, Φ_1, \cdots and show that their failure (to compute a solution to c) gives rise to a sequence $X \in ENSH_k^d$.

 (\Rightarrow)

- ▶ Take advantage of some particular algorithms Φ_0, Φ_1, \cdots and show that their failure (to compute a solution to c) gives rise to a sequence $X \in ENSH_k^d$.
- \bullet $\Phi_0^c, \Phi_1^c, \cdots$ are greedy algorithms in the sense that they extend their current computation (which is a finitely long variable word) whenever possible. More precisely,

(\Rightarrow)

- ▶ Take advantage of some particular algorithms Φ_0, Φ_1, \cdots and show that their failure (to compute a solution to c) gives rise to a sequence $X \in ENSH_k^d$.
- \bullet $\Phi_0^c, \Phi_1^c, \cdots$ are greedy algorithms in the sense that they extend their current computation (which is a finitely long variable word) whenever possible. More precisely,
- $lackbox{\Phi}_{r+1}^c$ extends its current computation from v_{r+1} to some $\hat{v} \succeq v_{r+1}$ where \hat{v} has more variables than v_{r+1} , whenever it is found that for some $\vec{a} \in d^{|v_r|+1}$, \hat{v}/\vec{a} is monochromatic for c.

(\Rightarrow)

- ▶ Take advantage of some particular algorithms Φ_0, Φ_1, \cdots and show that their failure (to compute a solution to c) gives rise to a sequence $X \in ENSH_{k}^{d}$.
- \bullet $\Phi_0^c, \Phi_1^c, \cdots$ are greedy algorithms in the sense that they extend their current computation (which is a finitely long variable word) whenever possible. More precisely,
- $lackbox{\Phi}_{r+1}^c$ extends its current computation from v_{r+1} to some $\hat{v} \succeq v_{r+1}$ where \hat{v} has more variables than v_{r+1} , whenever it is found that for some $\vec{a} \in d^{|v_r|+1}$, \hat{v}/\vec{a} is monochromatic for c.
- ▶ Moreover, Φ_{r+1}^c will build its solution v_{r+1} based on $\Phi_0^c, \dots, \Phi_r^c$ in the sense that all variables in v_{r+1} occur after $|v_r|$ and if some $\Phi_{\tilde{x}}^c$ extends its current computation, then all Φ_r^c (where $r > \tilde{r}$) will restart all over again.

 \triangleright Since c does not admit a c-computable solution, for every $r \in \omega$, the computation of Φ_r^c stucks at some v_r .

- \triangleright Since c does not admit a c-computable solution, for every $r \in \omega$, the computation of Φ_r^c stucks at some v_r .
- More precisely, let $\hat{v}_r = v_r x_{n_r-1}$ (where we assume that all variables in v_r are $\{x_0, \dots, x_{n_r-2}\}\)$, we have

- \triangleright Since c does not admit a c-computable solution, for every $r \in \omega$, the computation of Φ_r^c stucks at some v_r .
- More precisely, let $\hat{v}_r = v_r x_{n_r-1}$ (where we assume that all variables in v_r are $\{x_0, \dots, x_{n_r-2}\}\)$, we have
- ▶ there is no $\hat{v} \succeq \hat{v}_r$ such that for some $\vec{a} \in d^{|\hat{v}_{r-1}|}$, \hat{v}/\vec{a} is monochromatic for c; moreover, all variables in v_r occur after $|v_{r-1}|$ and $|v_r| > |v_{r-1}|$.
- ▶ We show that $n_0 n_1 n_2 \cdots \in ENSH_k^d$.

► Fix an $r \in \omega$, let $N = n_0 + \cdots + n_r$. To define $f: d^N \to k$ witnessing $ENSH_{\iota}^{d}(n_{0}\cdots n_{r})$, for every $\vec{a}\in d^{N}$ we map \vec{a} to a word $\vec{\hat{a}} = h(\vec{a}) \in d^{|\hat{v}_r|}$ and let $f(\vec{a}) = c(\vec{\hat{a}})$.

- Fix an $r \in \omega$, let $N = n_0 + \cdots + n_r$. To define $f: d^N \to k$ witnessing $ENSH_{k}^{d}(n_{0}\cdots n_{r})$, for every $\vec{a}\in d^{N}$ we map \vec{a} to a word $\vec{\hat{a}} = h(\vec{a}) \in d^{|\hat{v}_r|}$ and let $f(\vec{a}) = c(\vec{\hat{a}})$.
- Intuitively, h is defined by connecting each element of N, say $n_0 + \cdots + n_{s-1} + m$, to a set $P_{x_m}(\hat{v}_s)$ and copy the value $\vec{a}(n_0 + \cdots + n_{s-1} + m)$ to $\hat{a}(t)$ for all $t \in P_{x_m}(\hat{v}_s)$. More precisely,

Proof of theorem 13

- Fix an $r \in \omega$, let $N = n_0 + \cdots + n_r$. To define $f: d^N \to k$ witnessing $ENSH_{k}^{d}(n_{0}\cdots n_{r})$, for every $\vec{a}\in d^{N}$ we map \vec{a} to a word $\vec{\hat{a}} = h(\vec{a}) \in d^{|\hat{v}_r|}$ and let $f(\vec{a}) = c(\vec{\hat{a}})$.
- Intuitively, h is defined by connecting each element of N, say $n_0 + \cdots + n_{s-1} + m$, to a set $P_{x_m}(\hat{v}_s)$ and copy the value $\vec{a}(n_0 + \cdots + n_{s-1} + m)$ to $\hat{a}(t)$ for all $t \in P_{x_m}(\hat{v}_s)$. More precisely,
- ▶ Suppose $\vec{a} = \vec{a}_0 \cdots \vec{a}_r$ where $|\vec{a}_s| = n_s$ for all s < r. Let

$$\vec{\hat{a}}_s = \hat{v}_s(\vec{a}_s) \mid_{|\hat{v}_{s-1}|}^{|\hat{v}_s|-1} \text{ and } h(\vec{a}) = \vec{\hat{a}}_0 \cdots \vec{\hat{a}}_r.$$



Theorem 17

The following two classes of oracles are equal:

 $\{D \subseteq \omega : D' \text{ computes a member in } ENSH_k^d.\}$ $\{D \subseteq \omega : D \ computes \ a \ VWI(d, k) \text{-instance } c$ that does not admit a c-computable solution.

Relation to Hales-Jewett theorem

 \triangleright Disproving $ENSH_k^d$ on certain sequences is a generalization of Hales-Jewett theorem.

Relation to Hales-Jewett theorem

Hales-Jewett theorem. ▶ For $d, k, n \in \omega$, let HJ(d, k, n) denote the assertion that

 \triangleright Disproving $ENSH_k^d$ on certain sequences is a generalization of

there exists an N such that for every $c: d^N \to k$, there exists an *n*-variable word v of length N monochromatic for c.

Relation to Hales-Jewett theorem

- \triangleright Disproving $ENSH_k^d$ on certain sequences is a generalization of Hales-Jewett theorem.
- ▶ For $d, k, n \in \omega$, let HJ(d, k, n) denote the assertion that there exists an N such that for every $c: d^N \to k$, there exists an *n*-variable word v of length N monochromatic for c.

Theorem 18 (Hales-Jewett theorem)

For every $d, k, n \in \omega$, HJ(d, k, n) holds.

- ► HJ theorem is of fundamental importance in combinatorics.
- HJ theorem \Rightarrow van der Waerden theorem (which says that for every partition of integers, every $r \in \omega$, there exists an arithmetical progression of length r in one part).

- ► HJ theorem is of fundamental importance in combinatorics.
- \triangleright HJ theorem \Rightarrow van der Waerden theorem (which says that for every partition of integers, every $r \in \omega$, there exists an arithmetical progression of length r in one part).
- \triangleright The density HJ theorem \Rightarrow the density van der Waerden theorem, namely Szemerédi's theorem, which asserts that for every set A of integers of positive density (meaning $\limsup_{n\to\infty} |A\cap n|/n > 0$), every $r \in \omega$, there exists an arithmetical progression in A of length r (conjectured by Erdős and Turán).

- ▶ We show that: $\forall d, k, n[n^{\omega} \notin ENSH_k^d] \Leftrightarrow HJ$ theorem.
- Actually,

$$n^{\omega} \notin ENSH_k^d \Rightarrow HJ(d, k, n)$$
 and $HJ(d^n, k, 1) \Rightarrow n^{\omega} \notin ENSH_k^d$.

For every $d, k, n \in \omega$, $n^{\omega} \notin ENSH_k^d$.

Proof.

▶ For example we prove this for d, n = 2.

For every $d, k, n \in \omega$, $n^{\omega} \notin ENSH_k^d$.

Proof.

- For example we prove this for d, n = 2.
- ▶ Using HJ(4, k, 1), let r be the witness.
- Show that $ENSH_k^d(2 \underbrace{\cdots}_{r \text{ many}} 2)$ does not hold.

For every $d, k, n \in \omega$, $n^{\omega} \notin ENSH_k^d$.

Proof.

- For example we prove this for d, n = 2.
- \blacktriangleright Using HJ(4, k, 1), let r be the witness.
- Show that $ENSH_k^d(2) 2)$ does not hold. r many
- ightharpoonup Code 2^{2r} into 4^r where $\vec{a}(2t)\vec{a}(2t+1)$ (00,01,10,11 respectively) is coded into $\hat{a}(t)$ (0, 1, 2, 3 respectively).

For every $d, k, n \in \omega$, $n^{\omega} \notin ENSH_k^d$.

Proof.

- For example we prove this for d, n = 2.
- \blacktriangleright Using HJ(4, k, 1), let r be the witness.
- ▶ Show that $ENSH_k^d(2)$ does not hold. r many
- ightharpoonup Code 2^{2r} into 4^r where $\vec{a}(2t)\vec{a}(2t+1)$ (00,01,10,11 respectively) is coded into $\hat{a}(t)$ (0, 1, 2, 3 respectively).
- Given a coloring $c: 2^{2r} \to k$, consider $\hat{c}: 4^r \ni \vec{\hat{a}} \mapsto c(\vec{a})$.
- Let \hat{v} be a 1-variable word monochromatic for \hat{c} and consider vsuch that $v(2t)v(2t+1) = 00, 01, 10, 11, x_0x_1$ respectively if $\hat{v}(t) = 0, 1, 2, 3, x_0$ respectively.



The following theorem generalizes HJ theorem on d=2, k=2, n=2.

Theorem 20 ([?])

For every sequence $n_0 n_1 \cdots$ of positive integers with $n_s = 2$ for some s, $n_0 n_1 \cdots \notin ENSH_2^2$.

Lemma 21

There exists a sequence $n_0 \cdots n_r$ such that $ENSH_2^2(n_0 \cdots n_r)$ holds but $ENSH_2^2(n_0 \cdots n_r n)$ does not hold for all n.

Proof.

For example, $n_0 \cdots n_r = 1$ and note that $ENSH_2^2(1)$ is true but $ENSH_2^2(1n)$ is not true for any n.



Some open questions

Question 22

- ▶ Does $ENSH_2^2(2223)$ holds? Does $ENSH_2^2(222n)$ holds for sufficiently large n?
- ▶ Is it true that for every n, \hat{n} $ENSH_2^2(n \ \hat{\underline{n}} \cdots \hat{\underline{n}})$ does not hold. n+1 many

How computability question transform to non computability question

Further discussion——Is this rare?

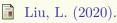
Many thanks!



Carlson, T. J. and Simpson, S. G. (1984).

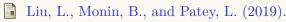
A dual form of ramsey's theorem.

Advances in Mathematics, 53(3):265-290.



The combinatorial equivalence of a computability theoretic question.

arXiv preprint arXiv:2012.13588.



A computable analysis of variable words theorems.

Proceedings of the American Mathematical Society, 147(2):823–834.



Miller, J. S. and Solomon, R. (2004).

Effectiveness for infinite variable words and the dual ramsey theorem.

Archive for Mathematical Logic, 43(4):543–555.



Montalbán, A. (2011). Open questions in reverse mathematics.

Bulletin of Symbolic Logic, 17(03):431-454.