Dekodiranje grupe M_{12}

Andrea Švob (asvob@math.uniri.hr)

Odjel za matematiku
Sveučilište u Rijeci

6.veljače 2009.

- **R.F.Bailey, J.N.Bray,** Decoding the Mathieu group M_{12}, Advances in Mathematics Communications, Volume 1, No.4, 2007, 477-487

"Error-correcting" kodovi

Teorija kodova bavi se proučavanjem točnosti podataka.
Kriptografija se bavi proučavanjem tajnosti podataka.

uređaj za kodiranje$\longrightarrow$ odašiljač$\longrightarrow$ kanal$\longrightarrow$ prijemnik $\longrightarrow$ dekodiranje

Teorija kodova se bavi pojmovima: odašiljač i dekodiranje

Želimo poslati ispravnu poruku kroz bučan kanal, kod kojeg može doći do interferencije. Ukoliko primljena poruka sadrži greške, želimo ju dekodirati tako da otkrijemo originalnu poruku. Ako su moguće pogreške dovoljno različite, tada je postupak moguć.
Pitanje: Što znači da su pogreške dovoljno različite?

Osnovna ideja: Korištenje permutacijske grupe kao koda, gdje su riječi koda elementi te permutacijske grupe.

Definicija

"Error-correcting" kod je skup **C**, *znakova ili simbola koji se nazivaju* **riječi koda**, *a izabrani su iz nekog skupa* **F**, *kojeg nazivamo* **skup alfabeta**.

Definicija

Neka je $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in F$. *Broj*

$$d(x, y) = |\{i | 1 \leq i \leq n, x_i \neq y_i\}|$$

se naziva **Hammingova udaljenost**.

Definicija

Minimalna udaljenost *riječi koda* **C** *je*

$$d(C) = min\{d(x, y), x, y \in C, x \neq y\}.$$

Vrijedi: Neka je poslana riječ koda g i neka je primljena riječ koda w. Tada, ako w sadrži najviše r grešaka, postoji jedinstveni najbliži susjed u C. Primljenu riječ w možemo uspješno dekodirati.

$$r = \left\lfloor \frac{d(C) - 1}{2} \right\rfloor$$

r - sposobnost ispravljanja primljene riječi koda

Elementi grupe permutacija, za koju je poznata minimalna udaljenost, mogu predstavljati poslane riječi koda. Općenito, nije jednostavno koristiti bilo koje permutacijske grupe.

Permutacijska grupa

Definicija

Označimo sa $S(\Omega)$ skup svih bijekcija na skupu Ω. Skup S s obzirom na kompoziciju preslikavanja tvori grupu. Za konačan skup Ω, bijekcije na skupu Ω se nazivaju **permutacije** *skupa Ω.*

Definicija

Grupa svih permutacija skupa Ω zove se **simetrična grupa** *i označava S_n, $n = |\Omega|$.Podgrupa grupe S_n naziva se* **permutacijska grupa**.

Broj svih permutacija skupa Ω: $|S_n| = n!$

Korolar (Cayleyev teorem)

Svaka konačna grupa je izomorfna nekoj permutacijskoj grupi.

Djelovanje grupe na skup

Definicija

Grupa G **djeluje** *na skup Ω ako postoji preslikavanje $f : G \times \Omega \to \Omega$ takvo da vrijedi*

1. $f(g_1, f(g_2, x)) = f(g_1 g_2, x),\ \forall x \in \Omega,\ \forall g_1,\ g_2 \in G,$
2. $f(1, x) = x,\ \forall x \in \Omega.$

Slika djelovanja elementa $g \in G$ na element $x \in \Omega$ označava se $g.x$.

Skup $G_x = \{g \in G \mid g.x = x\} \leq G$ naziva se **stabilizator** elementa x za djelovanje grupe G.

Na skupu Ω na kojeg djeluje grupa G može se definirati relacija

$$x \sim y \Leftrightarrow (\exists g \in G)\text{t.d.je} g.x = y.$$

Relacija $\sim$ je relacija ekvivalencije na skupu Ω.

Klasa ekvivalencije elementa x s obzirom na relaciju $\sim$, $G.x = \{g.x \mid g \in G\}$, naziva se **orbita** elementa x za djelovanje grupe G.

Definicija

Grupa G djeluje **tranzitivno** *na skup Ω ako postoji element $x \in \Omega$ takav da je $G.x = \Omega$. Odnosno, cijeli skup Ω je jedna orbita.*

Definicija

*Grupa G je k-***tranzitivna** *ako za svake dvije k-torke različitih elemenata iz Ω, postoji element $g \in \Omega$ takav da jednu preslikava u drugu tj. $(a_1, \ldots, a_k)$ i $(b_1, \ldots, b_k) \in \Omega$, $\exists g \in G$ t.d. $g.a_i = b_i$, $\forall i = 1, ..., k$*

Definicija

Grupa G je **strogo** *k-***tranzitivna** *ako je element g (iz gornje definicije) jedinstven.*

Vrijedi: Ako je u permutaciji poznato k mjesta, tada je permutacija jedinstveno određena i vrijedi sljedeća Propozicija.

Propozicija

Neke je G strogo k- tranzitivna permutacijska grupa stupnja n. Tada je minimalna udaljenost grupe G jednaka

$$n - k + 1.$$

Mathieve grupe

- 5 **jednostavnih** konačnih grupa koje je pronašao francuski matematičar Emile Leonard Mathieu, u razdoblju između 1861. i 1873.
- Označavamo ih simbolima $M_{11}, M_{12}, M_{22}, M_{23}, M_{24}$ i promatramo kao pemutacijske grupe koje djeluju na skupovima od 11, 12, 22, 23, 24 elemenata
- Postoji 26 sporadičnih jednostavnih grupa i Mathieve grupe su jedne od njih
- Nakon pronalaska Mathieovih grupa - 1965. pronalazak grupe J_1, Jankove grupe reda 175560
- M_{24}; 5-tranzitivna grupa, M_{23}; 4-tranzitivna grupa, M_{22}; 3-tranzitivna grupa, M_{12}; strogo 5-tranzitivna grupa, M_{11}; strogo 4-tranzitivna grupa
- Mathieve grupe se najjednostavnije definiraju kao: **Grupe automorfizama Steinerovog sustava**

Mathieva grupa M_{12}

- Mathieva grupa M_{12} je strogo 5-tranzitivna grupa, stupnja 12, reda 95040
- Generirana je permutacijama: (1 2)(3 4)(5 6)(7 8)(9 10)(11 12) i (1 3 2)(4 7 5)(8 9 11)
- Minimalna udaljenost: $12 - 5 + 1 = 8$

Steinerov sustav

Definicija

Incidencijska struktura $\mathcal{D}$ *je uređena trojka* $(\mathcal{P}, \mathcal{B}, \mathcal{I})$, *gdje su* $\mathcal{P}$ *i* $\mathcal{B}$ *neprazni disjunktni skupovi i* $\mathcal{I} \subseteq \mathcal{P} \times \mathcal{B}$. *Elementi skupa* $\mathcal{P}$ *se nazivaju* **točke**, *elementi skupa* $\mathcal{B}$ **blokovi**, *a relacija* $\mathcal{I}$ **relacija incidencije**.

Broj blokova koji su incidentni s točkom P naziva se **stupanj točke** P i broj točaka koje su incidentne s blokom x naziva se **stupanj bloka** x.

Za incidencijsku strukturu u kojoj je svaka od v točaka stupnja r i svaki od b blokova stupnja k vrijedi $vr = bk$.

Definicija

Neka je $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ incidencijska struktura. Strukturu $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I}^{'})$, gdje je $\mathcal{I}^{'} = \mathcal{P} \times \mathcal{B} - \mathcal{I}$ naziva se **komplementarna struktura** *strukture $\mathcal{D}$.*

Definicija

Konačna incidencijska struktura $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ *je* $t-(v,k,\lambda)$ **dizajn** *ako vrijedi sljedeće:*

1. $|\mathcal{P}| = v$,
2. *svaki element skupa* $\mathcal{B}$ *incidentan je s točno* k *elemenata skupa* $\mathcal{P}$,
3. *svakih* t *elemenata skupa* $\mathcal{P}$ *incidentno je s točno* λ *elemenata skupa* $\mathcal{B}$.

t-dizajn s parametrom λ=1 naziva se **Steinerov sustav** te vrijedi zapis: $S(t,k,v)$

Steinerov sustav trojki $\Longrightarrow$ za $t = 2, k = 3$

Definicija

Neka su $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$ *i* $\mathcal{D}' = (\mathcal{P}', \mathcal{B}', \mathcal{I}')$ *incidencijske strukture. Bijektivno preslikavanje* $f : \mathcal{P} \times \mathcal{B} \to \mathcal{P}' \times \mathcal{B}'$ *je* **izomorfizam** *iz* $\mathcal{D}$ *na* $\mathcal{D}'$ *ako vrijedi:*

1. f *preslikava* $\mathcal{P}$ *na* $\mathcal{P}'$ *i* $\mathcal{B}$ *na* $\mathcal{B}'$
2. $(P, x) \in \mathcal{I} \Rightarrow (f(P), f(x)) \in \mathcal{I}', \ \forall P \in \mathcal{P}$ *i* $\forall x \in \mathcal{B}$

Ako je $\mathcal{D}' = \mathcal{D}$*, onda se preslikavanje* f *naziva* **automorfizam**. *Skup svih automorfizama je grupa s obzirom na kompoziciju funkcija i naziva se* **puna grupa automorfizama** *strukture* $\mathcal{D}$.

Grupe automorfizama Steinerovog sustava:

- $S(4,5,11) \rightarrow M_{11}$
- $S(5,6,12) \rightarrow M_{12}$, $S(5,6,12) = W_{12}$ - mali Wittov dizajn
- $S(3,6,22) \rightarrow M_{22}$
- $S(4,7,23) \rightarrow M_{23}$
- $S(5,8,24) \rightarrow M_{24}$, $S(5,8,24) = W_{24}$ - veliki Wittov dizajn

Algoritam za dekodiranje - UBB

primljena riječ w sadrži r grešaka $\Rightarrow$ sadrži $n - r$ ispravnih simbola

Osnovni problem: Ne možemo točno odrediti u kojim se pozicijama nalaze greške.
Rješenje: Tražimo skup baza.

Definicija

Neka grupa G djeluje na konačnom skupu Ω. **Baza** *grupe G u ovako opisanom djelovanju je niz točaka $(x_1, \ldots, x_b) \in \Omega$ takvih da vrijedi:$G_{(x_1,\ldots,x_b)} = \langle 1 \rangle$. Odnosno, stabilizator ovog niza točaka je jedinični element.*

Definicija

Skup U k-podskupova od Ω zove se (n, k, r)-**otkrivajući***(uncovering) dizajn, ako ima svojstvo: Za svaki r-podskup $R \in \Omega$ postoji skup $S \in \Omega$ t.d. je $R \cap S = \phi$*

Problem $\longrightarrow$ traženje najmanjeg mogućeg otkrivajućeg dizajna za zadani skup parametara

Definicija

Pokrivajući*(covering)* (n, m, r) *dizajn je komplement otkrivajućeg dizajna;* $m = n - k$

Problem $\longrightarrow$ traženje skupa od m-podskupova iz skupa Ω t.d. je $\forall r$-podskup iz $\Omega(r < m)$ sadržan u barem jednom m-skupu.

Ovakav postupak dekodiranja naziva se **otkrivanje pomoću baza** ili "uncovering-by-bases", UBB

- Pronalazak skupa ovakvih baza nije jednostavan, no korištenjem strogo k-tranzitivnih grupa, postupak se pojednostavljuje
- Za strogo k-tranzitivne grupe vrijedi: Bilo koja k-torka elemenata tvori bazu
- Tražimo skup od k- podskupova od $\{1, \ldots, n\}$ takvih da za bilo koji r-skup grešaka vrijedi da je različit od barem jednog k-skupa
- k-skupovi $\Rightarrow$ komplementi blokova od $(n, n-k, r)$ *pokrivajućeg dizajna*
- Skup od svih k-podskupova iz $\{1, \ldots, n\}$ formira *otkrivajuću* bazu

Dokazano je:

- Za ostale grupe (one koje nisu strogo k-tranzitivne) algoritam je vrlo kompliciran
- UBB zaista postoji

Algoritam:

1. START
2. Biranje 1.baze u UBB
3. Identificiranje odgovarajućeg elementa grupe
4. $d_H(g, w) \leq r$
5. STOP
6. U slučaju $d_H(g, w) > r$ vraćamo se i biramo 2.bazu

Primjer

Za strogo 5*-tranzitivnu grupu* M_{12} *imamo:*$n = 12, k = 5, r = 3$.
Potreban je $(12, 5, 3)$*-otkrivajući dizajn.*

1	2	3	4	5
1	2	6	11	12
1	3	7	8	9
1	4	6	7	10
1	5	8	9	11
2	4	8	9	12
2	5	7	10	11
3	4	7	11	12
3	5	6	10	12
3	6	8	9	11
6	7	8	9	10

Pretpostavimo da smo poslali riječ g, a primili riječ w.
g=(1 2 3 4 5 6 7 8 9 10 11 12),
w=(6 2 1 4 6 6 7 8 9 10 11 12)
w sadrži greške na pozicijama 1,3 i 5.
Pustimo algoritam kroz okrivajući dizajn i dobivamo rješenje:

- error (repeated symbol)
- error (repeated symbol)
- (6 3 1 4 12 2 7 8 9 5 10 11), na udaljenosti 6 od w-odobijeno
- error (repeated symbol)
- error (repeated symbol)
- (1 2 3 4 5 6 7 8 9 10 11 12), na udaljenosti 3 od w-prihvaćeno

Ispravno, neispravno i neodređeno dekodiranje

- Poslana riječ$\longrightarrow$ permutacija g, $g \in G$
- Primljena riječ$\longrightarrow$Hammingova riječ w, $w \in H_n$
- Hammingova udaljenost, $i = d(g, w)$=broj grešaka u riječi w

Vrijedi: Riječ koja sadrži i grešaka može se uspješno dekodirati ako postoji jedinstveni element $g \in G$ na udaljenosti i od w, a ne postoji niti jedan element g na udaljenosti manjoj od i.

Riječ w dekodirati će se **ispravno** ako vrijedi:

$$d(w, G) = d(w, g) = i,$$

te ako vrijedi $h \in G$ t.d. $d(w, h) = i \Rightarrow h = g$, pri čemu je:
$d(w, G) = min\{d(w, g) : g \in G\}$

Vjerojatnost da će se riječ na udaljenosti i od g dekodirati ispravno:

$$P(G, g, i) = \frac{a}{b},$$

gdje je:

- a-broj riječi na udaljenosti i od g koje se dekodiraju ispravno
- b-broj riječi na udaljenosti i od g

Napomena

Broj riječi na udaljenosti i od g jednak je broju $(n-1)^i \binom{n}{i} \neq f(g)$. Broj riječi na udaljenosti i od g ne ovisi o g te stoga umjesto $P(G, g, i)$ možemo pisati $P(G, i)$

Ako G ima minimalnu udaljenost d te ako vrijedi $i \leq \lfloor \frac{d-1}{2} \rfloor$ sve riječi na udaljenosti i od g će se dekodirati ispravno i vrijedi, $P(G, i) = 1$

Riječ w na udaljenosti i od g dekodirati će se **neispravno** ako vrijedi:

$$d(w, G) < d(w, g) = i$$

Riječ w na udaljenosti i od g dekodirati će se **neodređeno** ako vrijedi:

$$d(w, G) = d(w, g) = i$$

te postoji $h \in G/\{g\}$ takav da je $d(h, w) = i$

Označimo:

- $Q(G, i)$- vjerojatnost da će se riječ na udaljenosti i dekodirati neodređeno
- $R(G, i)$- vjerojatnost da će se riječ na udaljenosti i dekodirati neispravno

Vrijedi:

$$P(G, i) + Q(G, i) + R(G, i) = 1$$

Riječ je:

- ZELENA: *id* dekodira ispravno
- ŽUTA: *id* dekodira neodređeno
- CRVENA: *id* dekodira neispravno

Riječi na udaljenosti 4

Vrijedi: Za $d(w, M_{12}) \leq 3$, postoji jedinstveno određena riječ koda $\Rightarrow$ za $i \leq 3, P(M_{12}, i) = 1, Q(M_{12}, i) = R(M_{12}, i) = 0$

- poslana je riječ $g \in M_{12}$, primljena je riječ $w \in H_{12}$ sa 4 grešaka
- Minimalna udaljenost grupe M_{12} jednaka je 8 pa postoji element iz M_{12} na udaljenosti 8 od g
- g nije jedinstveno određen kao riječ koda na udaljenosti 4 od w

Zanima nas: Koliko riječi sa 4 greške nemaju jedinstveno određenog najbližeg susjeda?

Lema

Ako se w nalazi na udaljenosti 4 od dviju riječi koda g, h, tada vrijedi:

$$d(g, h) = 8$$

Dokaz: primjenom nejednakosti trokuta

Lema

Neka je $g \in M_{12}$. Tada postoji točno $7\binom{12}{4}$ elementa $h \in M_{12}$ koji zadovoljavaju jednakost: $d(g, h) = 8$

Lema

Za $g, h \in M_{12}$ koji zadovoljavaju $d(g,h) = 8$, postoji točno $\binom{8}{4}$ riječi w koje zadovoljavaju: $d(g,w) = d(h,w) = 4$

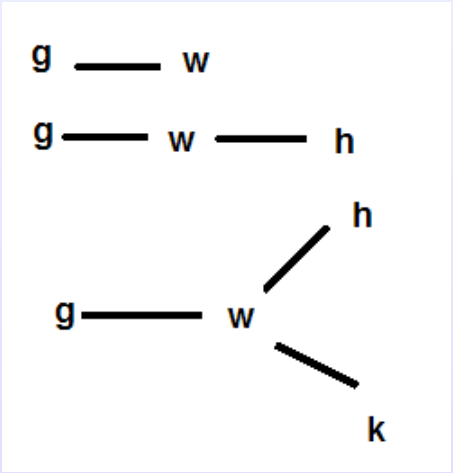
Lema

Neka g, h, w zadovoljavaju uvjete gornje Leme. Tada postoji još najviše jedna riječ koda $k \in M_{12}$ koja zadovoljava: $d(k,w) = 4$

Lema

Neka su $w \in H_{12}$, $g, h, k \in M_{12}$ međusobno različiti te neka vrijedi: $d(g, w) = d(h, w) = d(k, w) = 4$,$g = id$ i h fiksira 1,2,3,4. Broj uređene k-torke elemenata w, g, h, k jednak je 18.

Ako imamo riječ sa 4 greške može se dogoditi:



Osnovni teorem

Teorem

Vjerojatnost da će se riječ koja sadrži 4 greške jedinstveno dekodirati: $P(M_{12}, 4) = \frac{14160}{14641} \cong 0.967147$

Dokaz:

- NSOMP: $g \in M_{12}$ je fiksan, $g = id$, g i h se slažu u prve 4 pozicije
- n_i=broj parova (g, w) takvih da: $d(g, w) = 4$
- postoji točno i riječi koda h, $h \neq g, d(h, w) = 4$
- $i = 1, 2, 3$

Vrijedi:

$$n_1 + n_2 + n_3 = \binom{12}{4} 11^4$$

$$n_2 + 2n_3 = 7\binom{12}{4}\binom{8}{4}$$

$$2n_3 = \binom{12}{4} 18$$

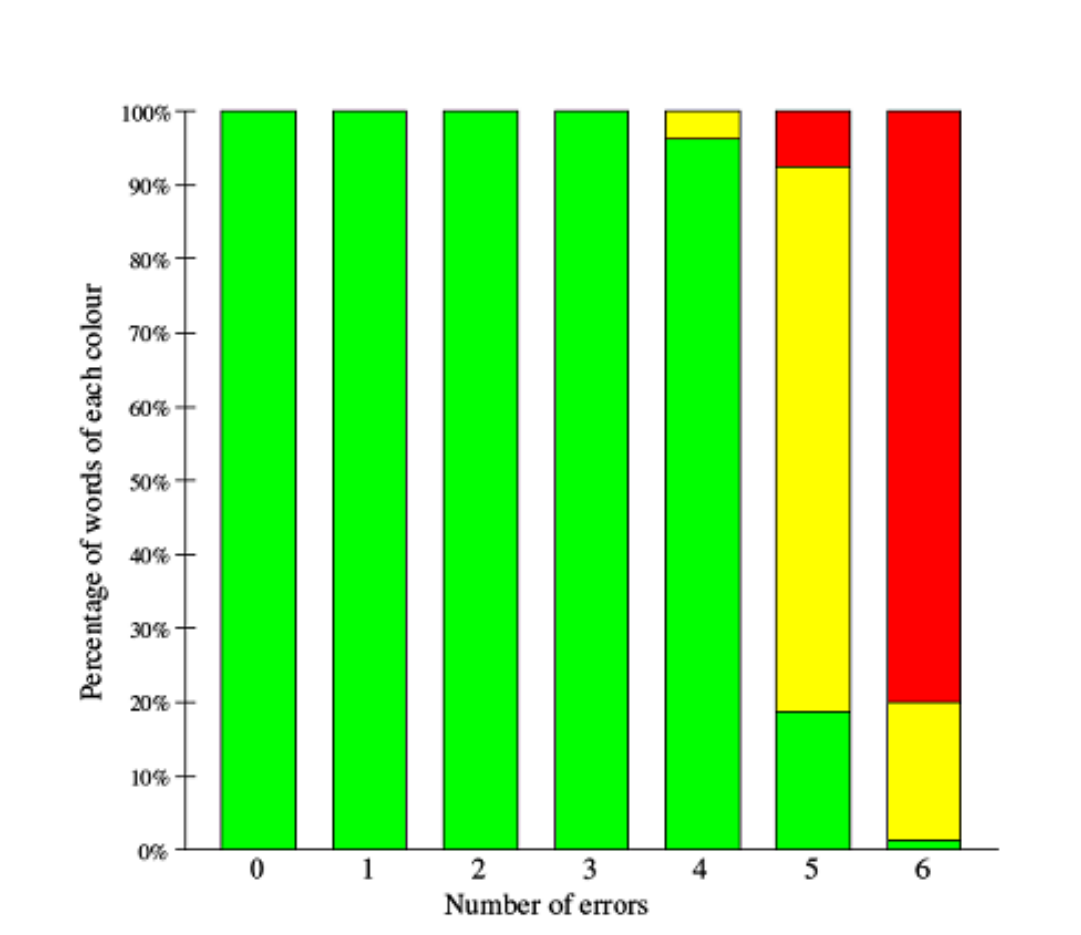
- $n_2 + n_3$ =broj konfiguracija koje se neće dekodirati jedinstveno
- $Q(M_{12}, 4) = \frac{n_2+n_3}{n_1+n_2+n_3} = \frac{481}{14641} \cong 0.032853$
- $P(M_{12}, 4) \cong 0.967147$

Q.E.D.

Riječi na udaljenosti 5,6 i 7

- poslana riječ g=(1 2 3 4 5 6 7 8 9 10 11 12)
- primljena riječ w=(4 1 2 3 6 6 7 8 9 10 11 12)
- $d(g, w) = 5$
- h= (4 1 2 3 6 7 8 5 9 10 11 12)
- $d(h, w) = 3$
- neispravno dekodiranje

Pretpostavke: Poslana riječ je *id*, greške se pojavljuju na prvih 5, 6 ili 7 pozicija. Pomoću programskog paketa GAP, koristeći UBB (za dobivanje najbližih susjeda) određene su boje riječi:



Teorem

Ne postoje zelene riječi na udaljenosti 7 od id.

Dokaz:

- $d(w, id) = 7$, $|w| \geq 5$.
- Postoje različiti 5-podskupovi $\{i_1, i_2, i_3, i_4, i_5\}$ i $\{j_1, j_2, j_3, j_4, j_5\} \in \Omega$ t.d. w ima različite simbole upravo na pozicijama: $i_1 \ldots i_5$ i $j_1 \ldots j_5$
- $g, h \in M_{12}$, slažu se s w na pozicijama $i_1 \ldots i_5$ i $j_1 \ldots j_5$
- $g \neq h$, w sigurno nije ZELENA
- $g = h$, w i h se slažu na najmanje 6 pozicija $\rightarrow$ w je CRVENA

Q.E.D.

Riječi na udaljenosti većoj od 8

- w-primljena riječ, $w \in H_{12}$
- $|w|$- broj pozicija u kojima se w i g ne razlikuju, $1 \leq |w| \leq 12$, $d(w, S_{12}) = 12 - |w|$
- $|w| \geq 5$, w se razlikuje na pozicijama $i_1 < i_2 < i_3 < i_4 < i_5$
- M_{12}- strogo 5-tranzitivna $\Rightarrow \exists! g \in M_{12}$ koji odgovara sa w u tim pozicijama
- $d(w, M_{12}) \leq 7$

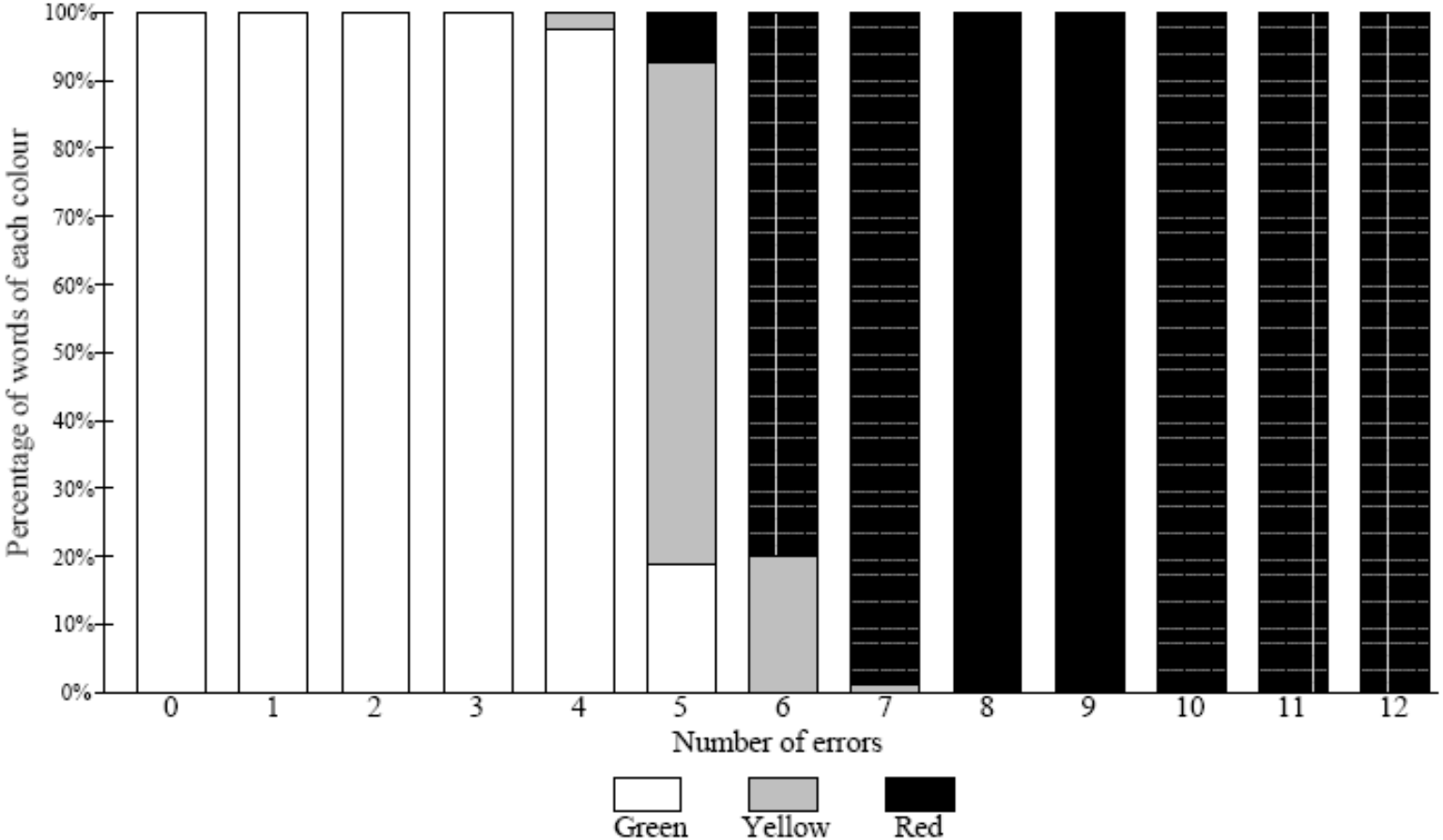
Što se događa kada je $|w| = m, m \leq 5$?

- promatramo pozicije $i_1 < ... < i_m$ te m-tranzitivnost od M_{12}
- $d(w, M_{12}) \leq 12 - m = d(w, S_{12}) \leq d(w, M_{12})$,
  $\Rightarrow d(w, M_{12}) = 12 - m$
- $m \leq 4$, stabilizator m-točaka nije trivijalan pa element iz M_{12} koji se slaže sa w nije jedinstven

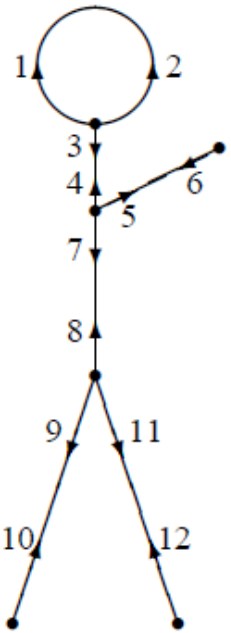
Zaključak: Za $i \leq 8$ ne postoje zelene riječi od M_{12}. Riječ w je žuta $\iff |w| = 12 - i$

Zaključak

i	$P(M_{12}, i)$	$Q(M_{12}, i)$	$R(M_{12}, i)$
0	1	0	0
1	1	0	0
2	1	0	0
3	1	0	0
4	0.967147	0.032853	0
5	0.187531	0.733147	0.079323
6	0.000043	0.195886	0.804072
7	0	0.004069	0.995931
8	0	0.000306	0.999694
9	0	0.000008	0.999992
10	0	0.000000	1.000000
11	0	0.000000	1.000000
12	0	0	1


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Percentage of words of each colour
0
1
2
3
4
5
6
7
8
9
10
11
12
Number of errors
Green
Yellow
Red

Generatori grupe M_{12}



Generators for M_{12}