Study of the NP-completeness of the Compact Table problem

NP-completeness comes to wargaming

J.-C. Dubacq   J.-Y. Moyen

Laboratory LIPN – UMR 7030
CNRS – Université Paris 13

Journées Automates Cellulaires 2008

Outline

1. Compact table problem
   - Random-choices tables
   - Formal description
   - Other applications
2. NP-Completeness
   - General case
   - Fixed amplitude case
   - Bounded number of results case
3. Conclusion and perspectives

Outline

1. Compact table problem
   - Random-choices tables
   - Formal description
   - Other applications

2. NP-Completeness
   - General case
   - Fixed amplitude case
   - Bounded number of results case

3. Conclusion and perspectives

Random tables

- Set of initial conditions
- Finite number of results
- $\rightarrow$ 2-D table:

	1	2	3	4	5	6	7	8	9	10
A	α	α	α	β	β	β	β	γ	γ	γ
B	α	β	β	β	γ	γ	δ	δ	δ	δ
C	α	γ	γ	γ	γ	γ	γ	δ	δ	δ
D	α	α	β	β	γ	γ	γ	γ	δ	δ

- Dimension reduction: A: $+0$, B: $+10$, C: $+20$, D: $+30$

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
α	α	α	β	β	β	β	γ	γ	γ	α	β	β	β	γ	γ	δ	δ	δ	δ

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
α	γ	γ	γ	γ	γ	γ	δ	δ	δ	α	α	β	β	γ	γ	γ	γ	δ	δ

- Tiny font, because very long!

Compact tables

Shortening the information

- Some lines may overlap partially...
- Known problem! This is the superword problem. NP-complete. But here, zero overlap!
- However, we can also shuffle around the lines...

Example

$$\alpha\beta\beta\beta\beta\alpha\alpha\gamma\gamma\gamma$$

- A: $+6 \rightarrow \alpha\alpha\alpha\beta\beta\beta\beta\gamma\gamma\gamma$
- B: $+0 \rightarrow \alpha\beta\beta\beta\gamma\gamma\delta\delta\delta\delta$
- C: $+12 \rightarrow \alpha\gamma\gamma\gamma\gamma\gamma\gamma\delta\delta\delta$
- D: $+9 \rightarrow \alpha\alpha\beta\beta\gamma\gamma\gamma\gamma\delta\delta$

Compact tables
Shortening the information

- Some lines may overlap partially...
- Known problem! This is the superword problem. NP-complete. But here, zero overlap!
- However, we can also shuffle around the lines...

Example

$$\delta\delta\delta\delta\gamma\gamma\alpha\beta\beta\beta\beta\alpha\alpha\gamma\gamma\gamma$$

- A: $+6 \rightarrow \alpha\alpha\alpha\beta\beta\beta\beta\gamma\gamma\gamma$
- B: $+0 \rightarrow \alpha\beta\beta\beta\gamma\gamma\delta\delta\delta\delta$
- C: $+12 \rightarrow \alpha\gamma\gamma\gamma\gamma\gamma\gamma\delta\delta\delta$
- D: $+9 \rightarrow \alpha\alpha\beta\beta\gamma\gamma\gamma\gamma\delta\delta$

Compact tables
Shortening the information

- Some lines may overlap partially...
- Known problem! This is the superword problem. NP-complete. But here, zero overlap!
- However, we can also shuffle around the lines...

Example

$$\delta\delta\delta\delta\gamma\gamma\alpha\beta\beta\beta\beta\alpha\alpha\gamma\gamma\gamma\gamma\delta\delta\delta\gamma\gamma$$

- A: $+6 \to \alpha\alpha\alpha\beta\beta\beta\beta\gamma\gamma\gamma$
- B: $+0 \to \alpha\beta\beta\beta\gamma\gamma\delta\delta\delta\delta$
- C: $+12 \to \alpha\gamma\gamma\gamma\gamma\gamma\gamma\delta\delta\delta$
- D: $+9 \to \alpha\alpha\beta\beta\gamma\gamma\gamma\gamma\delta\delta$

Compact tables

Shortening the information

- Some lines may overlap partially...
- Known problem! This is the superword problem. NP-complete. But here, zero overlap!
- However, we can also shuffle around the lines...

Example

$$\delta\delta\delta\delta\gamma\gamma\alpha\beta\beta\beta\beta\alpha\alpha\gamma\gamma\gamma\gamma\delta\delta\delta\gamma\gamma$$

- A: $+6 \to \alpha\alpha\alpha\beta\beta\beta\beta\gamma\gamma\gamma$
- B: $+0 \to \alpha\beta\beta\beta\gamma\gamma\delta\delta\delta\delta$
- C: $+12 \to \alpha\gamma\gamma\gamma\gamma\gamma\gamma\delta\delta\delta$
- D: $+9 \to \alpha\alpha\beta\beta\gamma\gamma\gamma\gamma\delta\delta$

Initial motivation

Real wargames, no computers involved

- Dimension reduction is crucial (easily-readable tables)
- Size of reduced table important (easily-learnable tables)

Initial motivation

Real wargames, no computers involved

- Dimension reduction is crucial (easily-readable tables)
- Size of reduced table important (easily-learnable tables)

Initial motivation

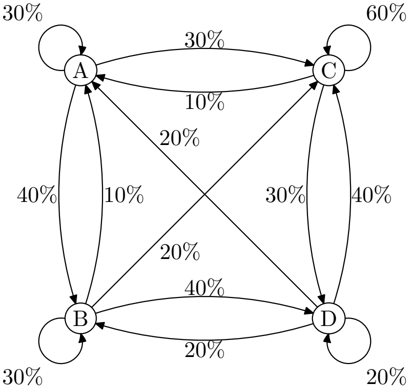
Real wargames, no computers involved

- Dimension reduction is crucial (easily-readable tables)
- Size of reduced table important (easily-learnable tables)

Probabilistic automata

Automata rule the world

If the set of initial conditions matches the set of outcomes, we get a probabilistic automaton. Efficient representation of probabilistic automata.



Outline

1. Compact table problem
   - Random-choices tables
   - Formal description
   - Other applications

2. NP-Completeness
   - General case
   - Fixed amplitude case
   - Bounded number of results case

3. Conclusion and perspectives

Formal Descriptions of Decision Problems

Compact Table problem

Instance Alphabet Σ, integer ℓ, set of words $S \subset \Sigma^\ell$, integer k

Answer YES if there exists a word $\tau \in \Sigma^k$ such that for any word $u \in S$, there exists a permutation σ and words v and w such that $\tau = v \cdot \sigma(u) \cdot w$, NO in all other cases.

Compact table of order ℓ

Instance Alphabet Σ, set of words $S \subset \Sigma^\ell$, integer k

Answer YES if there exists a word $\tau \in \Sigma^k$ such that for any word $u \in S$, there exists a permutation σ and words v and w such that $\tau = v \cdot \sigma(u) \cdot w$, NO in all other cases.

Formal Descriptions of Decision Problems

Compact Table problem

Instance Alphabet Σ, integer ℓ, set of words $S \subset \Sigma^\ell$, integer k

Answer YES if there exists a word $\tau \in \Sigma^k$ such that for any word $u \in S$, there exists a permutation σ and words v and w such that $\tau = v \cdot \sigma(u) \cdot w$, NO in all other cases.

Compact table of order ℓ

Instance Alphabet Σ, set of words $S \subset \Sigma^\ell$, integer k

Answer YES if there exists a word $\tau \in \Sigma^k$ such that for any word $u \in S$, there exists a permutation σ and words v and w such that $\tau = v \cdot \sigma(u) \cdot w$, NO in all other cases.

Outline

1. Compact table problem
   - Random-choices tables
   - Formal description
   - Other applications
2. NP-Completeness
   - General case
   - Fixed amplitude case
   - Bounded number of results case
3. Conclusion and perspectives

DNA single strand analysis

- Weights of A, C, G and T molecules are different
- Replicate an unknown DNA strand, cut it in small pieces
- Centrifugate and weight each small piece
- Infer the ACGT percentages
- Reconstruct the shortest possible single-strand DNA sequence possible with CT.
- Will not work, since it is NP-complete.

Outline

NP-Completeness and Hamiltonian Path

HP is definitely not Harry Potter!

Theorem

The Hamiltonian Path problem can be reduced to the Compact Table problem. Thus, the Compact Table problem is NP-complete.

Proof.

We define Σ to be the set $E \cup V$. Each vertex v is associated to a word τ_v of Σ^ℓ which is the set of edges adjacent to v (in no particular order) and padded (since G is not forced to be regular) by as many occurrences of v as deemed necessary. k is determined to be $n(\ell - 1) + 1$.

Being in NP is straightforward. □

NP-Completeness and Hamiltonian Path

HP is definitely not Harry Potter!

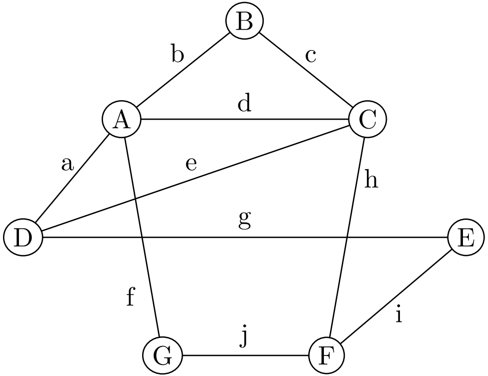
Theorem

The Hamiltonian Path problem can be reduced to the Compact Table problem. Thus, the Compact Table problem is NP-complete.

Proof.

We define Σ to be the set $E \cup V$. Each vertex v is associated to a word τ_v of Σ^ℓ which is the set of edges adjacent to v (in no particular order) and padded (since G is not forced to be regular) by as many occurrences of v as deemed necessary. k is determined to be $n(\ell - 1) + 1$.
Being in NP is straightforward. $\square$

Example of construction

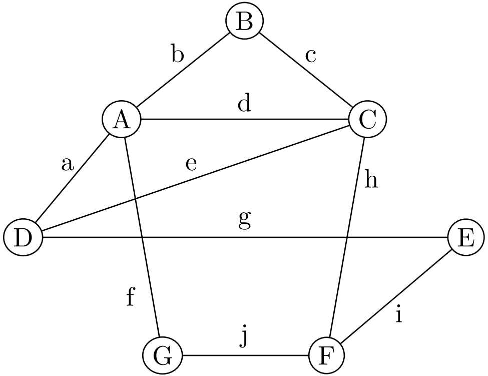


- $\tau_A = abdf$
- $\tau_B = BBbc$
- $\tau_C = cdeh$
- $\tau_D = Daeg$
- $\tau_E = EEgi$
- $\tau_F = Fhij$
- $\tau_G = GGjf$

Hamiltonian path *ABCDEFG* corresponding to the word (of length $n(\ell - 1) + 1 = 22$)

$$\tau = adf\mathbf{b}BB\mathbf{c}dh\mathbf{e}Da\mathbf{g}EE\mathbf{i}Fh\mathbf{j}GGf$$

Example of construction



- $\tau_A = abdf$
- $\tau_B = BBbc$
- $\tau_C = cdeh$
- $\tau_D = Daeg$
- $\tau_E = EEgi$
- $\tau_F = Fhij$
- $\tau_G = GGjf$

Hamiltonian path $ABCDEFG$ corresponding to the word (of length $n(\ell - 1) + 1 = 22$)

$$\tau = adf\mathbf{b}BB\mathbf{c}dh\mathbf{e}Da\mathbf{g}EE\mathbf{i}Fh\mathbf{j}GGf$$

NP-Completeness proof

HP exists $\rightarrow$ CT exists

Along the Hamiltonian path, edges can be collapsed, yields a word of length $n(\ell-1)+1$. *An edge is never used twice!*

CT exists $\rightarrow$ HP exists

Overlap only on edges, so τ describes a path in G. Because of length constraints, the path goes only once through each vertex.

#P-completeness remark

Transformation is not parcimonious (many possible permutations). But given an instance, number of solutions is either 0 or $(\ell-1)^2 \prod_{1\leq i\leq n} \frac{(\ell-2)!}{(\ell-d(i))!}$.

Therefore, the problem is also #P-complete, even though the reduction is not (and probably cannot) be parcimonious.

NP-Completeness proof

HP exists $\rightarrow$ CT exists

Along the Hamiltonian path, edges can be collapsed, yields a word of length $n(\ell - 1) + 1$. *An edge is never used twice!*

CT exists $\rightarrow$ HP exists

Overlap only on edges, so τ describes a path in G. Because of length constraints, the path goes only once through each vertex.

#P-completeness remark

Transformation is not parcimonious (many possible permutations). But given an instance, number of solutions is either 0 or $(\ell - 1)^2 \prod_{1 \leq i \leq n} \frac{(\ell - 2)!}{(\ell - d(i))!}$.

Therefore, the problem is also #P-complete, even though the reduction is not (and probably cannot) be parcimonious.

NP-Completeness proof

HP exists $\rightarrow$ CT exists

Along the Hamiltonian path, edges can be collapsed, yields a word of length $n(\ell - 1) + 1$. *An edge is never used twice!*

CT exists $\rightarrow$ HP exists

Overlap only on edges, so τ describes a path in G. Because of length constraints, the path goes only once through each vertex.

#P-completeness remark

Transformation is not parcimonious (many possible permutations). But given an instance, number of solutions is either 0 or $(\ell - 1)^2 \prod_{1 \leq i \leq n} \frac{(\ell - 2)!}{(\ell - d(i))!}$.

Therefore, the problem is also #P-complete, even though the reduction is not (and probably cannot) be parcimonious.

Outline

Amplitude $\ell > 2$

Work is already done

Theorem

The Compact Table problem of order $\ell > 2$ is NP-complete.

Proof.

Our reduction reduces HP of degree ℓ to CT of order ℓ. Since HP is still NP-complete with degree 3, done. $\square$

Amplitude $\ell > 2$

Work is already done

Theorem

The Compact Table problem of order $\ell > 2$ is NP-complete.

Proof.

Our reduction reduces HP of degree ℓ to CT of order ℓ. Since HP is still NP-complete with degree 3, done. $\square$

Amplitude $\ell = 2$

From Hamilton to Euler

Theorem

The Compact Table problem of order $\ell = 2$ is in P.

Proof.

Consider the results as vertices, initial conditions are edges. One can see easily that giving the smallest word containing all lines of the table is akin to describe a graph containing all edges of the graph. Details about unconnected components are in the paper. ▸ Proof details

Amplitude $\ell = 2$
From Hamilton to Euler

Theorem

The Compact Table problem of order $\ell = 2$ is in P.

Proof.

Consider the results as vertices, initial conditions are edges. One can see easily that giving the smallest word containing all lines of the table is akin to describe a graph containing all edges of the graph. Details about unconnected components are in the paper. Proof details $\square$

Outline

2-results case

Everything is so easy now

Theorem

CTP is solvable in linear time in case there are only two possible results.

Proof.

Use a sequence of m_1 times the first result “0” followed by m_2 times the second result “1”, where m_1 is the largest number of “0” for any initial condition and m_2 is the largest number of “1”. □

- Limited amplitude+limited outcomes, trivial (finite number of words).
- 2 results: CT method not efficient (prob. success enough)

2-results case

Everything is so easy now

Theorem

CTP is solvable in linear time in case there are only two possible results.

Proof.

Use a sequence of m_1 times the first result “0” followed by m_2 times the second result “1”, where m_1 is the largest number of “0” for any initial condition and m_2 is the largest number of “1”. □

- Limited amplitude+limited outcomes, trivial (finite number of words).
- 2 results: CT method not efficient (prob. success enough)

2-results case

Everything is so easy now

Theorem

CTP is solvable in linear time in case there are only two possible results.

Proof.

Use a sequence of m_1 times the first result "0" followed by m_2 times the second result "1", where m_1 is the largest number of "0" for any initial condition and m_2 is the largest number of "1". □

- Limited amplitude+limited outcomes, trivial (finite number of words).
- 2 results: CT method not efficient (prob. success enough)

3-results case

Unfinished fun on triangles

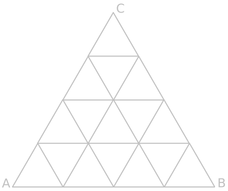
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). $\square$

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



3-results case
Unfinished fun on triangles

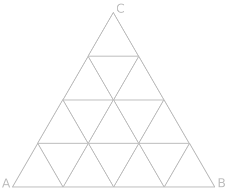
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). $\square$

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



3-results case

Unfinished fun on triangles

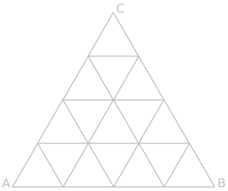
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). $\square$

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



3-results case

Unfinished fun on triangles

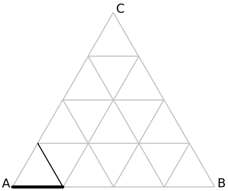
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). □

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



3-results case

Unfinished fun on triangles

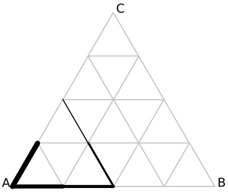
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). $\square$

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



3-results case

Unfinished fun on triangles

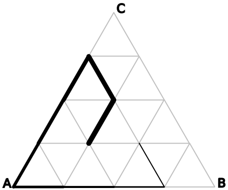
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). □

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



3-results case
Unfinished fun on triangles

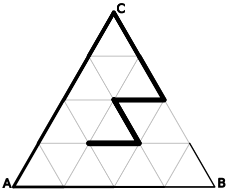
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). $\square$

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



3-results case

Unfinished fun on triangles

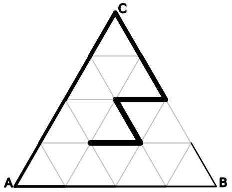
Number of possible words of amplitude ℓ: $\binom{\ell+k-1}{\ell}$.

Proof.

Number of occurrences of each result, including 0, in bijection with words on $\{x, y\}$ of length $\ell + k - 1$ with $k - 1$ y letters separating runs of x (run i is the number of occurrence of result i). □

Superword of size $\binom{\ell+k-1}{\ell} + \ell - 1$ containing all permutations? $\rightarrow$ **Open problem**.

- 3 outputs, amplitude 1: *abc*.
- Amplitude 2: *caabbcc*.
- $\ell = 3$: *abcccaaabbbc*.
- $\ell = 4$: *abacbcbccccaaaabbbbc*.
- Recurrence? 4-results? Beyond?



Summary and Open problems

Stuff we couldn't do in time

Answered Questions

- Works even if words of different lengths;
- Permutations do not help;
- They may even make things harder;
- Some things remain simple.

Open Questions

- The superword problem is known to be NP-hard but approximable;
- For Compact table: not clear. Heuristics may apply, but ratio is not a constant.
- Restriction to 3 or more results: still open. 3 results may be possible (winding out from the inside to the outside).

Summary and Open problems

Stuff we couldn't do in time

Answered Questions

- Works even if words of different lengths;
- Permutations do not help;
- They may even make things harder;
- Some things remain simple.

Open Questions

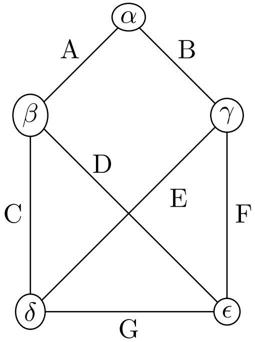
- The superword problem is known to be NP-hard but approximable;
- For Compact table: not clear. Heuristics may apply, but ratio is not a constant.
- Restriction to 3 or more results: still open. 3 results may be possible (winding out from the inside to the outside).

This slide intentionally left blank

Some details on proof in case $\ell = 2$

You probably asked for it!

	α	β	γ	δ	ϵ
A	1	1			
B	1		1		
C		1		1	
D		1			1
E			1	1	
F			1		1
G				1	1



Go back

Some details on proof in case $\ell = 2$ (cont.)

We separate in A (connected components with vertices of odd degree) and the other ones (B). We want to reach $(a, b, n/2) = (1, 0, 1)$.

- α Adding one edge going from one component to itself: either $[0, 0, 1]$ between two even vertices, $[0, 0, 0]$ between an even vertex and an odd vertex, $[0, 0, -1]$ between two odd vertices. There is a special case for the last one: the move could also be $[-1, 1, -1]$.
- β Adding one edge between two components of A: $[-1, 0, 1]$ between two even vertices, $[-1, 0, 0]$ between an even vertex and an odd vertex, $[-1, 0, -1]$ between two odd vertices.

Some details on proof in case $\ell = 2$ (cont.)

γ Adding one edge between one component of A and one of B: $[0, -1, 1]$ if the vertex in the component in A was of even degree, $[0, -1, 0]$ otherwise. There is always an even number of odd-degree vertices in a component, so a never decreases this way.

δ Adding one edge between two components of B: $[1, -2, 1]$ (always).

- If $a = 0$, then $n = 0$ and $b > 1$. The transformation $\delta\gamma^{b-2}$ leads us to the final state and is of length $b + n - 1 = b - 1$.
- If $a > 0$, then transformation $\beta^{a-1}\gamma^{b}\alpha^{n-a}$ leads us to the final state and is of length $b + n - 1$.
- In each case, there is only one subcase that decreases $b + n$; there may be some choice for the exact edge to be added.

Go back