FIT 2016

直径3のグラフにおける平均最短経路長の近似

清水 伸高
(東京大学)

森 立平
(東京工業大学)

0. Introduction

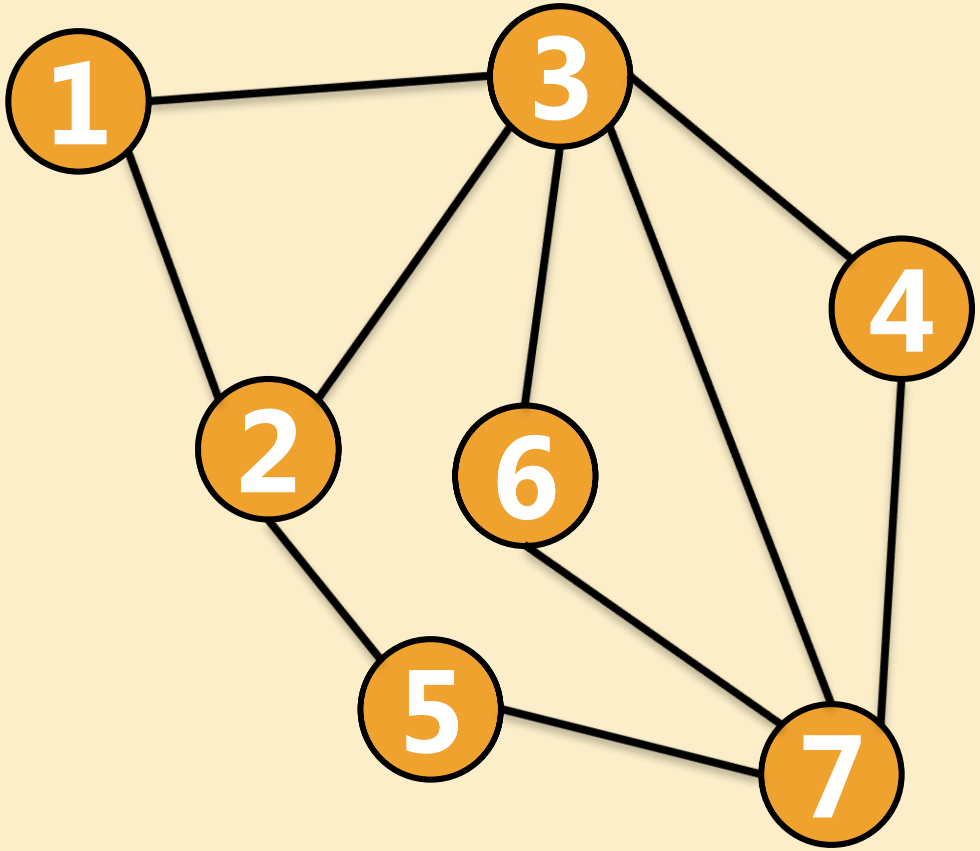
Average Shortest Path Length (ASPL)

Definition

- $G = (V, E)$ ：単純無向グラフ
- $\mathrm{dist}(i, j)$ ：*i*-*j* 間の平均距離
- $\mathrm{ASPL} = \mathrm{ASPL}(G) := \sum_{i \neq j \in V} \mathrm{dist}(i, j) \Big/ \binom{n}{2}$
- $\mathrm{diam}(G) := \max_{i \neq j \in V} \mathrm{dist}(i, j)$ ：直径

Average Shortest Path Length (ASPL)

	1	2	3	4	5	6	7
1	0	1	1	2	2	2	2
2		0	1	2	1	2	2
3			0	1	2	1	1
4				0	2	2	1
5					0	2	1
6						0	1
7							0



$$\mathrm{ASPL} = \frac{1 \times 10 + 2 \times 11}{21} = 1.523\ldots \qquad \mathrm{diam}(G) = 2$$

Our Problem

Problem

min: $\mathrm{ASPL}(G)$

s.t. G は n 頂点 d-正則グラフ

(d-正則グラフ: 全ての次数 = d)

Related Problems

- Order / Degree Problem

unexplored

min: $\mathrm{diam}(G)$

s.t. G は n 頂点 d-正則グラフ

- Degree / Diameter Problem

explored
(graph theory)

max: n (頂点数)

s.t. G は d-正則グラフ

$diam(G) = D$

Contributions

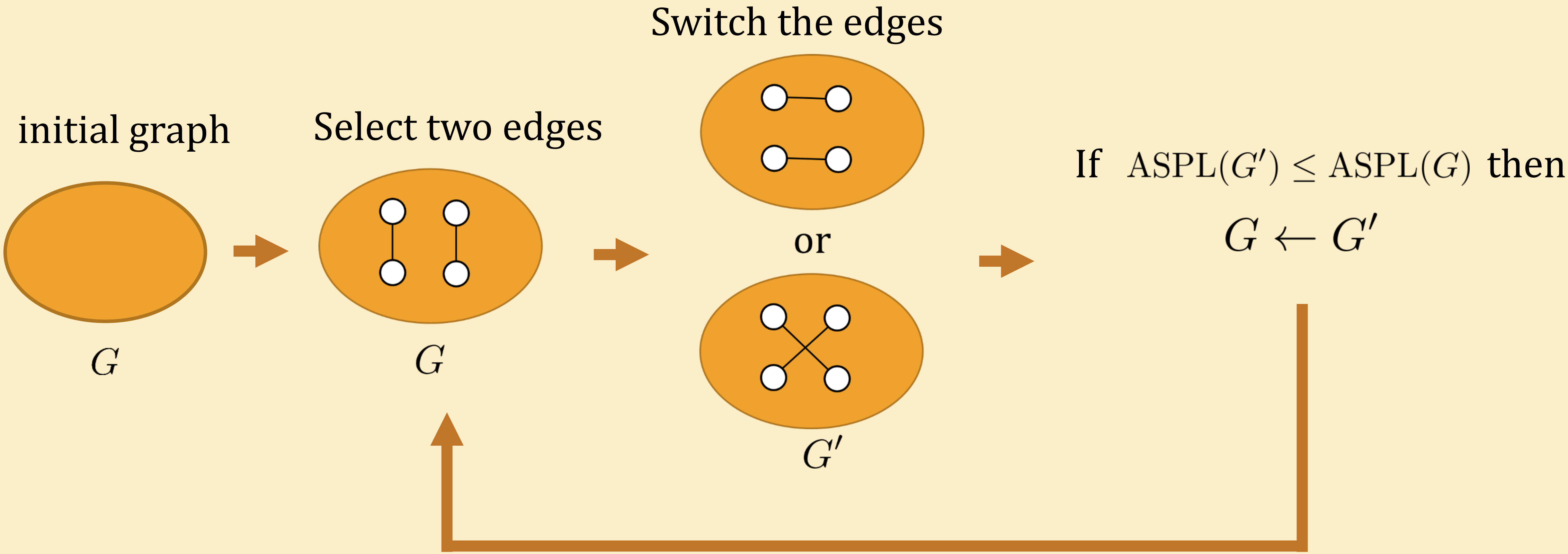
- 直径3のグラフのASPLに関する上界・下界・等式を示した (主定理)
- 直径3の小ASPLグラフを求める Local search algorithm の提案

1. Naïve Algorithm

Local Search

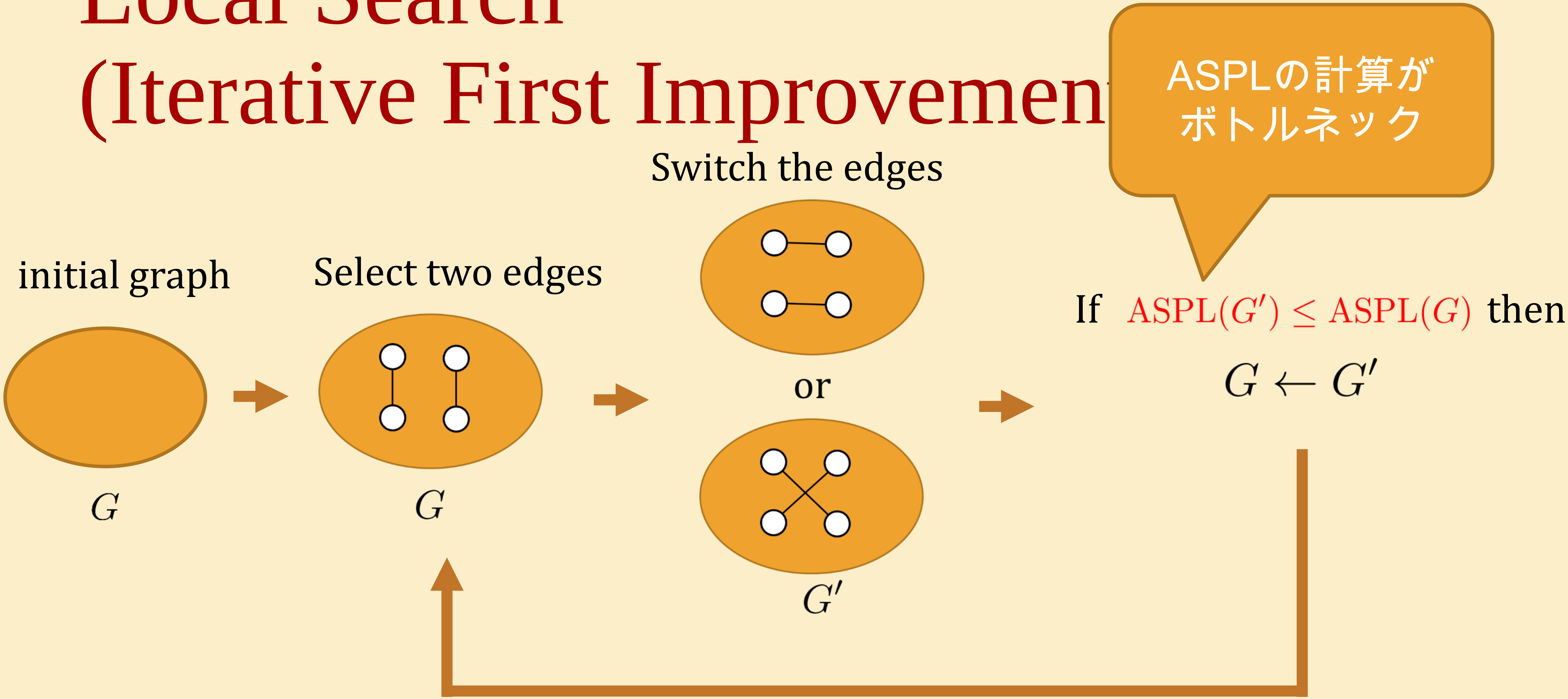
- 初期グラフを作る。
- 改善していく。
- なんらかの条件を満たしたら終了する。

Local Search (Iterative First Improvement)



・Gが更新できなくなったら終了する。

Local Search (Iterative First Improvement)



- If G cannot be improved for any edge pairs, terminate

Time Complexity

n = 頂点数
d = 次数

- Calculating $\mathrm{ASPL}(G) \cdots O(VE) = O(n^2 d)$

Time Complexity

n = 頂点数
d = 次数

・Calculatin

遅い！！

Time Complexity

n = 頂点数
d = 次数

- Calculating $\mathrm{ASPL}(G) \cdots O(VE) = O(n^2 d)$

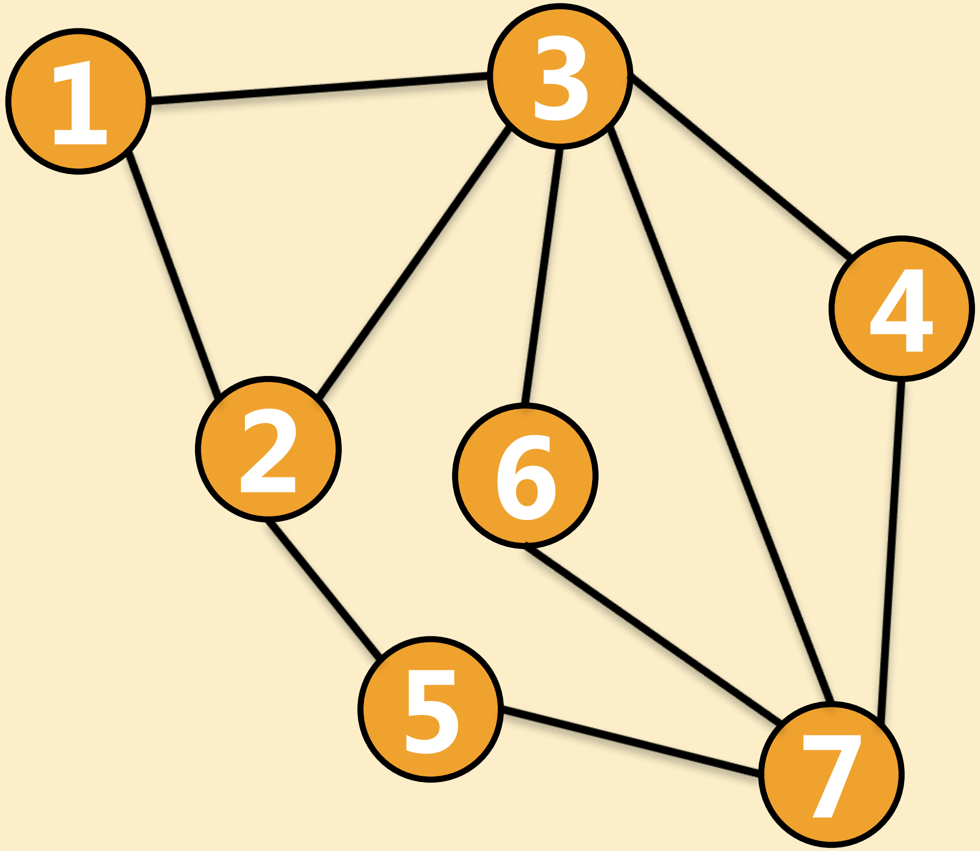
グラフの変化はちょっとだけ

→ ASPLの差分を上手く計算したい

→ 適当な配列を使ってできないか？？

Average Shortest Path Length (ASPL)

	1	2	3	4	5	6	7
1	0	1	1	2	2	2	2
2		0	1	2	1	2	2
3			0	1	2	1	1
4				0	2	2	1
5					0	2	1
6						0	1
7							0



$$\mathrm{ASPL} = \frac{1 \times 10 + 2 \times 11}{21} = 1.523\ldots \qquad \mathrm{diam}(G) = 2$$

ASPL Calculation

n = 頂点数
d = 次数

$\mathrm{ASPL}(G)$ ：頂点ペアの距離の総和が分かれば良い

→ 距離1のペア数、距離2のペア数,...が分かれば計算できる。

辺数 = nd/2

Case of Diameter 3

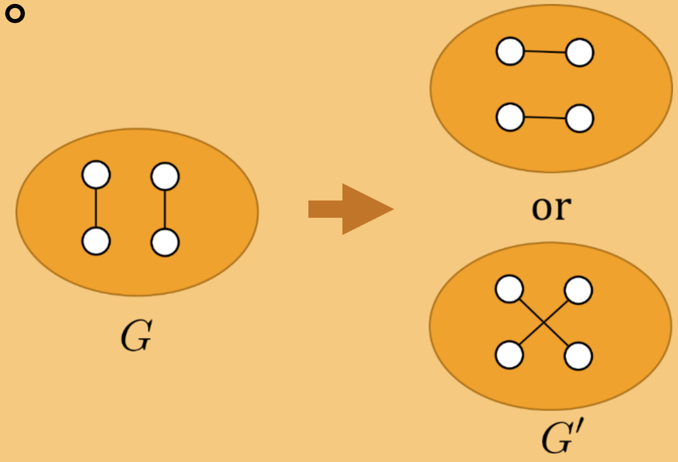
n = 頂点数
d = 次数

辺の入れ替え(switch)前後のグラフの直径が3とする。

距離1,2,3のペア数が分かれば良い。

全体のペア数$=\frac{n(n-1)}{2}$，距離1のペア数$=\frac{nd}{2}$

よって距離2のペア数だけ分かれば良い。



上手く配列を持てば$O(d)$でASPLの差分を計算可能

Time Complexity

n = 頂点数
d = 次数

- $\mathrm{ASPL}(G)$ の計算 $\cdots$ $O(VE) = O(n^2 d)$
- 直径 3 の場合 $\cdots$ $O(d)$ 時間で差分 ($\mathrm{ASPL}(G') - \mathrm{ASPL}(G)$)

Time Complexity

n = 頂点数
d = 次数

依然、遅い！！

・直径

Time Complexity

n = 頂点数
d = 次数

・直径

近似で評価しよう！

Our Target

n = 頂点数
d = 次数

- ほぼ全てのグラフが直径3となるような (n, d) について考える
- 実際には $n \approx d^2$ (n が d^2 に近い)

2. Our Idea

Diameter 3 (again)

n = # of nodes
d = degree

辺の入れ替え(switch)前後のグラフの直径が3とする。

距離1,2,3のペア数が分かれば良い。

全体のペア数 $= \frac{n(n-1)}{2}$, 距離1のペア数$= \frac{nd}{2}$

よって距離2のペア数だけ分かれば良い。

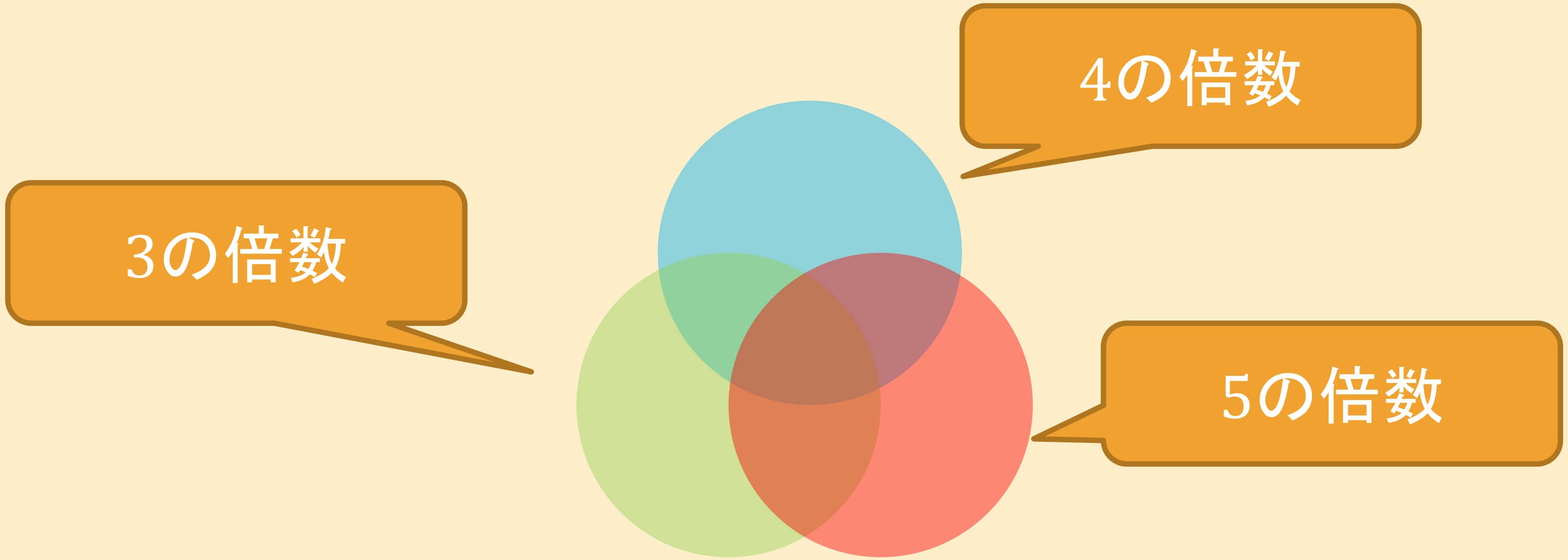
Our goal

距離2以下のペア数を求める！！！

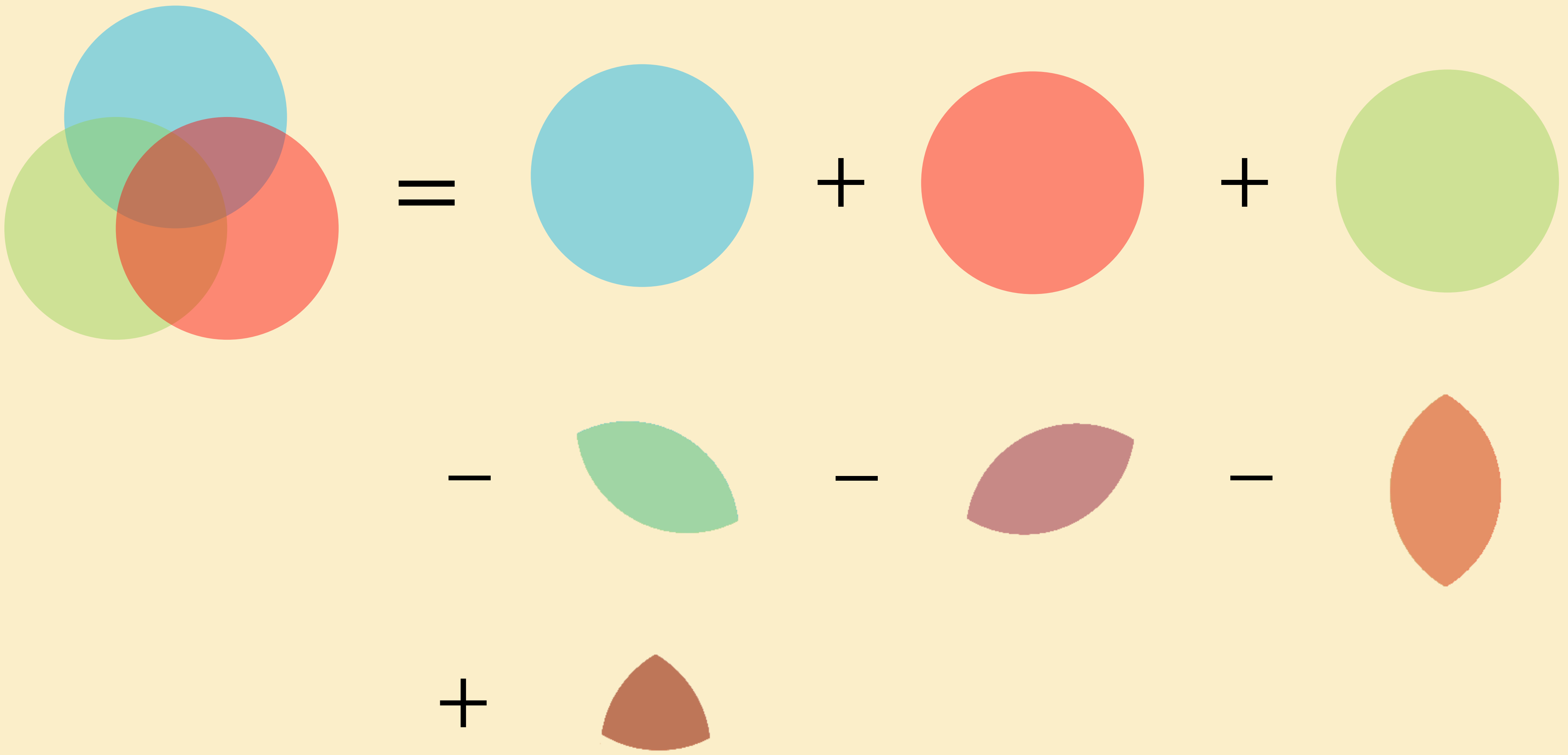
Inclusion Exclusion Principle

Exercise：

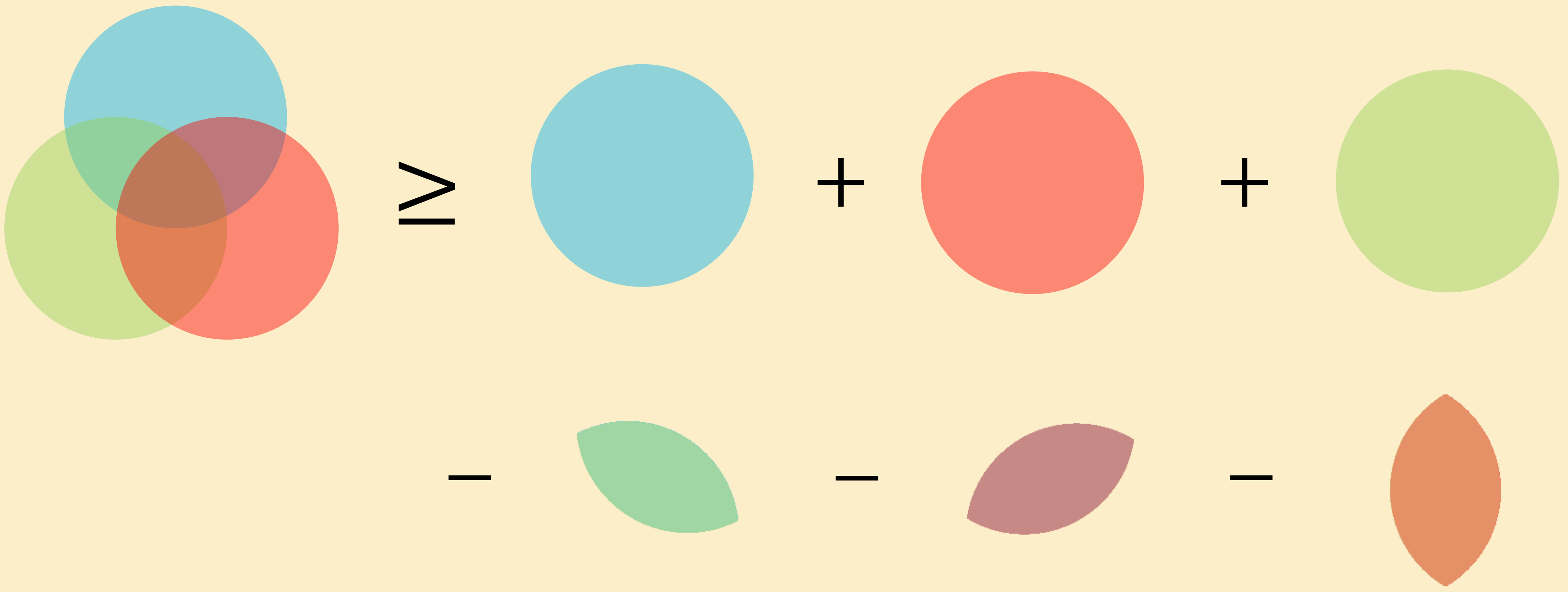
1~120 の中で、「3の倍数 or 4の倍数 or 5の倍数」である数をいくつ？



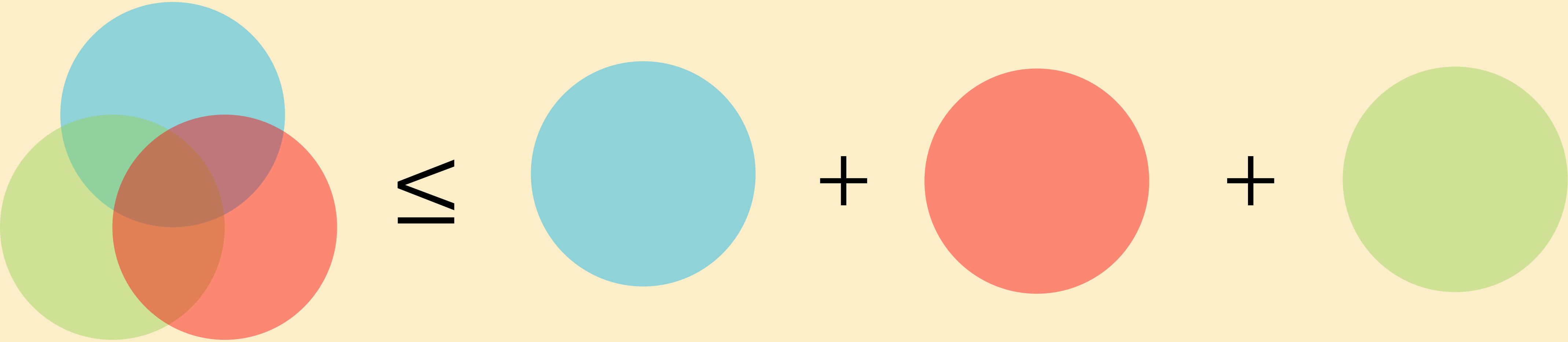
Inclusion Exclusion Principle



Bonferroni Inequality



Bonferroni Inequality



Key Point

和集合のサイズは

いくつかの “**共通部分**” のサイズから

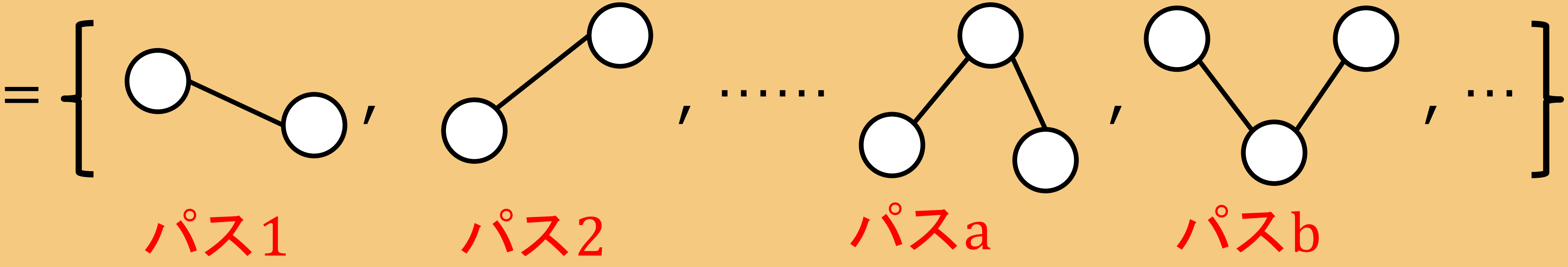
分かる。

Goal (again)

距離2<u>以下</u>のペア数を求める！！！

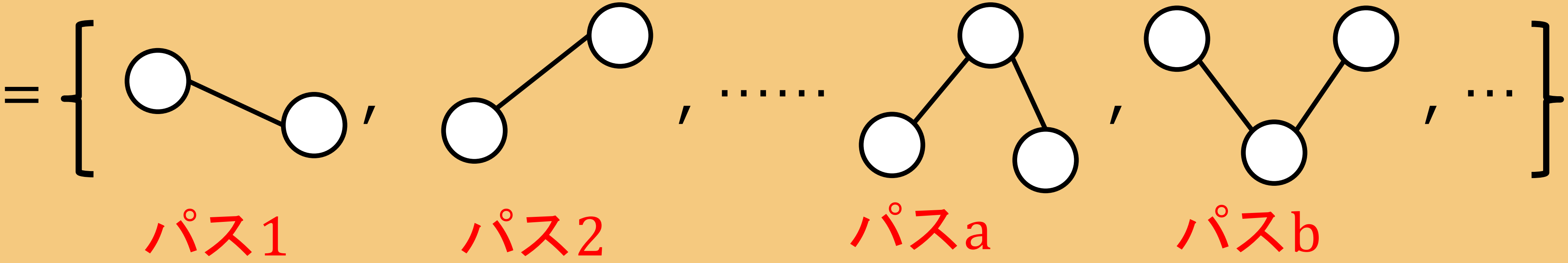
Pairs of distance at most 2

{距離2以下のパス}



Pairs of distance at most 2

{距離2以下のパス}



{距離2以下のペア}

= パス1の端点 ∪ パス2の端点 ∪ ⋯

Pairs of distance at most 2

{距離2以下のパス}

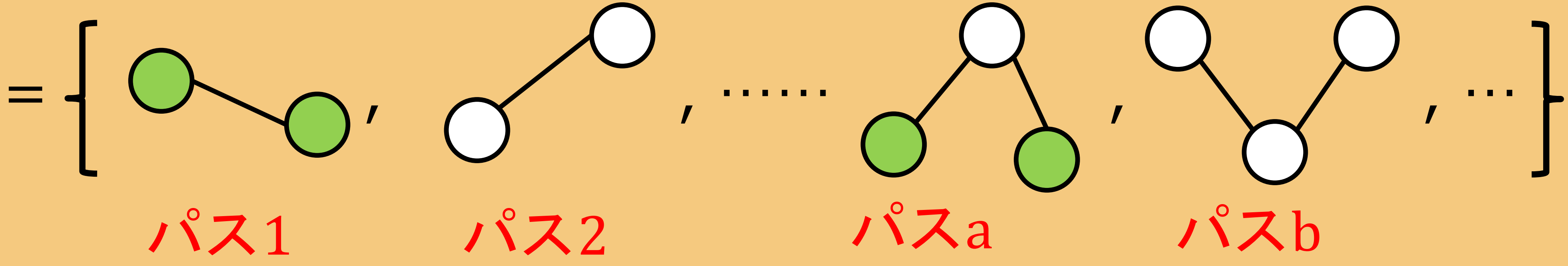
包除原理

{距離2以下のペア}

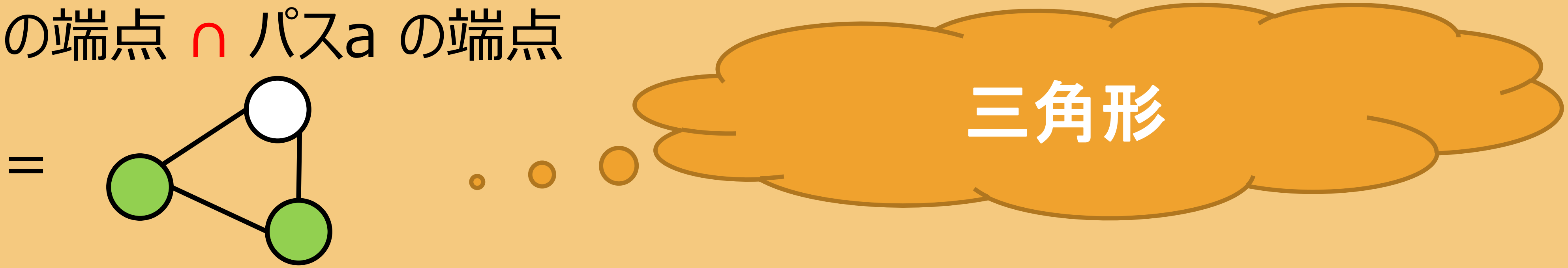
= パス1の端点 ∪ パス2の端点 ∪ …

Pairs of distance at most 2

{距離2以下のパス}

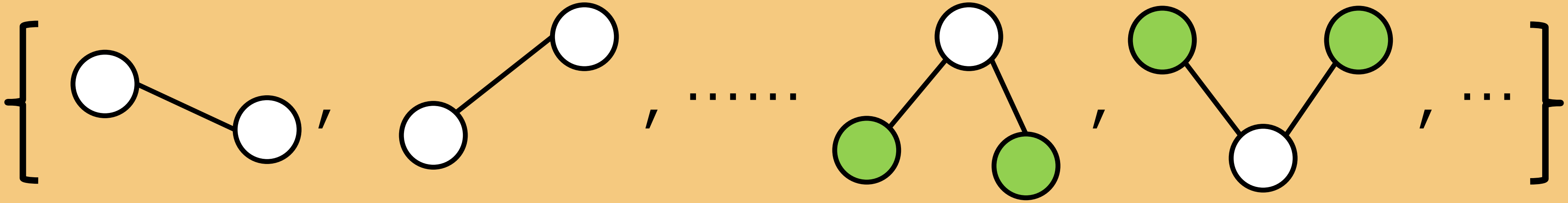


パス1 の端点 ∩ パスa の端点



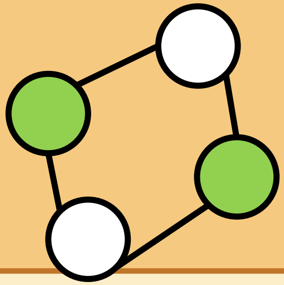
Pairs of distance at most 2

{距離2以下のパス}

= {  , … }

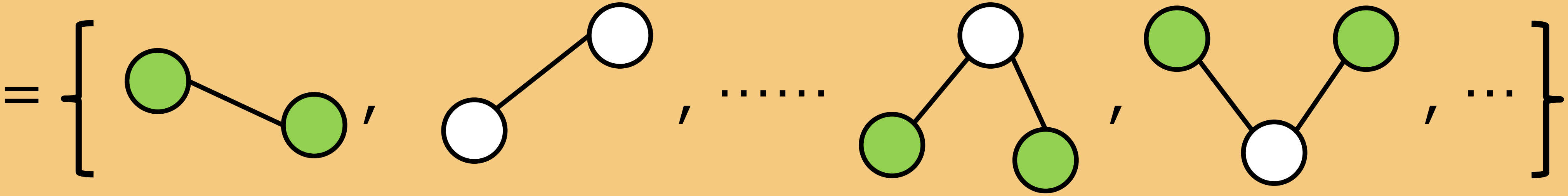
パス1　パス2　パスa　パスb

パスa の端点 $\cap$ パスb の端点

= 

Pairs of distance at most 2

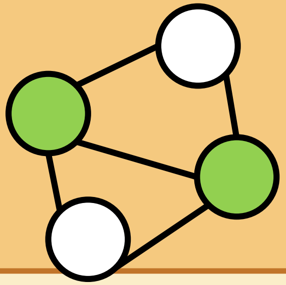
{距離2以下のパス}



パス1　パス2　パスa　パスb

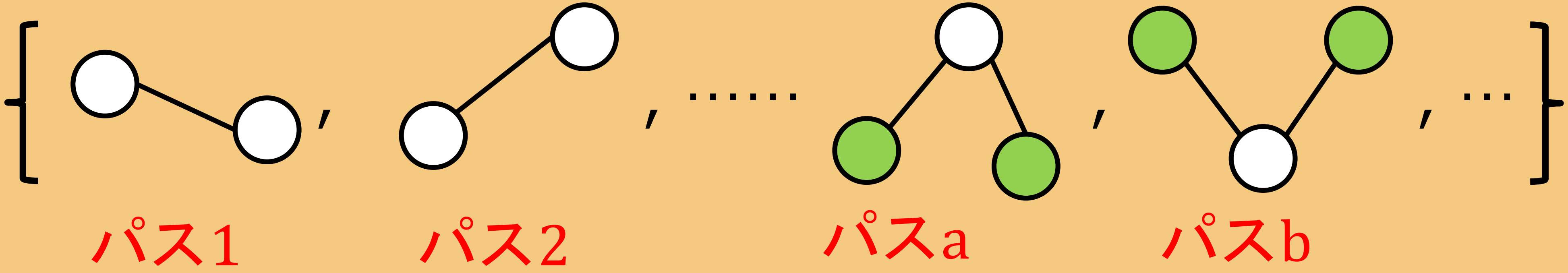
パス1 の端点 ∩ パスa の端点 ∩ パスb の端点

=

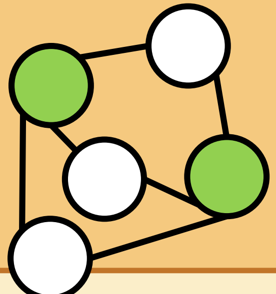


Pairs of distance at most 2

{距離2以下のパス}

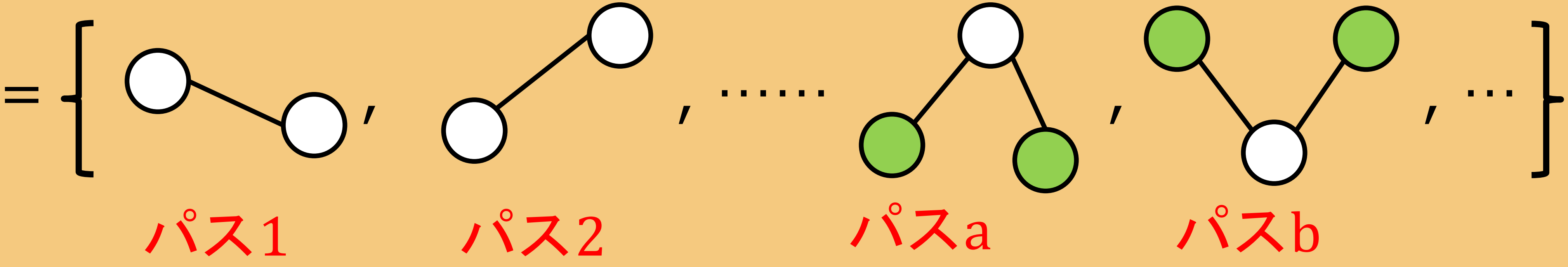
= 

パスa の端点 ∩ パスb の端点 ∩ パスc の端点

= 

Pairs of distance at most 2

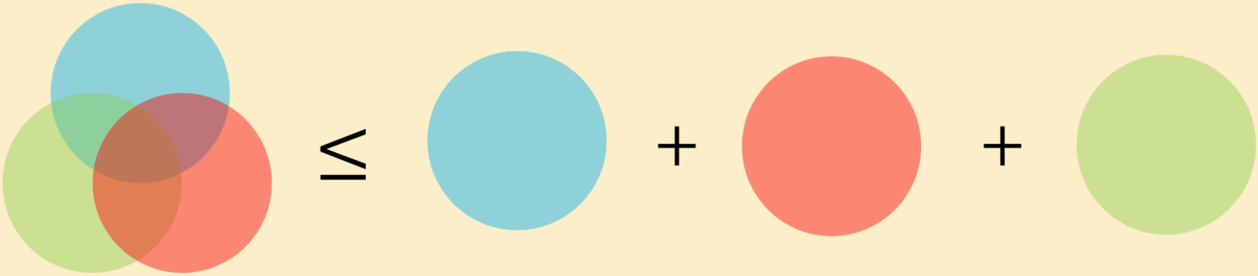
{距離2以下のパス}



距離2以下のペア数を"図形の個数"によって表せる！！

→ ASPLを"図形の個数"によって表せる！！

Moore Bound



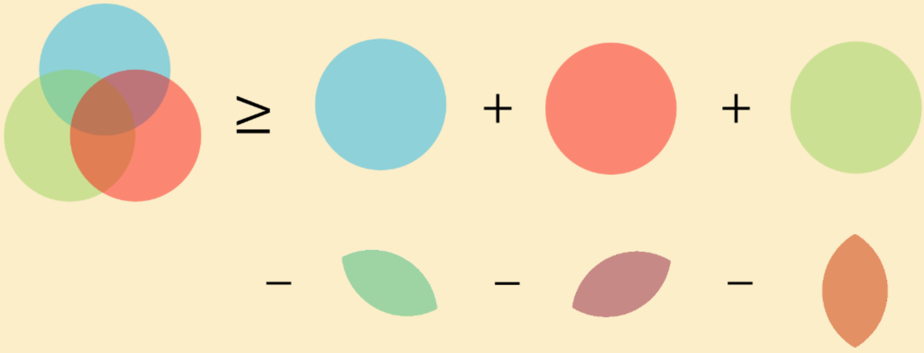
Fact (the Moore Bound)

$G=(V,E)$: d-正則, n 頂点, 直径3

$$\binom{n}{2}\cdot \mathrm{ASPL}(G) \geq 3\binom{n}{2} - \frac{nd}{2} - nd^2$$

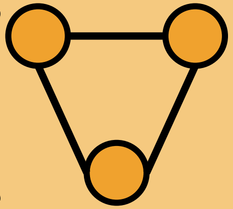
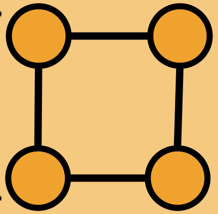
共通部分は考えていない

ASPL Upper Bound



Theorem

$G=(V,E)$: d-正則, n 頂点, 直径3

$\binom{n}{2}\cdot \mathrm{ASPL}(G) \le 3\binom{n}{2}-\frac{nd}{2}-nd^2+3\#\{$  $\}+2\#\{$  $\}$

2個の共通部分まで考えている

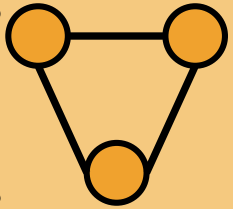
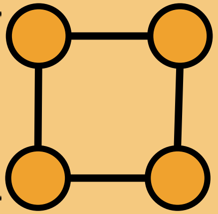
(この上界を近似値として見ることも出来る)

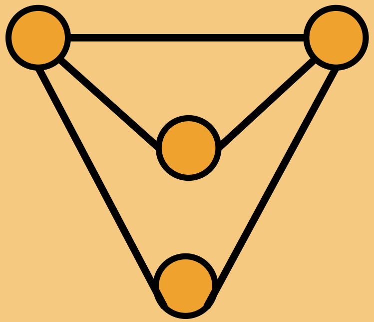
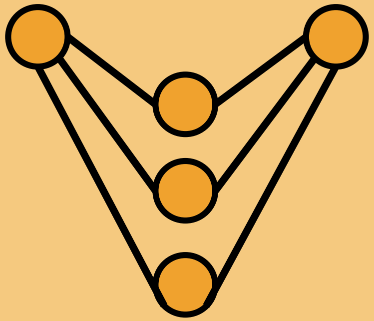
ASPL Lower Bound

Theorem

3個の共通部分まで考えている

$G = (V, E)$: d-正則, n 頂点, 直径3

$\binom{n}{2} \cdot \mathrm{ASPL}(G) \geq 3\binom{n}{2} - \frac{nd}{2} - nd^2 + 3\#\Big\{$  $\Big\} + 2\#\Big\{$  $\Big\}$

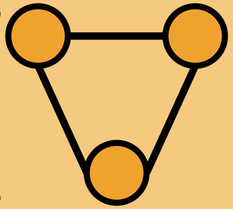
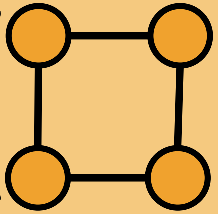
$-\#\Big\{$  $\Big\} - \#\Big\{$  $\Big\}$

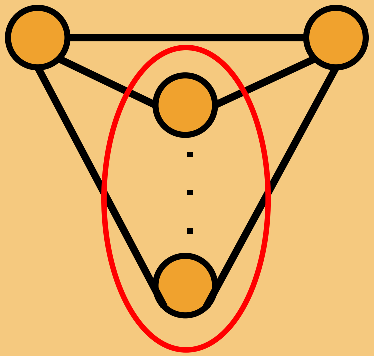
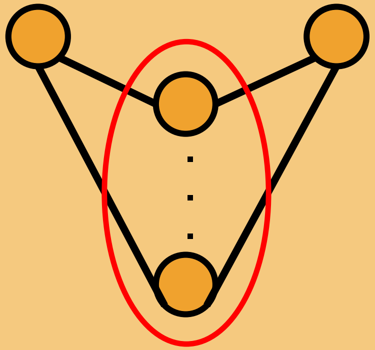
(この下界を近似値として見ることも出来る)

ASPL Characterization

Theorem

$G = (V, E)$: d-正則, n 頂点, 直径3

$\binom{n}{2} \cdot \mathrm{ASPL}(G) = 3\binom{n}{2} - \frac{nd}{2} - nd^2 + 3\#\{$  $\} + 2\#\{$  $\}$

$+ \sum_{m=3}^{n-1} (-1)^m \Big[\#\{$  $\} + \#\{$  $\} \Big]$

$m-1$ nodes　　m nodes

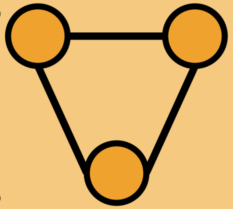
3. Proposed Algorithm

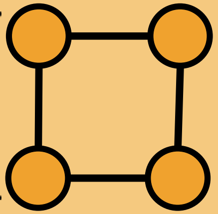
Proposed Algorithm

以下の評価関数でグラフを評価する:

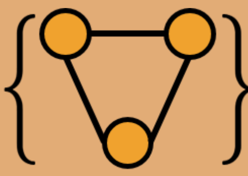
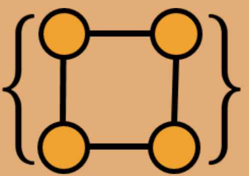
$$f(G) = 3\triangle + 2\square$$

ここで

$\triangle = \#\{$  $\}$

$\square = \#\{$  $\}$

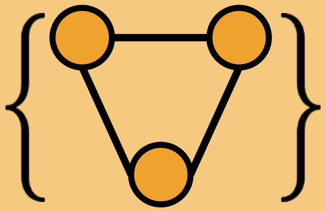
Proposed Algorithm

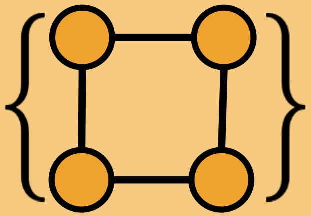
$\binom{n}{2} \cdot ASPL(G) \leq 3 - \frac{nd}{2} - nd^2 + 3\#$  $+2\#$ 

以下の評価関数でグラフを

$$f(G) = 3\triangle + 2\square$$

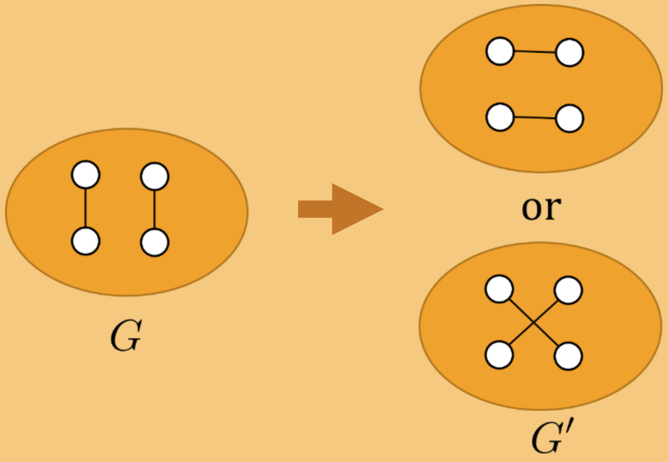
ここで

$\triangle = \#$ 

$\square = \#$ 

Time Complexity

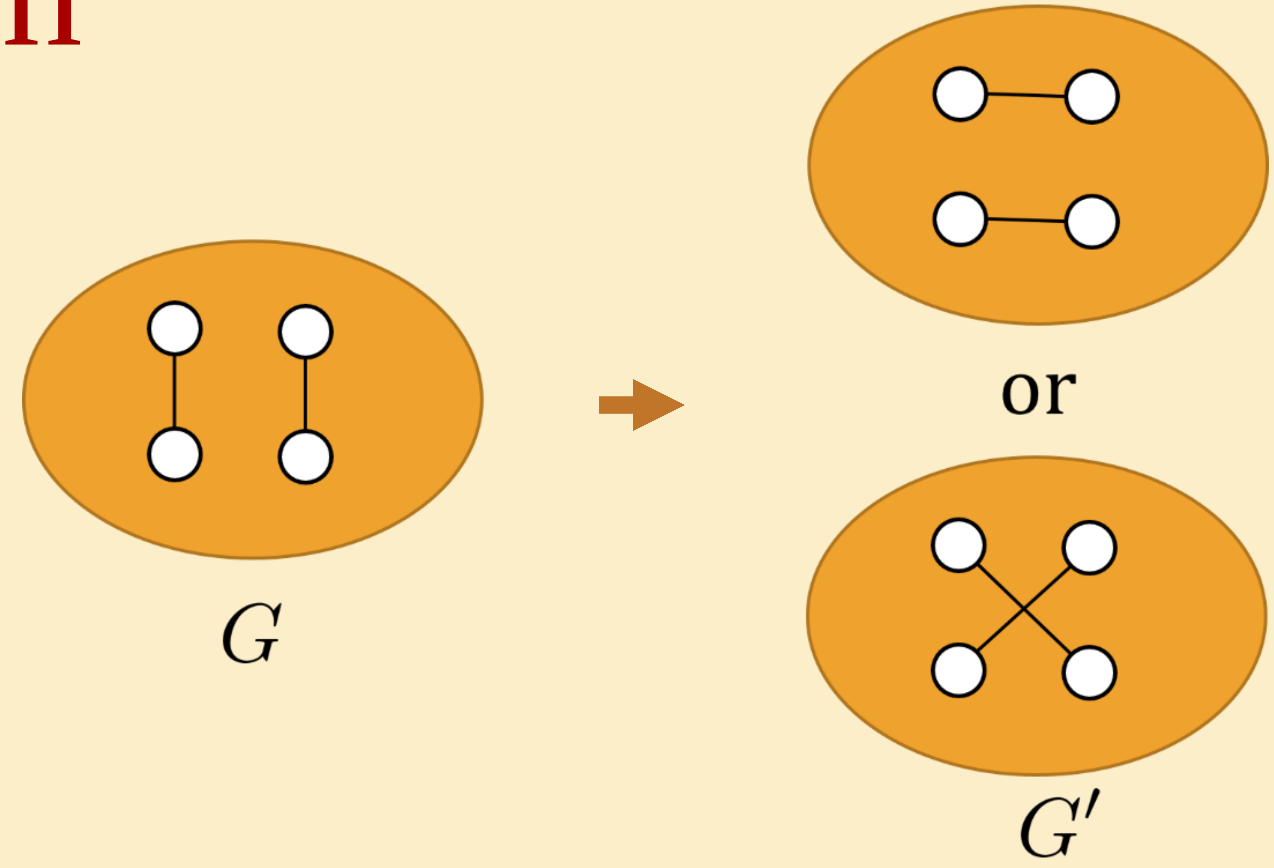
- $f(G') - f(G)$: $O(1)$ 時間で計算できる



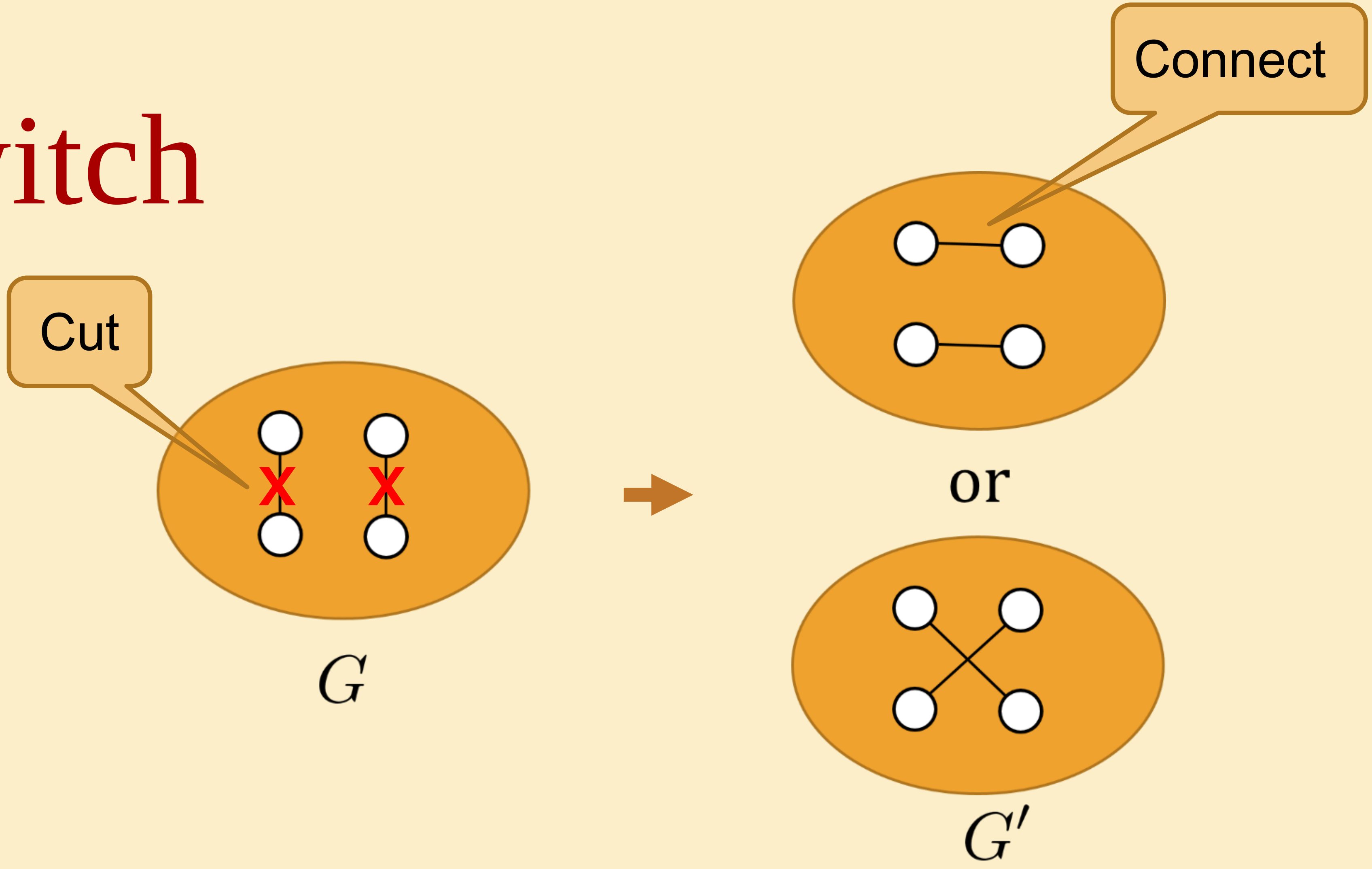
以下の配列を使う:

- T2$[i][j]$:= # of i-j paths of length 2
- T3$[i][j]$:= # of non-backtracking i-j paths of length 3

Switch



Switch

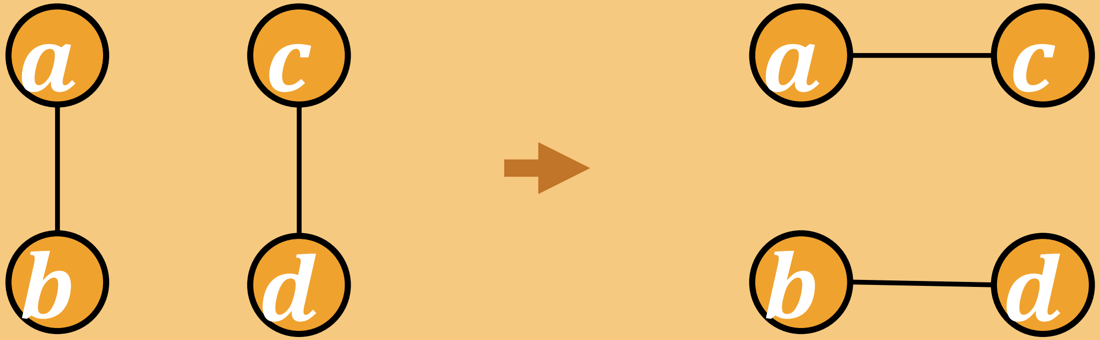


G $\rightarrow$ G'

この操作による三角形/四角形の増減をカウントしたい

Evaluation Algorithm

G' を以下のswitchによって得られたグラフとする：

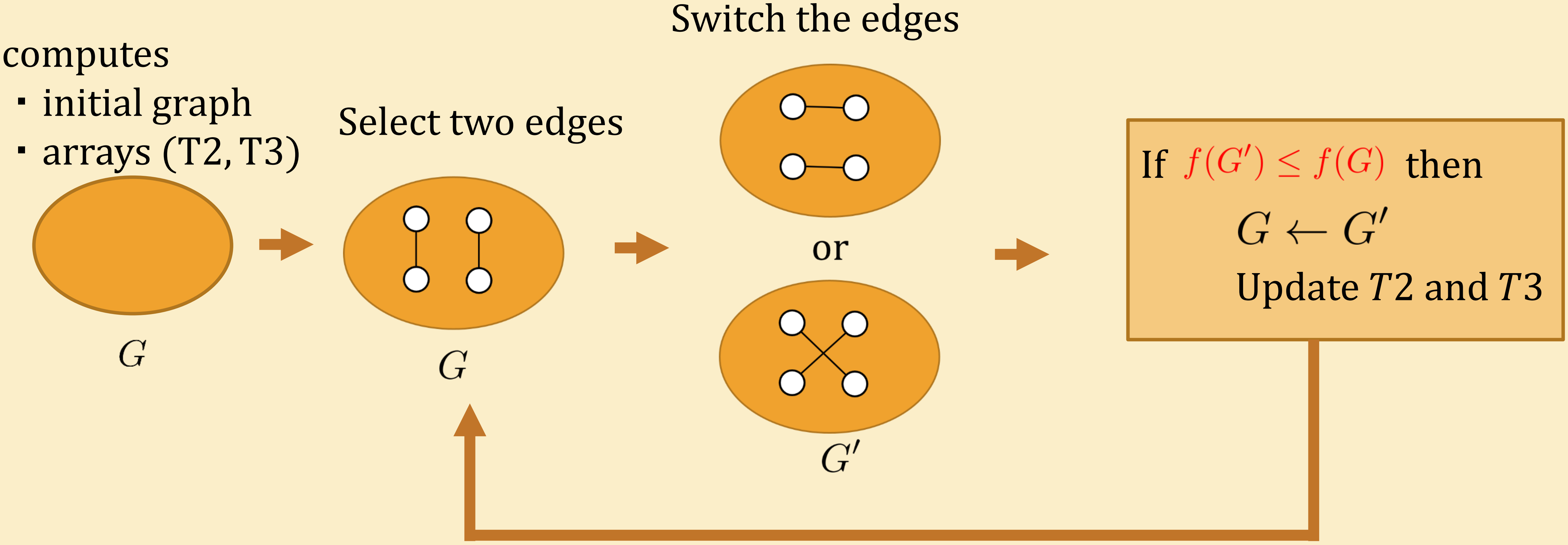


すると、

$$
\begin{aligned}
f(G') - f(G) = 3(&-\mathrm{T2}[a][b] - \mathrm{T2}[c][d] + \mathrm{T2}[a][c] + \mathrm{T2}[b][d] \\
&- 2(\mathrm{T1}[a][d] + \mathrm{T1}[b][c])) \\
&+2(-\mathrm{T3}[a][b] - \mathrm{T3}[c][d] + \mathrm{T3}[a][c] + \mathrm{T3}[b][d] \\
&- 2(\mathrm{T2}[a][d] + \mathrm{T2}[b][c] - \mathrm{T1}[a][d]\mathrm{T1}[b][c]))
\end{aligned}
$$

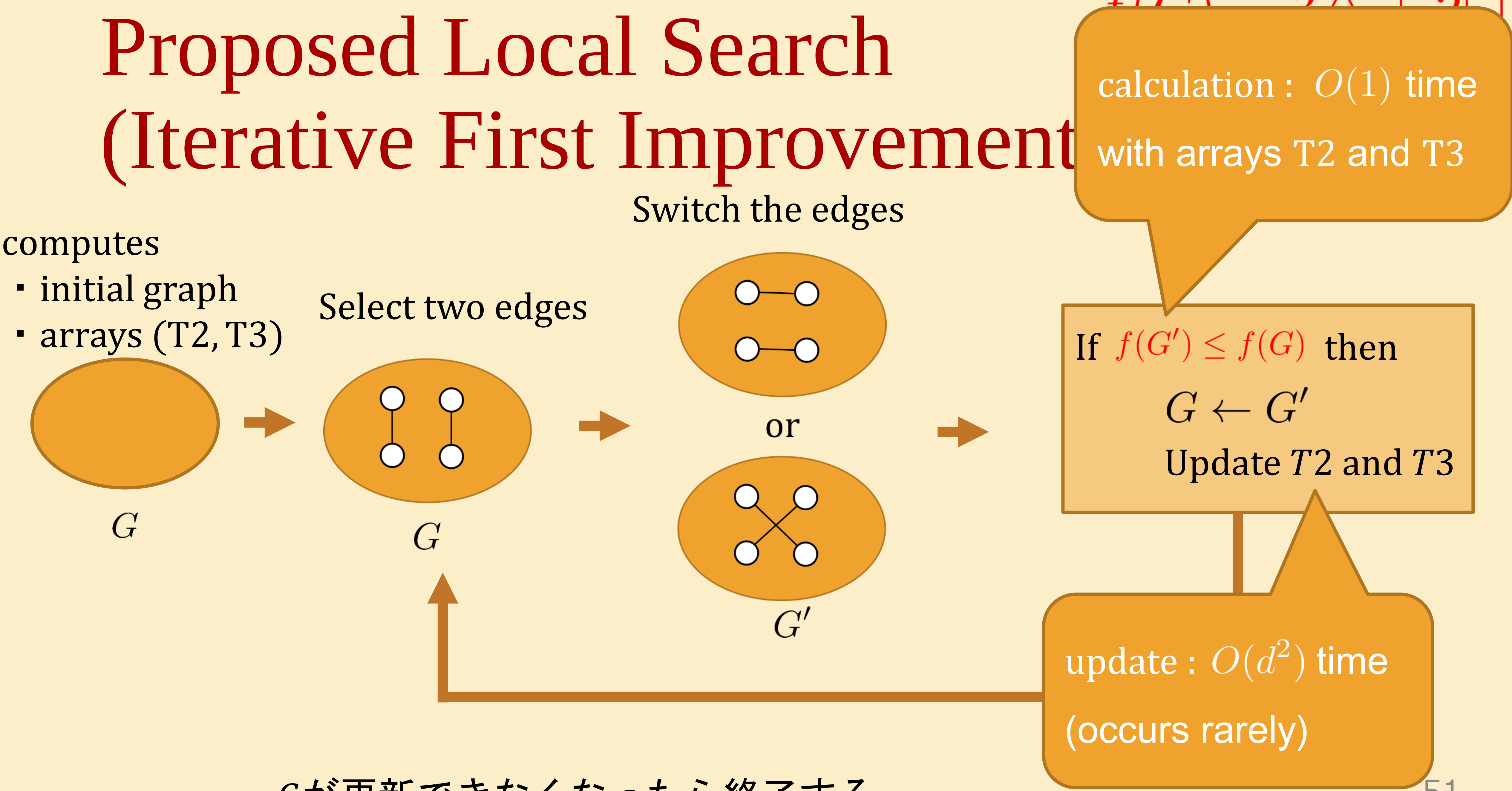
Proposed Local Search (Iterative First Improvement)

$f(G) = 3\triangle + 2\square$



・Gが更新できなくなったら終了する。

Proposed Local Search (Iterative First Improvement



・Gが更新できなくなったら終了する。

$f(G) = 3\triangle + 2\square$

Proposed Local Search (Iterative Fir… Improv…ent)

computes

・initial graph

then

G'

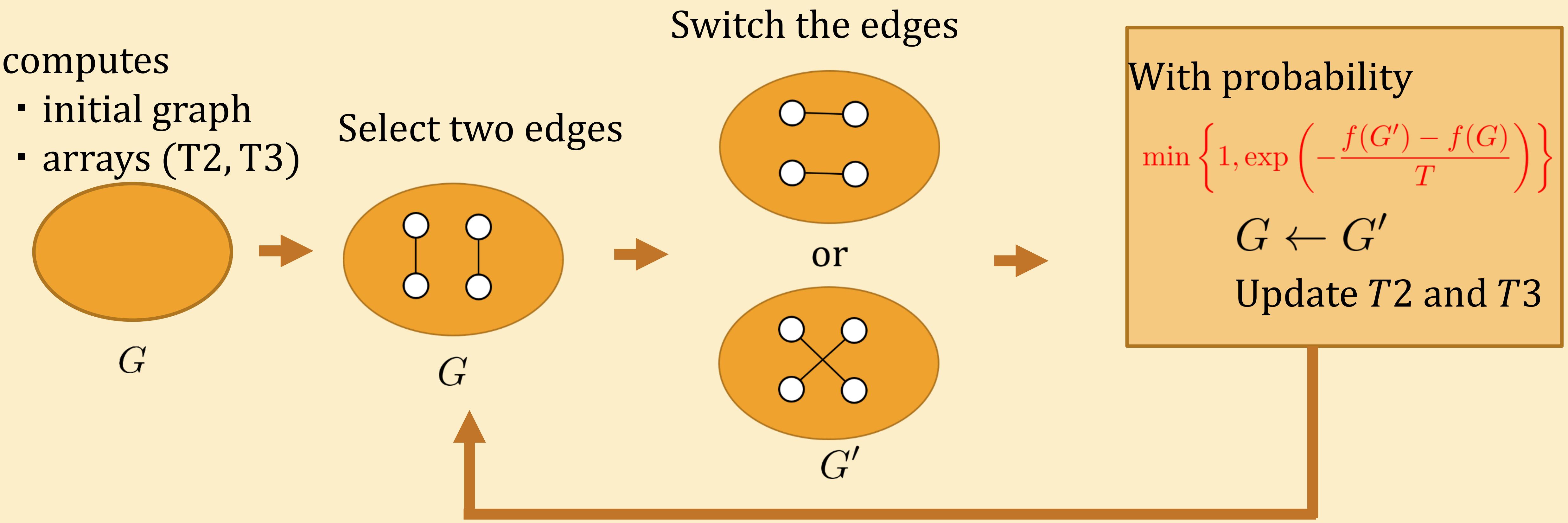
G

G

局所解を得た

・Gが更新できなくなったら終了する。

Simulated Annealing



・T は温度と呼ばれるパラメータ
・T が十分小さくなったら終了する。

Simulated Annealing

- G' is better than G
  → update <u>must</u> occur
- G' is worse than G
  → update <u>may</u> occur

computes
- initial graph
- arrays (T

Select two edges

Switch the edges

温度が高い → random search

温度が小さい → Iterative First Improvement に近い

T を徐々に小さくしていく。

G

G'

With probability

$$\min\left\{1, \exp\left(-\frac{f(G')-f(G)}{T}\right)\right\}$$

$$G \leftarrow G'$$

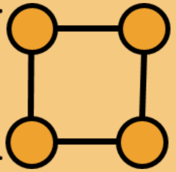
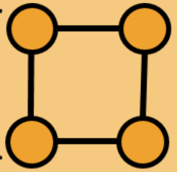
Update T2 and T3

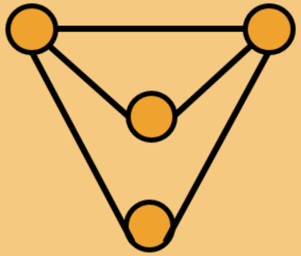
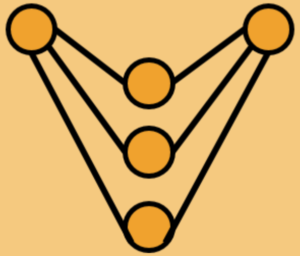
- T は温度と呼ばれるパラメータ
- T が十分小さくなったら終了する。

4. Numerical Experiments

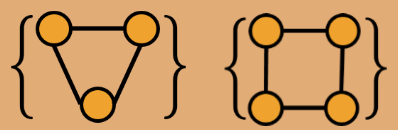
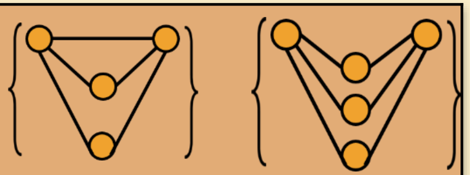
Numerical Experiments (1/2)

ASPLの近似の精度を調べる：

- $\binom{n}{2} \cdot ASPL(G) \geq 3 - \frac{nd}{2} - nd^2$　　the Moore Bound
- $\binom{n}{2} \cdot ASPL(G) \leq 3 - \frac{nd}{2} - nd^2 + 3\#$ {  } $+2\#$ {  }
- $\binom{n}{2} \cdot ASPL(G) \geq 3 - \frac{nd}{2} - nd^2 + 3\#$ {  } $+2\#$ {  }

  $-\#$ {  } $-\#$ {  }

Numerical Experiments (1/2)

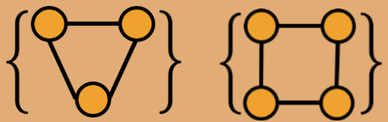
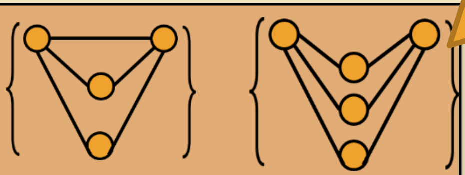
$(\boldsymbol{n}, \boldsymbol{d})$	**Moore**		
$(4096, 60)$	-0.1206	0.0355	-0.0074
$(4096, 64)$	-0.1544	0.0523	-0.0124
$(10000, 60)$	-0.0209	0.0024	-0.0002
$(10000, 64)$	-0.0270	0.0036	-0.0003

relative error:
(approx. val – ASPL) / ASPL

Numerical Experiments

Better than 

(計算が難しい)

$(\boldsymbol{n}, \boldsymbol{d})$	**Moore**		
$(4096, 60)$	-0.1206	0.0355	-0.0074
$(4096, 64)$	-0.1544	0.0523	-0.0124
$(10000, 60)$	-0.0209	0.0024	-0.0002
$(10000, 64)$	-0.0270	0.0036	-0.0003

グラフがスパースだとより精度が高い

Numerical Experiments (2/2)

- 実際にアルゴリズムを動かす。
- 直径3でスパースとなるような (n, d) を選んだ。
  - ➢ $(n, d) = (10000, 60), (10000, 64)$
- 初期グラフはランダム正則グラフとした。
- Iterative First Improvement (IFI) と Simulated Annealing (SA) を比較した。

Numerical Experiments (2/2)

初期温度 $T_0 = 11$
kステップ目の温度 $T_k = T_0 / \log k$

	Random	IFI		SA	
(n, d)	ASPL gap	ASPL gap	Time	ASPL gap	Time
$(10000, 60)$	21.3×10^{-3}	6.8×10^{-3}	40h 30m	6.2×10^{-3}	60 days
$(10000, 64)$	27.7×10^{-3}	10.8×10^{-3}	37h 00m	10.0×10^{-3}	60 days

ASPL gap = (solution ASPL – Moore Bound) / solution ASPL

Numerical Experiments (2)

The best graphs
in Graph Golf!!

	Random	IFI			
(n, d)	ASPL gap	ASPL gap	Time	ASPL gap	Time
$(10000, 60)$	21.3×10^{-3}	6.8×10^{-3}	40h 30m	6.2×10^{-3}	60 days
$(10000, 64)$	27.7×10^{-3}	10.8×10^{-3}	37h 00m	10.0×10^{-3}	60 days

ASPL gap = (solution ASPL – Moore Bound) / solution ASPL

5. Conclusion

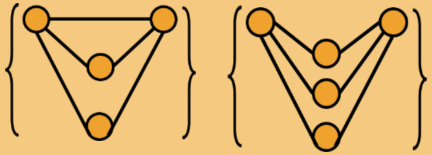
Conclusion

直径3のグラフに対して

・図形の個数によるASPLの特徴づけ

・三角形・四角形で評価する手法の提案

・実際に小ASPLなグラフを構成 (Graph Golf でベストなグラフ)

Future Work 1: 精度の高い近似による評価 : 

Future Work 2: 直径4以上のグラフ