BinN 理論談話会2022#7

Metropolis-Hastings sampling of paths

Gunnar Flötteröd , Michel Bierlaire
Transportation Research Part B 48 (2013) 53-66

2022/5/17 B4 白井帆香

Outline

1.導入

- 論文の目的：
  任意の分布に従って一般的なネットワークから経路を抽出する新手法の提案

- 特長：
  - 解析者がサンプリングプロトコルをコントロールできる
  - 実際のネットワークから経路を抽出できる
    - 論文では、17,000リンク・8,000ノードのネットワークで例証
  - 交通関連にとどまらず、所与の分布に従い経路を抽出するあらゆる場面に応用可能
    - 例：活動順序選択モデル、確率的マップマッチング、動的交通割り当て、経路誘導

1.導入　ー既往研究と問題意識

- 経路選択モデル
  - OD間の経路数が膨大になる可能性
  - 適切な経路を抽出することが必要
- 手法の整理　参考：BinN Studies シリーズ「ネットワーク行動学」（羽藤）

確定的に経路選択肢を列挙する	確率的に経路選択肢を列挙する	経路を明示しない方法
• ダイクストラ法 • K番目最短経路探索 • ラベリング • ペナルティ • Gateway Shortest Paths	• シミュレーション法 • ランダムウォーク • 提案手法	• Dialアルゴリズム • マルコフ連鎖による記述アプローチ

1.導入　ー既往研究と問題意識

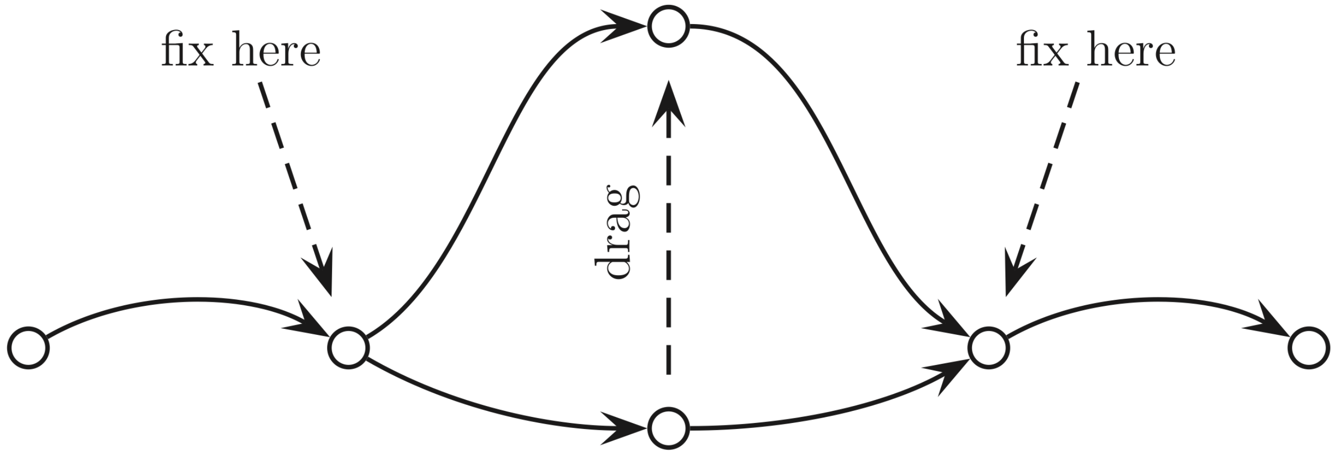
参考：BinN Studies シリーズ「ネットワーク行動学」（羽藤）

- 確率的に経路選択肢を列挙する既存手法
  - シミュレーション法
    - ネットワーク上の各リンクのコストが独立な確率分布に従う条件の下、最小コストとなる経路の抽出を繰り返す
      - 【Bekhor et al.】リンクコストが所要時間の正規分布に従うものとしてランダムサンプリング
    - <u>計算効率は良いが、サンプリング分布を制御できず、応用先が限られる</u>
  - 重み付きランダムウォーク【Frejinger et al.】
    - ノードごとの乱数発生により、確率的に次のリンクを選択して経路を生成
    - <u>サンプリング分布の定式化が必要</u>

→本論文では、任意の確率分布から選択確率を生成するアルゴリズムを提案

1.導入　ー提案手法の概要

参考：BinN Studies シリーズ「ネットワーク行動学」（羽藤）



<u>経路の一部置換による代替選択肢生成</u>

1. 最短経路上でノードを3点ランダムに選択
2. 2点(fix)を固定し、中間のノードを確率的に他のノードに置き換え新しい経路を生成
3. 生成された経路に対して選択確率が定義され、選択確率に応じた採択率によって採択or棄却が決定

1.導入　－MHアルゴリズム

参考：「基礎からのベイズ統計学」（豊田）

- 提案手法：**Metropolis-Hastings アルゴリズム**の一例
- Metropolis-Hastings（MH）アルゴリズムとは
  - ベイズ統計学の「マルコフ連鎖モンテカルロ法」の一つ
    - サンプリングしたい分布（＝目標分布）に対して、それを「定常分布」とするマルコフ連鎖を構成する方法

- マルコフ連鎖（MC)とは
  - マルコフ性：確率過程の未来の挙動が現在の状態のみに依存し、過去の状態と無関係であること
    マルコフ性を持つ確立過程をマルコフ過程という
    - 確率過程：確率論において、時間とともに変化する確率変数$\{Xt\}$のこと
  - マルコフ連鎖：マルコフ過程$\{Xt\}$のうち、取り得る状態が離散的なもの
    特に時間tが離散的なものを指すことが多い

1.導入　ーMHアルゴリズム

参考：「基礎からのベイズ統計学」（豊田）

- ベイズ統計学
  - ベイズの定理　$p(A_i|B_j) = \frac{p(B_j|A_i)p(A_i)}{p(B_j)}$
    - $p(A_i)$を事前確率、$p(A_i|B_j)$を事後確率という
  - 分布に関するベイズの定理
    - 条件付き確率の式を移行した$p(A_i, B_j) = p(B_j|A_i)p(A_i) = p(A_i|B_j)p(B_j)$について
      A_iを母数θ、B_jをxとして書き換えると
      $$f(\theta, x) = f(x|\theta)f(\theta) = f(\theta|x)f(x) \quad ①$$
      - ベイズ統計学では、$f(\theta)$を母数の分布として導入（母数θは確率変数）
    - ①を$f(x)$で割り、$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)}$
      - $f(\theta|x)$を事後確率分布（事後分布）、$f(\theta)$を事前確率分布（事前分布）と呼ぶ
      - $f(x|\theta)$は尤度

事前分布と尤度（データ）を結合し、確率を更新する

1.導入　ーMHアルゴリズム

参考：「基礎からのベイズ統計学」（豊田）

- マルコフ連鎖モンテカルロ法（MCMC法）
  - 目的：事後分布が明らかな状態で、事後分布に従う乱数を求める
    ＝事後分布が定常分布になるように、遷移核を設計する
    - 事後分布からの乱数の発生
      - 解析的な評価が難しい$f(\theta|x)$の積分を避け、母数を推論することができる
    - 遷移核
      - マルコフ連鎖を規定する条件付き確率（マルコフ連鎖では１期前のみが“条件”となる）
    - 定常分布
      - マルコフ連鎖の確率過程は定常分布に収束する（変化しなくなる）
      - 収束までの期間をバーンイン（焼き入れ）期間という
  - 詳細釣り合い条件
    ＝マルコフ連鎖が定常分布に収束するための十分条件
    - 分布が変化しないように制約をいれた全確率の式
    - 詳細釣り合い条件を確率密度関数で置き換えると

$f(\theta|\theta')f(\theta') = f(\theta'|\theta)f(\theta).$　目標分布： $f(\theta)$or$f(\theta')$　遷移核： $f(\theta|\theta')$or$f(\theta'|\theta)$

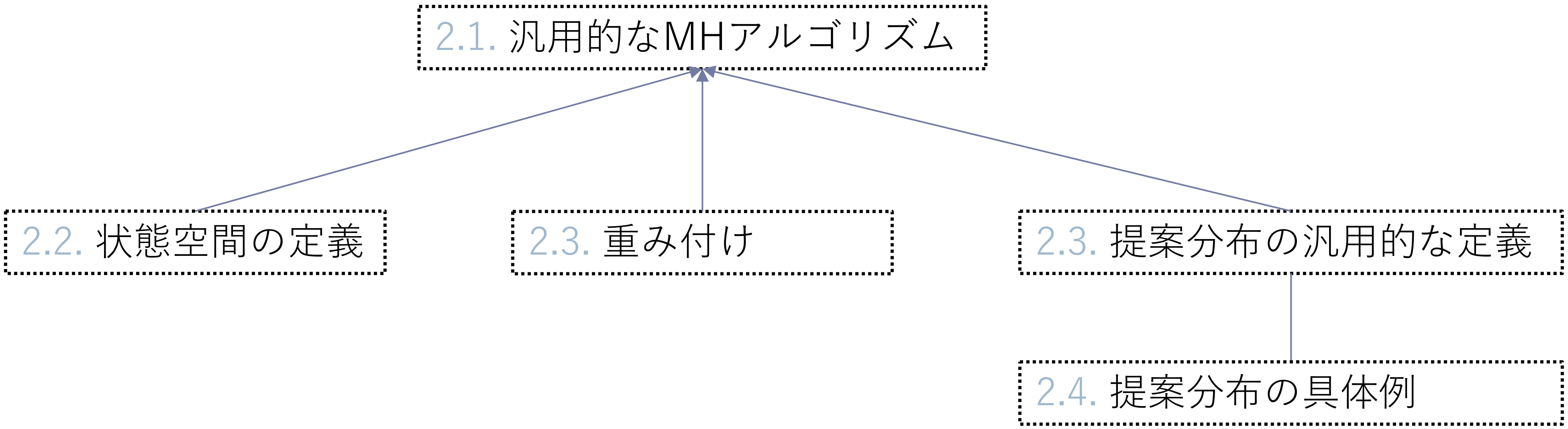
1.導入　ーMHアルゴリズム

参考：「基礎からのベイズ統計学」（豊田）

- 提案分布
  - 既知の事後分布（＝目標分布）$f(\theta)$を満たすような遷移核$f(|)$の探索は困難
  - $f(|)$の代わりに適当な遷移核$q(|)$を提案分布として利用
    - 「目標分布からのサンプルとしてこんなものは如何でしょう」と候補を提案してくれる条件付き確率

- 提案分布が詳細釣り合い条件を満たすよう確率補正をする方法＝MH法
  - 提案された候補点aを確率min(1,r)で受容($\theta^{(t+1)} = a$)し、さもなくばその場に留まる($\theta^{(t+1)} = \theta^{(t)}$)アルゴリズム
    - $r = \frac{q\left(\theta^{(t)}|a\right)f(a)}{q\left(a|\theta^{(t)}\right)f(\theta^{(t)})}$

2. フレームワーク


2.1. 汎用的なMHアルゴリズム
2.2. 状態空間の定義
2.3. 重み付け
2.3. 提案分布の汎用的な定義
2.4. 提案分布の具体例

2.1. 汎用的なMHアルゴリズム

MHアルゴリズムは、予め定義された定常分布からマルコフ連鎖を生成する。

定義する必要があるもの

S	マルコフ連鎖(MC)の有限状態空間	
$\{b(i)\}_{i \in S}$	正の重み	状態$i \in S$の定常確率に比例
$Q = (q(i,j))$	既約※提案分布 状態iから状態jへの遷移を提案する確率分布	

Qから得られるもの

$\alpha(i,j)$	採択率 状態iから状態jへの遷移提案を採択する確率	

※既約性
すべての状態jは、すべての状態iから1回以上の遷移によって到達され得ること

2.1. 汎用的なMHアルゴリズム

アルゴリズム1　Metropolitan-Hastings algorithm

1. 反復回数を示す添字kについて、$k=0$とする
2. 任意の初期状態i^kを選ぶ
3. 以下を繰り返す
    1. $\{q(i^k,j)\}_j$で遷移先の候補となる状態jを提案
    2. 採択率$\alpha(i^k,j) = \min(\frac{(b(j)q(j,i^k)}{b(i^k)q(i^k,j)},1)$を計算する
    3. $\alpha\left(i^k\right)$に応じて、$i^{k+1} = \mathrm{j}$、または$i^{k+1} = i^k$とする
    4. kを１つふやす

2.2. 状態空間の定義

前提

- 考慮するネットワーク上の全ての実行可能な経路集合を構成する必要がある
- 循環のない経路のみが対象であり、有限状態空間が保証されている

N	ノード集合
Γ	（循環のない）経路
$\lvert\Gamma\rvert$	経路Γに含まれるノードの数
$\Gamma(a)$	経路Γのa番目のノード
$\Gamma^{-1}(v)$	ノード$v \in \Gamma$の経路Γにおける位置、$1 \leq v \leq \lvert\Gamma\rvert$
$\Gamma(a, b)$	経路Γのノードab間のサブ経路、$1 \leq a \leq b \leq \lvert\Gamma\rvert$
$\Gamma_1 + \Gamma_2$	経路Γ_1と経路Γ_2の連結
$\delta(\Gamma)$	経路Γ_1の長さ（コスト）
$SP(v, w)$	ノードvからノードwの最短経路
$SP_M(v, w)$	$\mathrm{M} \subseteq \mathrm{N}$の点のみ辿れるとした場合の、ノード$v$からノード$w$の最短経路
$\delta_{SP}(v, w)$	ノードvからノードwの最短経路コスト
$\delta_{SP,M}(v, w)$	$\mathrm{M} \subseteq \mathrm{N}$の点のみ辿れるとした場合の、ノード$v$からノード$w$の最短経路コスト

解析者が設計できる
典型例）経路上のノード数
　　　　長さや移動時間などのリンクコストの合計

最短経路は効率的で計算可能な提案分布を定義するためだけに使われる。（目標分布は最短経路によって定義されなくてもよい。）

2.2. 状態空間の定義

要件１
すべてのリンクの上流ノードと下流ノードの間の最短経路(SP)は、
そのリンクを通る。

すべての可能な経路がSPの連結によって構築できる

定義１※
MHアルゴリズムの状態空間Sは、Γが出発地と目的地をつなぐ循環のない経路であるとき、タプル(Γ, a, b, c)で構成され、a, b, cは$1 \leq a < b < c \leq |\Gamma|$を満たす整数である。

- 経路Γに対し、$\binom{|\Gamma|}{3} = |\Gamma|(|\Gamma| - 1)(|\Gamma| - 2)/6$個の$a, b, c$の配置が存在
  - 状態空間のサイズと経路数の比が、平均経路長の3乗のオーダーとなる

※定義１からは３ノードより少ない経路が除外されているが、出発地と目的地を結ぶ１リンクの経路にしか影響しない。この経路もダミーノードを加えることで、状態空間に加えることができる。

2.3. 重み付け

- 経路上のノード数が選択確率に与える影響を排する
  - 経路Γを含む状態iの重みは以下のように定義できる

$$\mathrm{b}(i) = \frac{d(\Gamma)}{|\Gamma|(|\Gamma|-1)(|\Gamma|-2)/6} \qquad (1)$$

  - $d(\Gamma)$：経路Γの重みづけ　定義は応用例に依存する

- 経路選択など多くの状況では、
  重みは経路コストが指数関数的に減少する関数

$$\mathrm{b}(i) = \frac{exp[-\mu\delta|\Gamma|]}{|\Gamma|(|\Gamma|-1)(|\Gamma|-2)/6} \qquad (2)$$

  - μは非負のスケールパラメーター $\begin{cases} \mu = 0 & \text{異なるコストの経路について無差別} \\ \mu \to \infty & \text{SPしか考慮に入れない} \end{cases}$
  - 本論文の実験では(2)を採用

2.4.　提案分布の汎用的な定義

- 計算効率のため、提案分布はSP計算に大きく依存→要件２

要件２

複数のSPで繋がったすべてのノードペア(v, w)について、

関数$SP(v, w)$は再現性よくこの集合から特定の1つの要素を返す

- 要件２の目的
  - 単一の選択されたSPに制限することで、アルゴリズムの複雑さを軽減
- 要件２を満たすため
  - 経路の一意な順位づけが必要
  - vとwを同じ最小コストで結ぶ経路が複数ある場合、任意の基準に従って、順位づけと選択がなされる　基準の例）辞書
- 既約性
  - 既約性は保証（証明は略）→要件２はアルゴリズムが到達できる経路集合を制限しない

2.4. 提案分布の汎用的な定義

- 提案分布Qは、2つの操作によって定義する

SPLICE

2.4.1.

SHUFFLE

2.4.2.

結合

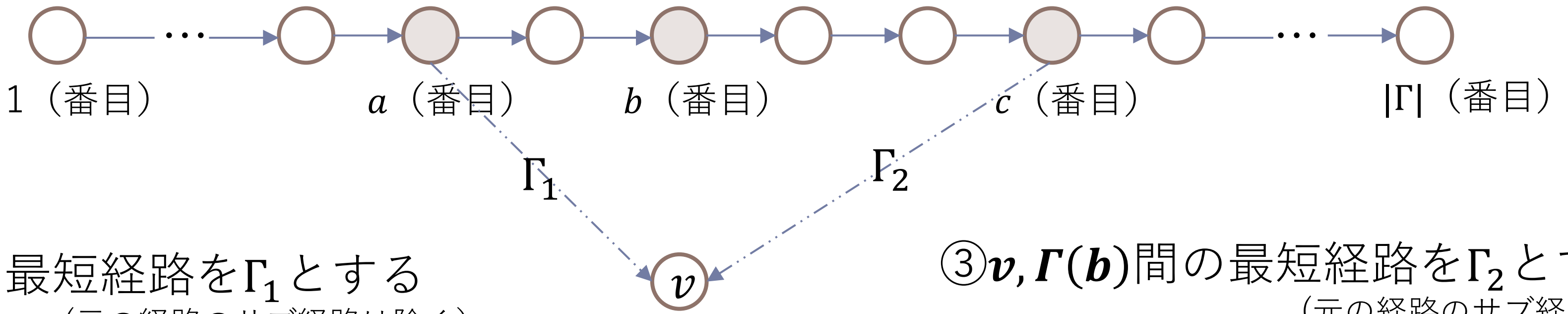
2.4.3.

2.4.　提案分布の汎用的な定義

- 提案分布Qは、2つの操作によって定義する

SPLICE

P_{splice}で操作が発生



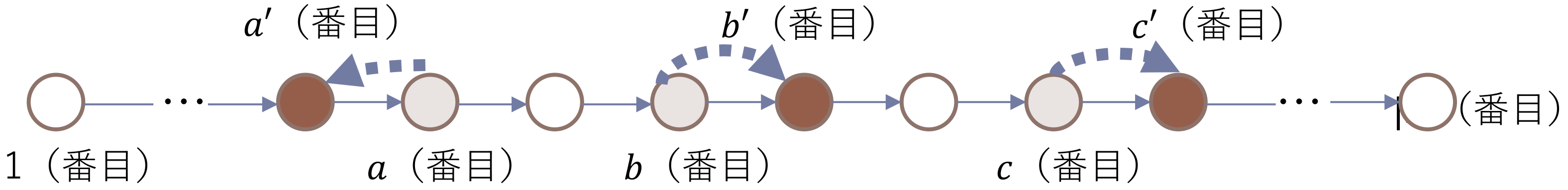
② $\boldsymbol{\Gamma(a), v}$間の最短経路を$\Gamma_1$とする
(元の経路のサブ経路は除く)

①$P_{insert}(v)$でvを挿入

③$\boldsymbol{v, \Gamma(b)}$間の最短経路を$\Gamma_2$とする
(元の経路のサブ経路は除く)

④元の経路のab間を$\Gamma_1 + \Gamma_2$で置換し、新経路を生成

SHUFFLE



$P_{shuffle}$で組み合わせ(a, b, c)の値を変更

2.4.1.　提案分布の汎用的な定義 ーSPLICE

- SPLICE の操作方法
  - 現在の状態$i = (\Gamma, a, b, c)$が与えられれば、候補となる状態$I' = (\Gamma', a'. b', c')$は以下の操作で生成される

1. すべての$v \in N$について正の値をもつ所与の選択分布$P_{insert}(v)$によって、挿入ノードvを選択。
2. $\boldsymbol{N_1(i) = N \setminus \Gamma(1, a-1) \setminus \Gamma(c, |\Gamma|)}$**のとき、**$\boldsymbol{\Gamma_1 = SP_{N_1(i)} = (\Gamma(a), v)}$とする※1　このような経路が存在しない場合、$i' = i$としてもとの状態を維持し、SPLICEを終了する。
3. $\boldsymbol{N_2(i) = N \setminus \Gamma(1, a) \setminus \Gamma(c+1, |\Gamma|)}$**のとき、**$\boldsymbol{\Gamma_2 = SP_{N_2(i)} = (v, \Gamma(c))}$とする※2　このような経路が存在しない場合、$i' = i$としてもとの状態を維持し、SPLICEを終了する。

※1 Γの全てのノードからa未満のものとc以上のものを除外した時の、 $\Gamma(a)$ からvへの最短経路

※2 Γの全てのノードからa以下のものとcより大きいものを除外した時の、vから$\Gamma(c)$への最短経路

2.4.1.　提案分布の汎用的な定義 ーSPLICE

- SPLICEの操作方法（続き）

4. $\tilde{\Gamma} = \Gamma(1, a) + \Gamma_1 + \Gamma_2 + \Gamma(c, |\Gamma|)$とする※1

   $\tilde{\Gamma}$に循環がある場合は、$i' = i$としてもとの状態を維持し、*SPLICE* 操作を終了する

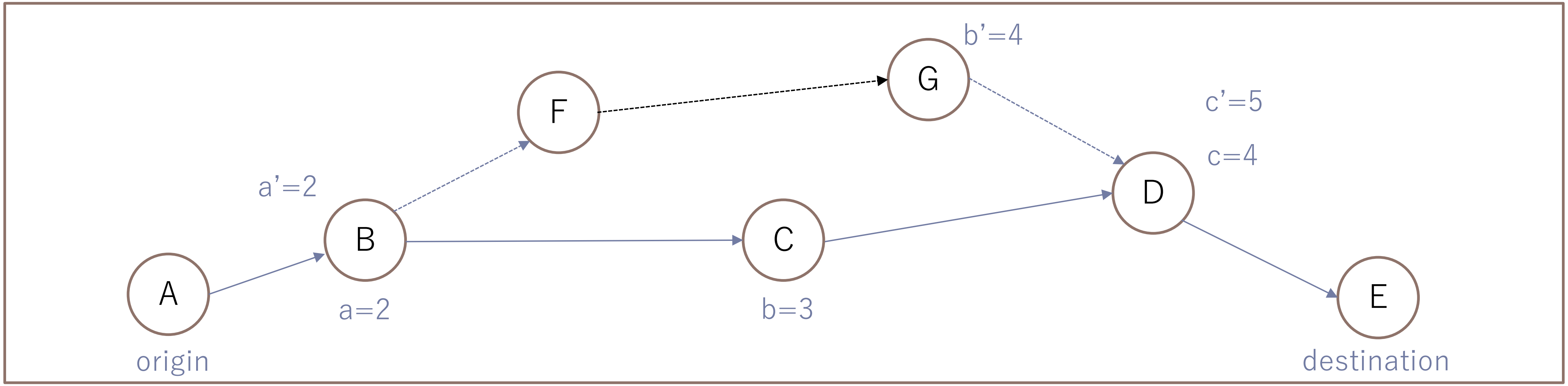
5. $\Gamma' = \tilde{\Gamma}, a' = a, b' = \tilde{\Gamma}^{-1(v)}, c' = b' + |\Gamma_2|$-1とする※2

※1 $\Gamma(a)$から$\Gamma(c)$への迂回路がvを通過する経路
※2 つまり、新しい経路を引き継ぎ、a’とc’は以前と同じノードを指し、b’は挿入ノードvを指すように指標を調整する。

SPLICE =経路Γをa, cに固定したまま、迂回路または近道を導入する操作

2.4.1. 提案分布の汎用的な定義 ーSPLICE

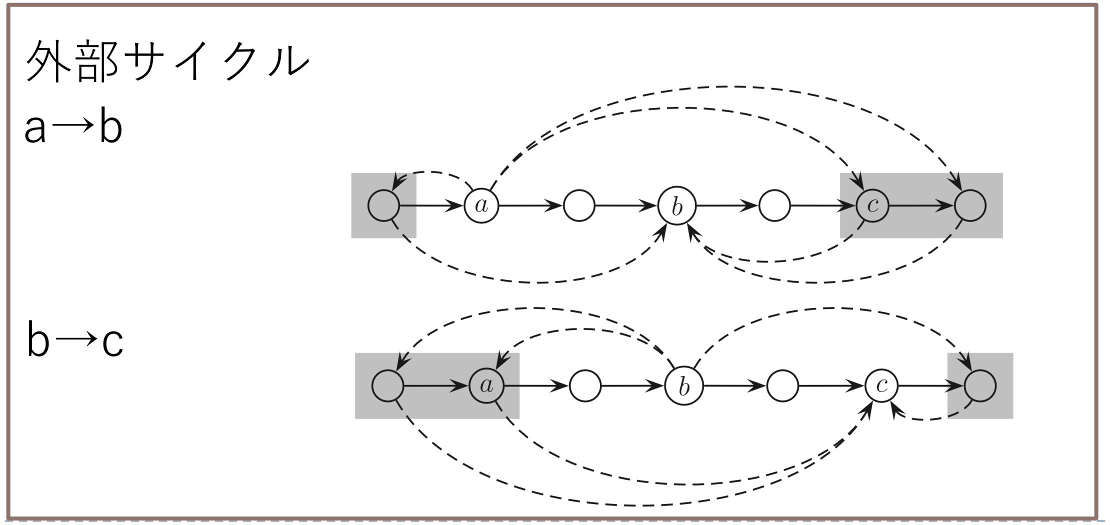
- SPLICEの具体例
  - 初期状態iを構成する経路Γについて$\Gamma = \{A, B, C, D, E\}. (a.b.c) = (2,3,4)$
  - 挿入ノードGは、$\Gamma(a) = B$に$\{B, F, G\}$を通して接続、$\Gamma(c) = D$に$\{G, D\}$を通じて接続される。
  - Γのうち$\{B, C, D\}$の部分を迂回路で置き換えると、新しい経路$\Gamma' = \{A, B, F, G, D, E\}$を得られる。
  - 新しい位置指標は、以下のように設定する
    1. 新しい指標$a' = 2$と$c' = 5$が元の状態との境界ノードである$B = \Gamma'(a')$と$D = \Gamma'(c')$を指す
    2. $b' = 4$が新しい挿入ノード$\Gamma'(b') = G$を指す

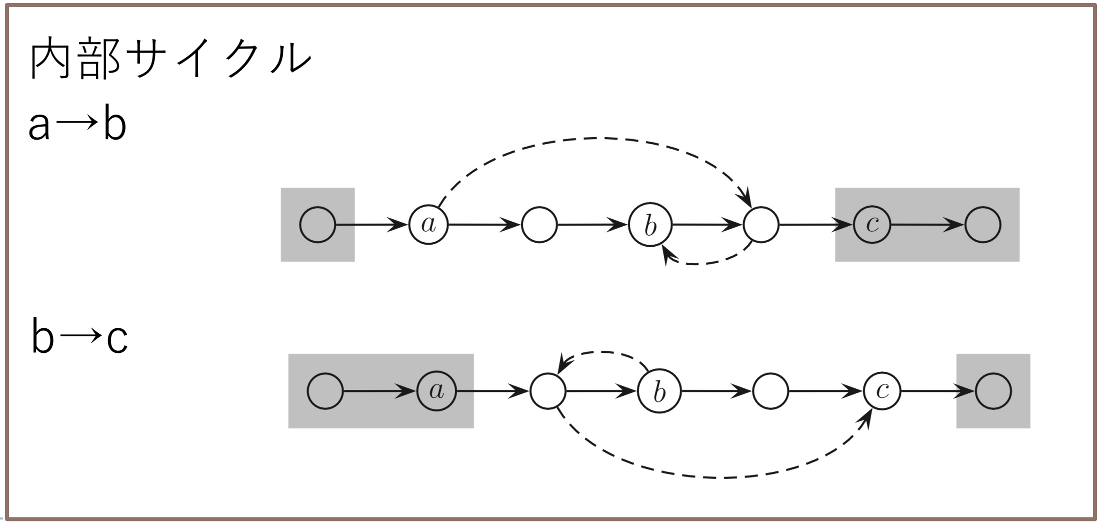


2.4.1. 提案分布の汎用的な定義 ーSPLICE

- SPLICEは２種類の方法で循環を禁じている
  1. 外部サイクルの禁止
     - 「SPLICEの操作」2.と3.で、最短経路の計算から、それぞれノードセット$N \setminus N_1$と$N \setminus N_2$が除外されている

     → もとの経路において「aより上流」または「bより上流」で交差するSPLICE部分は計算できない
  2. 内部サイクルの禁止
     - 「SPLICEの操作」4.で、循環がある場合はもとの状態が維持される。





2.4.2.　提案分布の汎用的な定義 －SHUFFLE

- SHUFFLE の操作方法
  - 現在の状態$i = (\Gamma, a, b, c)$が与えられれば、候補となる状態$I' = (\Gamma', a'. b', c')$は以下の操作で生成される

1. ある分布$P_{shuffle}(a', b', c' | i)$から、3つの整数の組$(a', b', c')$を引用（$1 \leq a' \leq b' < c' \leq |\Gamma|$）
   - $P_{shuffle}$は、任意の配列$1 \leq a'' \leq b'' < c'' \leq |\Gamma|$に有限回のSHUFFLE操作で到達できるものでなければならない
2. $\Gamma' = \Gamma$とする。

SHUFFLE =SPLICE操作での固定点をランダムに変更する操作
3つの整数をシャフルするだけで、経路には影響しない

2.4.3.　提案分布の汎用的な定義 ー操作の結合

- “*spliceable*”の定義
  - アルゴリズム1より、$q(i,j) > 0 \Leftrightarrow q(j,i) > 0$が必要　“遷移は可逆”

定義2

経路を構成する$\Gamma(a)$から$\Gamma(b)$の部分および$\Gamma(b)$から$\Gamma(c)$の部分がそれぞれ最短経路であるとき、状態$i = (\Gamma, a, b, c)$は”*spliceable*”である。

- つまり、$\Gamma(a,b) = SP(\Gamma(a), \Gamma(b))\ and\ \Gamma(b,c) = SP(\Gamma(b), \Gamma(c))$ であるということ
- SPLICE操作から生まれるすべての状態は”*spliceable*”である

- 定義の目的
  - “*spliceable*”でないものにSPLICE操作をすると、高確率で経路のノードが変更されるが、もとの状態に回復できない→遷移が可逆でない→アルゴリズム１で拒否される
  - “*spliceable*”でないものは（どうせ拒否されるので）計算したくない

2.4.3.　提案分布の汎用的な定義 ー操作の結合

- SPLICEとSHUFFLEの結合操作
  - 実行可能な状態iから実行可能な状態i'への遷移は以下のように選択される

1. Γが”$spliceable$”であるとき、確率P_{splice}※で
   a. $i' = SPLICE(i)$

Γが”$spliceable$”でない、または確率P_{splice}※で操作が起こらなかったとき
   a. $i' = SHUFFLE(i)$

- ※ 3.2.の感度分析から、2.5.で定義した具体的なSPLICEおよびSHUFFLEに対して、P_{splice}の値が0.5から0.75の間で良好かつ堅牢な演算性能が得られることが示された。

2.5. 提案分布の具体例

- SPLICEでノード$v \in N$を挿入する選択確率$P_{insert(v)}$

$$P_{insert(v)} = \frac{\exp[-\tilde{\mu}(\delta_{Sp}(origin, v) + \delta_{Sp}(v, destination))]}{\sum_{w \in N} \exp[-\tilde{\mu}(\delta_{Sp}(origin, w) + \delta_{Sp}(w, destination))]} \quad (3)$$

  - $\tilde{\mu}$は非負のスケールパラメータ、挿入ノードをOD間の最短経路からどれだけ遠くにとるかを制御する
  - すべてのノードは正の確率で選択され得る
  - 分母の計算量は多いが、この分布は状態iに依存しないため、MHアルゴリズムの前に一度計算するだけで良い（各ノードは経路サンプリング前に選択確率を割り当てられ、この確率はサンプリングプロセス中ずっと固定される）

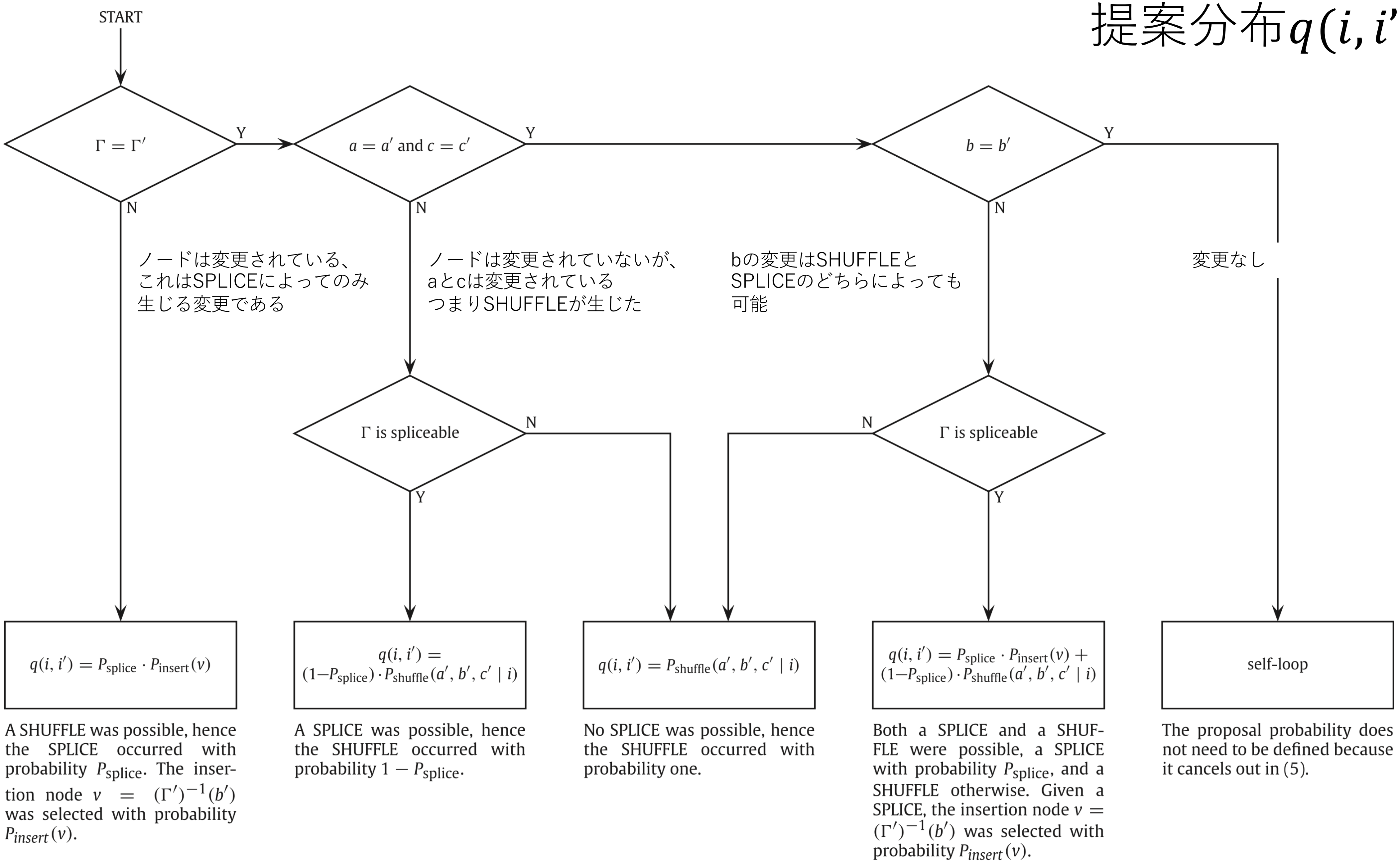
2.5.　提案分布の具体例

- SHUFFLEの選択確率$P_{shuffle}(a', b', c'|i)$

$$P_{\text{shuffle}}(a', b', c'|i) = \begin{cases} \left(\frac{|\Gamma|(|\Gamma|-1)(|\Gamma|-2)}{6}\right)^{-1} & \text{if } 1 \leqslant a' < b' < c' \leqslant |\Gamma| \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

  - すべての指標$(a', b', c'|i)$について同一
  - すべての指標$(a', b', c'|i)$の組み合わせは、有限回の遷移で到達できる

提案分布$q(i, i')$の定義


START
Γ = Γ′
Y
N
a = a′ and c = c′
Y
N
b = b′
Y
N
ノードは変更されている、
これはSPLICEによってのみ
生じる変更である
ノードは変更されていないが、
aとcは変更されている
つまりSHUFFLEが生じた
bの変更はSHUFFLEと
SPLICEのどちらによっても
可能
変更なし
Γ is spliceable
N
Y
Γ is spliceable
N
Y
q(i, i′) = P_splice · P_insert(v)
q(i, i′) = (1−P_splice) · P_shuffle(a′, b′, c′ | i)
q(i, i′) = P_shuffle(a′, b′, c′ | i)
q(i, i′) = P_splice · P_insert(v) + (1−P_splice) · P_shuffle(a′, b′, c′ | i)
self-loop
A SHUFFLE was possible, hence the SPLICE occurred with probability P_splice. The insertion node v = (Γ′)^−1(b′) was selected with probability P_insert(v).
A SPLICE was possible, hence the SHUFFLE occurred with probability 1 − P_splice.
No SPLICE was possible, hence the SHUFFLE occurred with probability one.
Both a SPLICE and a SHUFFLE were possible, a SPLICE with probability P_splice, and a SHUFFLE otherwise. Given a SPLICE, the insertion node v = (Γ′)^−1(b′) was selected with probability P_insert(v).
The proposal probability does not need to be defined because it cancels out in (5).

2.5.　提案分布の具体例

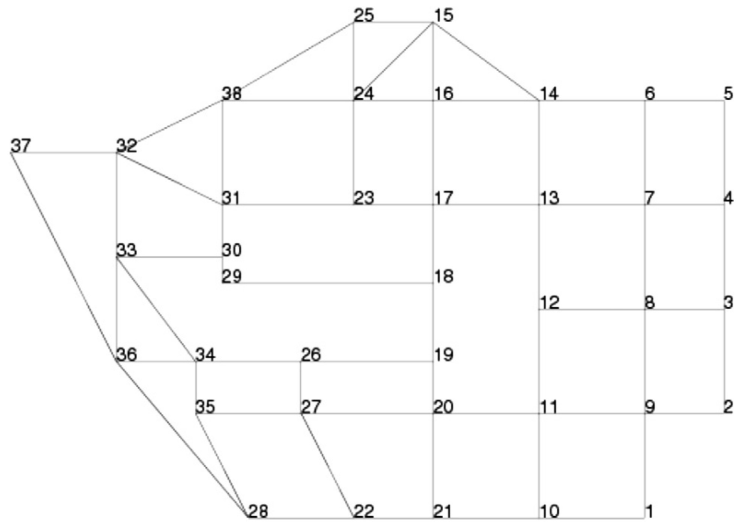
アルゴリズム２　提案手法の要約

1. 反復回数を示す添字kについて、$k = 0$とする
2. 任意の初期状態$i^k = (\Gamma, a, b, c)^k$を選ぶ. 例えば
   a. Γ^{k}を最短経路とする
   b. $(a, b, c)^k$を$1 \leq a^k < b^k < c^k \leq |\Gamma^k|$をなるよう一様におく
3. 以下を繰り返す
   a. Γ^kが”$spliceable$”であるとき、確率P_{splice}で
      i. 具体的な挿入ノード確率を用いて、SPLICEで候補となる状態jを選択する

   Γ^kが”$spliceable$”でないとき、または確率P_{splice}で何も起きないとき
      i. 具体的なインデックス選択分布を用いて、SHUFFLEで候補となる状態jを選択する
   b. 提案分布$q(i^k, j)$と$q(j, i^k)$を計算する
   c. 採択率$\alpha(i, k\ ) = \min(\frac{(b(j)q(j, i^k)}{(b(i^k)q(i^k, j,)})$を計算する
   d. $\alpha\left(i^k, j\right)$に応じて、$i^{k+1} = \mathrm{j}$、または$i^{k+1} = i^k$とする
   e. kを１つふやす

3.1. 数値実験　ー単純な例

- 実験条件・内容
  - 循環のないネットワーク（右図）
    - 64のリンク
    - 38のノード
    - 出発地（ノード番号1）と目的地（ノード番号38）
    - ODを結ぶ170の経路を有する
    - 最大12ステップで目的地に到達
  - 3回の実験でMHアルゴリズムを10^8回繰り返す
    - $P_{slice} = 0.5$
    - $\mu = \tilde{\mu}$はそれぞれ0,2,4
    - $\mathrm{b}(i) = \frac{exp[-\mu\delta|\Gamma|]}{|\Gamma|(|\Gamma|-1)(|\Gamma|-2)/6}$ (2)



3.1. 数値実験　ー単純な例

- 類似性測定
  - 2つの状態空間の類似性は、間に行われるアルゴリズムの反復回数が増えるにつれ減少

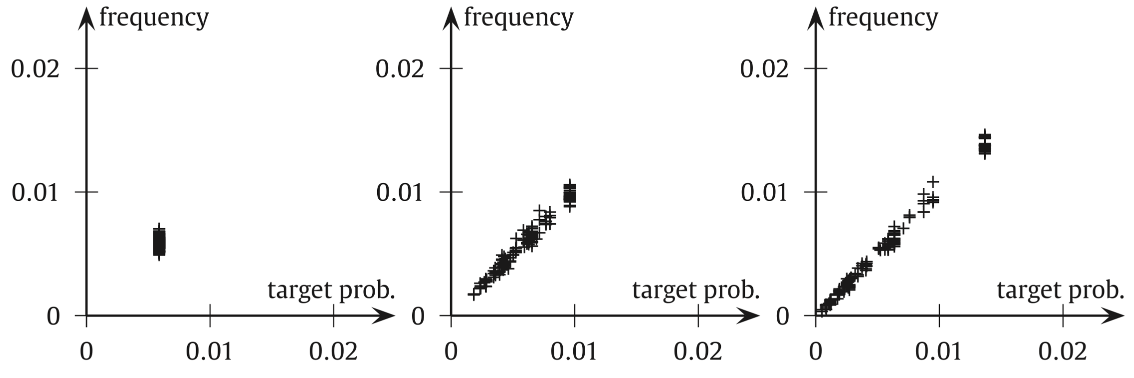
Γ^kとΓ^{k+d}の経路の平均的な類似度$\phi(d)$

$$\phi(d) = \frac{1}{K}\sum_{k+1}^{K} \frac{|\Gamma^k \cap \Gamma^{k+d}|}{\frac{1}{2}(|\Gamma^k| + |\Gamma^{k+d}|)} \tag{8}$$

  - $|\Gamma^k \cap \Gamma^{k+d}|$は k 回目と $k+d$ 回目で生成されたパスにおける同一ノードの数
- 3回の実験では$\phi(d) = 0.5$付近で安定する
  - ネットワークがかなり制約されており、完全にランダムなパスでも重なりやすいため
- 経路の独立性が担保される距離値d
  - 線形回帰モデルより、 $\mu = 0$ で2900、$\mu = 2$で2400、 $\mu = 4$で2100
  - 経路の初期値はSPなので、その後それぞれ2900,2400,2100番目ごとの経路が独立なものとして抽出される

3.1. 数値実験　ー単純な例

- 実験結果
  - MCが望ましい目標分布を生成する能力
    - x座標：各経路の目標確率　y座標：サンプル中の経路の頻度



(a) $\mu = 0$　(b) $\mu = 2$　(c) $\mu = 4$

- 点はy=xを中心に対称的に分布
  →目標分布を偏りなく再現
- $\mu = 0$ではすべての経路が同じ目標確率

    - χ^2検定による目標分布とサンプルの一貫性の検証
      - 3 回の実験すべてについて、有意差なし（有意水準10%）

提案手法は、目標分布に合致した経路サンプルを生成できる

3.2. 数値実験　ー複雑な例

- 実験の目的
  - μによって定義された異なる提案分布が与えられたとき、アルゴリズムの計算性能に対する$\tilde{\mu}$とP_{splice}の効果を確認
- 実験条件
  - 実在するイスラエル・テルアビブのネットワーク
    - 17,118のリンク　7,879のノード
  - 以下のパラメーターのすべての組み合わせ(27組)に対し、10^7回の反復
    - $\mu = 0.01/0.02/0.04$　$\tilde{\mu} = 0.5\mu/1.0\mu/2.0\mu$　$P_{splice} = 0.25/0.5/0.75$
- 実験結果
  - 各経路を抽出するのにかかる時間はμが小さくなるにつれて大幅に増加

計算速度は良好であり、十分に実用可能

4. まとめ

- 任意の分布に従う効率的な経路抽出手法の提案
  - MHアルゴリズムに合致する「状態空間」と「提案分布」を定義
- 数値実験で提案手法の実用性を確認
  - 小規模ネットワークで生成される経路の頻度は、目標分布に十分に適合
  - 大規模ネットワークでも、良好な計算速度を確認
- 経路選択モデルにとどまらない応用が可能
  - 例）動順序選択モデル、確率的マップマッチング、動的交通割り当て、経路誘導のための選択肢のサンプリング

参考文献

- Gunnar Flötteröd , Michel Bierlaire.(2013).Metropolis-Hastings sampling of paths. *Transportation Research*, Part B,48,53-66
- BinN Studies シリーズ「ネットワーク行動学」羽藤英二・伊藤創太・伊藤篤志　篇（http://bin.t.u-tokyo.ac.jp/kaken/data/full-20140926.pdf）
- 豊田秀樹.「基礎からのベイズ統計学ーハミルトニアンモンテカルロ法による実践的入門ー」(2015).朝倉書店