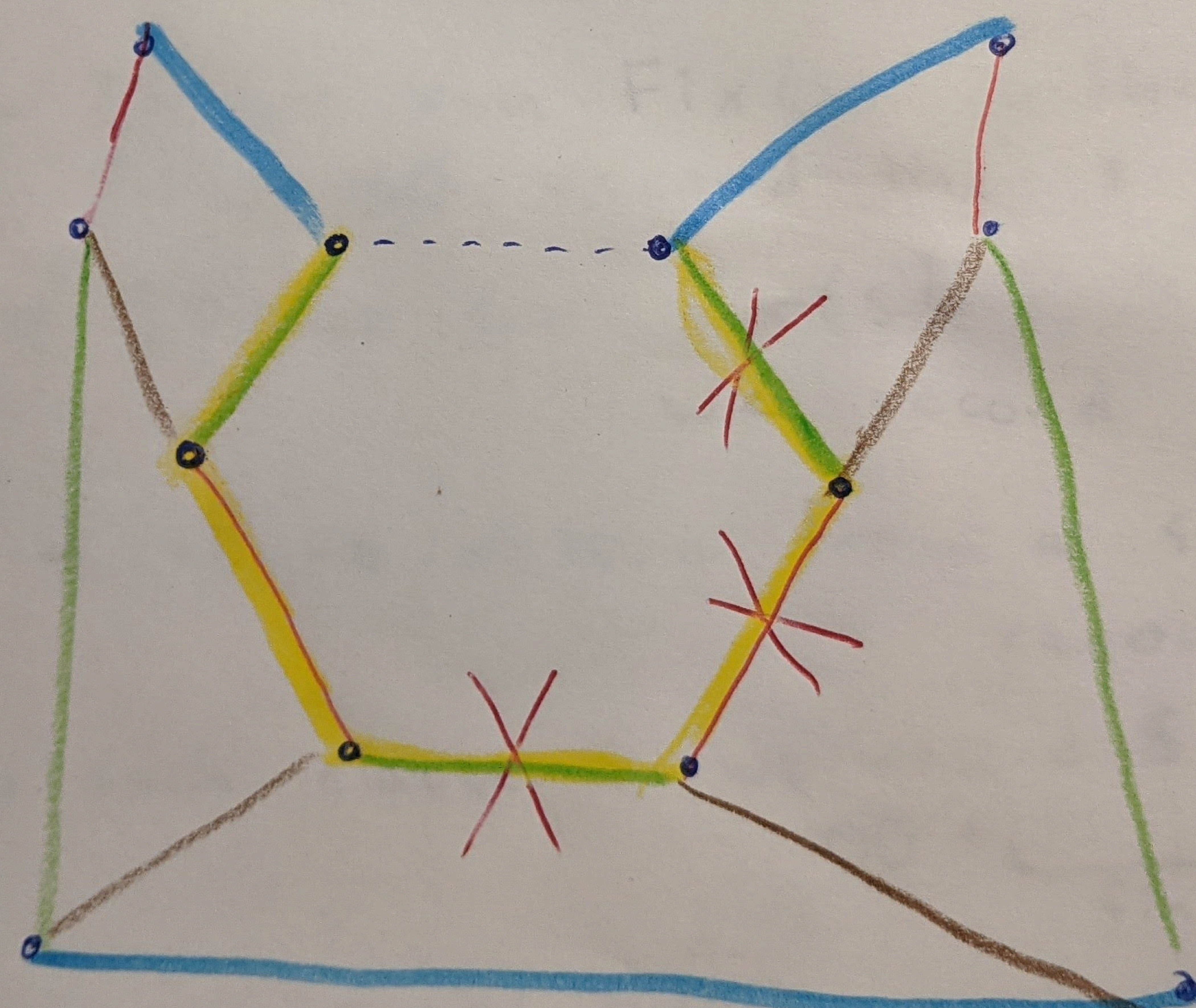
Lecture 20: Entropy compression II



A k-CNF formula

$$(\underbrace{x_1 \vee \neg x_2 \vee x_3}_{\text{each clause has size } k}) \wedge (x_3 \vee x_4 \vee \neg x_5) \wedge (\neg x_1 \vee \neg x_3 \vee \neg x_5)$$

Theorem 10.1: There exists $C>0$ such that for every k [redacted] every k-CNF formula F s.t. each clause shares variables with $\leq 2^{k-C}$ others is satisfiable.

Moser 2009

(+ Tardos generalization)

(and there exists a randomized algorithm finding the satisfying assignment efficiently.)

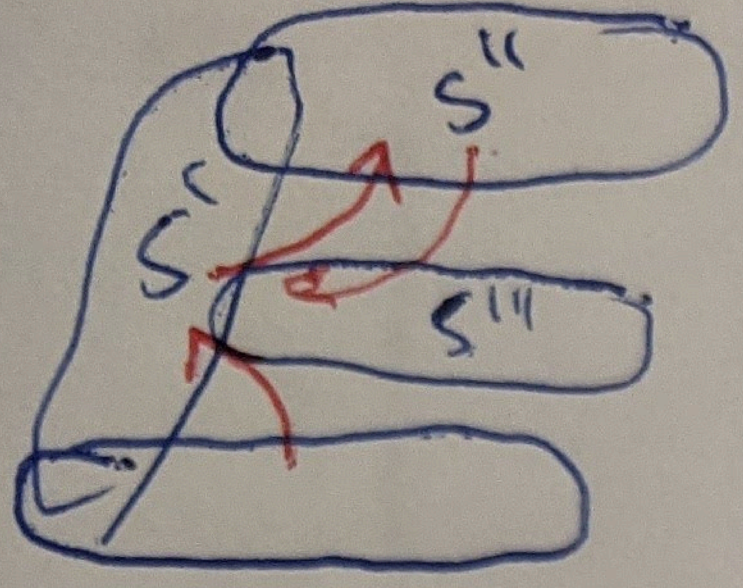
Proof: Let M be an integer depending on F, large and let R be a sequence of M "true, false" assignments.

Algorithm: Subroutine | Fix(s): s is a violated clause.

- reassign the values of variables in S using the next k values in R.
- For each clause s' sharing variables with s | if s' is violated run Fix (s')
- return.

Assign values to all variables say all, true.
- if there exists a violated clause s run Fix(s),

Algorithm terminates if we ~~got~~ succeed or R is empty.



s

φ

History:

For each clause s let

~~$\boxed{\varphi_s}: [n] \leftrightarrow$ set of clauses sharing variables with s~~ $n \leq \boxed{2^{k-c}}$

~~be a bijection. called index~~

$\underline{s, \varphi_s(s'), \varphi_{s'}(s''), R,}$ φ_s: set of causes sharing var. with s $\rightarrow [n]$

$\underline{\varphi_{s'}(s''')}$. be a bijection called index.

- If we run Fix(s) in the main algorithm we record the s in our history.
- If we run Fix(s') from Fix(s) we record $\underbrace{\varphi_s(s')}$
- if we return from a subroutine we record this.

~~constant~~ linear in $|F|$ length (depending on F) $\leftarrow$ For each clause s we record s as $00 + \underbrace{\qquad\qquad}_{\text{finite identifying } s}$.

$\underbrace{k-c+2}$ $\leftarrow$ record $\underbrace{\text{index}}$ as $01 + \underbrace{k-c}$ symbols

$1 \leftarrow$ record $\underbrace{\text{return}}$ as 1.

Total length of History $\leq (k-c+2)$ # Fix calls $+ O(|F|)$.

We can deduce the values from R used in the run of the algorithm from its history and current values.

If algorithm did not succeed then R is empty so we made $M/k \cdot \text{Fix}$ calls.

R is determined by History + Final state.

$$(k - c + 2) \cdot \frac{M}{k} + O(|F|).$$

So for $c > 2$

Entropy compression

if M is large History length is at most $(1-\varepsilon) M$. For $\varepsilon > 0$.

If R is random 2^M choices only $2^{(1-\varepsilon)M}$ historys

So with prob. $\leq \boxed{2^{-\varepsilon M}}$ R can be recovered and algorithm fails.

Non-repetitive sequences

Let W be a sequence ~~in alpha~~ of symbols $\{a, b, c\}$

W is non-repetitive ~~if for~~ if no word appear within W twice consequtively.

$$\underbrace{W'}\underbrace{W'}$$

Examples: $abcab \to$ non repetitive

$b\underbrace{abc}\underbrace{abc} \to$ repetitive

$\underline{ab}\,\underline{ab}$, $\quad a\underline{bc}\,\underline{bc}$

Thue 1906: There exists arbitrarily long non-repetitive sequences in the alphabet with 3 symbols.

Proof idea: If W is non-repetitive, then we can obtain non-repetitive W' from it by substitution

Automated sequences

$$\begin{cases} a \to abcab \\ b \to acabcb \\ c \to acbcacb \end{cases}$$

Conjecture: Let $L_1, L_2, \dots, L_n$ be subsets of an alphabet A
$|L_i| = 3$

Then there exist $a_1 \in L_1, \dots, a_n \in L_n$
s.t. $a_1 a_2 \dots a_n$ is non-repetitive.

(If $L_1 = L_2 = \dots = L_n$ → Thue's theorem).

Theorem 10.2: Conjecture above is true if $|L_i| = 4$ for every i.

Grytczuk, Przybyło, Zhu 2010
Grytczuk, Kozik, Micek 2012

Proof: Idea: Extend our string at random
if we generate a repetition

$a_1 a_2 \dots a_i b_1 \dots b_s \underbrace{b_1 \dots b_s}$

~~eras~~ erase the second half of the ~~word~~ repetition

Analyzing history as in 10.1 it is possible to show that this succeeds.

Simplified proof:

Let N_i be the number of non-repetitive words $a_1 a_2 \dots, a_i$ s.t. $a_j \in L_j$ for every j.

Want $N_n \geq 1$.

We will show by induction on i that

$$N_i / N_{i-1} \geq \boxed{x} \text{ for some } x > 1 \qquad \left(\begin{array}{c} x=2 \\ \text{works} \end{array}\right)$$

(This implies $N_j \leq x^{j-i} N_i$ for any $j<i$).

Base case: $i=1$ $\quad N_0 = 1,\ N_1 = 4$ so if $x \leq 4$ we are ok.

Induction step: Consider N_i non-repetitive words of length i and extend them by a symbol from L_{i+1}

$$\underbrace{a_1 \ldots a_i\, b}$$

$$4 N_i \leq \underbrace{N_{i+1}}_{\text{result non-repetitive words}} + \underbrace{N_i + N_{i-1} + \ldots + N_{i-2}}_{\text{results}}$$

$a_1 \ldots a_j\, w\, \boxed{w}$

w has $\frac{i+1-j}{2}$ symbols

by ind hypothesis $\frac{N_{i-s}}{N_i} \leq x^{-s}$

$\frac{1}{1-\frac{1}{x}} = \frac{x}{x-1}$

$a_1 \ldots a_i\, \boxed{a_i} \leq N_i$

$a_1 \ldots a_{i-3}\, bc\, bc \leq N_{i-1}$

$a_{i-T}\, bcd\, bcd \leq N_{i-2}$

$$\frac{N_{i+1}}{N_i} \geq 4 - \underbrace{1 - \frac{1}{x} - \frac{1}{x^2} - \ldots}$$? $\geq$? x

$$4 \geq x + \frac{x}{x-1} \qquad x = 2 \ \checkmark$$

An <u>acyclic edge coloring</u> of a graph G

is a map $c: E(G) \to [k]$

s.t. it is proper - edges sharing an end receive different colors.

and for every cycle the edges of it receive at least 3 colors.

– ~~Conjecture~~

Let Δ be the maximum degree of G.

Let $a'(G)$ denote the minimum number of colors needed for acyclic edge coloring.

<u>Conjecture:</u> $a'(G) \leq \Delta + 2$.

Fiamčik 1978

<u>Homework problem:</u> $a'(G) \leq O(\Delta)$.

assignment 2

(not easy, use LLL, constant is fairly large).

Theorem 10.3: $a'(G) \le 4\Delta - 4$ for every graph of maximum degree Δ. .

Esperet, Parreau 2013

Proof: ~~We will p~~ Let $k = 4\Delta - 4$.

For a subgraph H of G let N_H be the number of acyclic edge colorings of H.

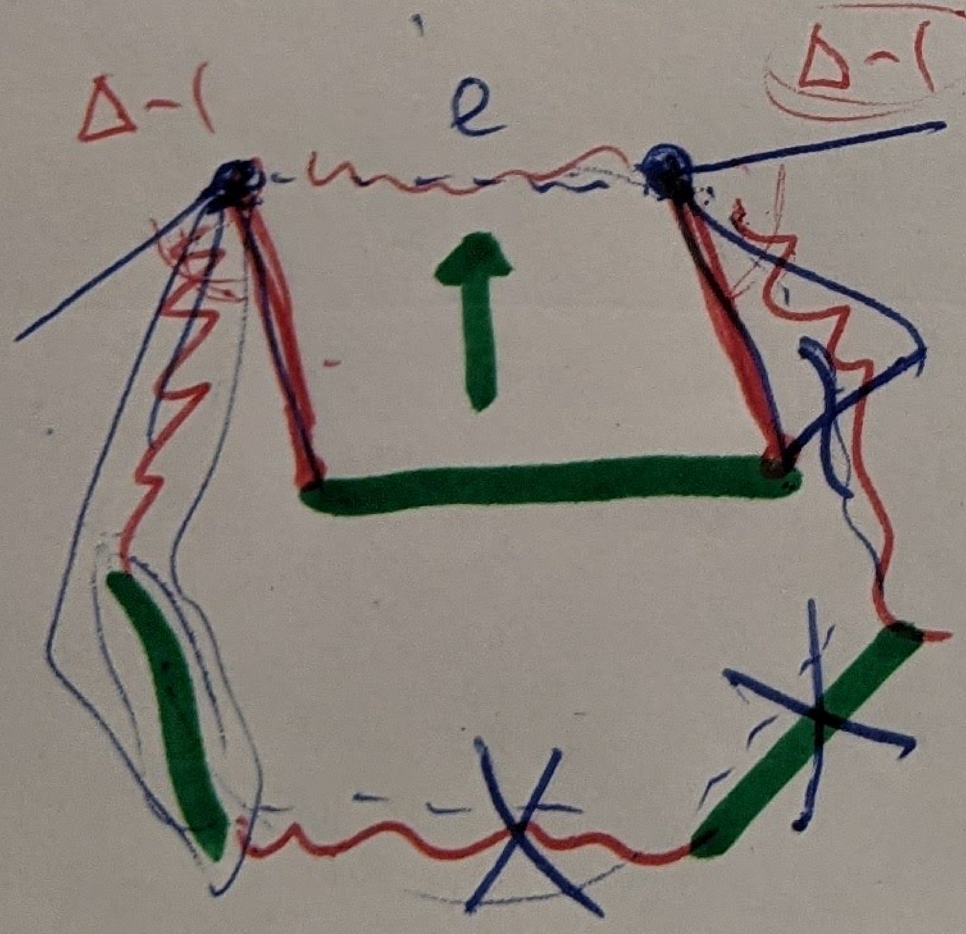
We will prove by induction on $|E(H)|$ that

$N_H \ge x \, N_{H \setminus \{e\}}$ for any edge e in H, where x to be determined

(this implies theorem)

$\left(x = \frac{1+\sqrt{5}}{2} (\Delta - 1) \right)$

~~Proof:~~ **Induction step:**



$$k \cdot N_{H \setminus \{e\}} \le N_H + 2(\Delta-1) N_{H \setminus \{e\}} + \sum_{\substack{\text{cycles } C \\ \text{containing } e \\ \text{of even length} \le 2k}} N_{\underbrace{H-(C-e_{1,C}-e_{2,C})}_{\text{delete } (2k-2) \text{ edges.}}}$$

extension of acyclic coloring of $H \setminus \{e\}$ to H

$2(\Delta-1)N_{H\setminus\{e\}}$: non-proper colorings or colorings giving cycle of length 4.

$$N_H / N_{H \setminus \{e\}} \geq \underset{\substack{\| \\ 4(\Delta-1)}}{k} - 2(\Delta-1) - \underset{\substack{\uparrow \\ \text{max} \\ \#\text{ of cycles} \\ \text{of length} \\ 6}}{(\Delta-1)^4} \cancel{/} \underset{\substack{\uparrow \\ \text{ind hypothesis}}}{x^3} - \frac{(\Delta-1)^6}{x^5} - \frac{(\Delta-1)^8}{x^7}.$$

So if x satisfies

$$2(\Delta-1) - \frac{(\Delta-1)^4}{x^3} - \frac{(\Delta-1)^6}{x^5} - \dots \geq x$$

then the induction step succeeds.

$x = \frac{1+\sqrt{5}}{2}(\Delta-1)$ is as required.