

凸凹と中立性を含む適応度景観の中立ネットワークに基づく構造解析

片田 喜章[†],

摂南大学工学部[†],

1 はじめに

中立性(neutrality)は遺伝子型-表現型写像もしくは表現型-適応度写像に多くの冗長性を含むことで現れる。実問題においても,進化ロボティクスにおけるニューラルコントローラ¹⁾や進化型ハードウェア²⁾などで中立性が確認されている。これらの問題では一般に,上述の写像を一对一に設計することが困難で,冗長になるという共通の特徴がある。

中立性を含む適応度景観の概念は,Harvey³⁾によって中立ネットワーク(neutral networks)としてEC研究分野に初めて導入された。その定義を原文のまま以下に引用する。“A neutral network of a fitness landscape is defined as a set of connected points of equivalent fitness, each representing a separate genotype: here connected means that there exists a path of single (neutral) mutations which can traverse the network between any two points on it without affecting fitness.” また,この考え方をもとに,冗長性を排しコンパクトに設計することが良いとされてきた遺伝子型-表現型写像,つまり,遺伝子コーディングに意図的に冗長性を導入し,中立ネットワークを構成し遺伝的アルゴリズムの解探索能力を高めようという研究も多くなされている⁴⁾。

中立ネットワークを含む問題における進化ダイナミクスは平衡期と過渡期に分類することができる⁵⁾。平衡期では,遺伝子型空間において個体集団は,“準種”とよばれるクラスタを形成し,中立ネットワークに沿って,より高い適応度を示す中立ネットワークへと導くポータルと呼ばれる地点(遺伝子型)を見つけるまで移動し続ける。その後,個体集団はポータルから適応度の高い中立ネットワークへ移るべく,過渡期が始

まる。この過渡期は平衡期に比べてその長さが極端に短いものが典型である。

中立ネットワークにある個体集団の分布は,突然変異率に関連して明確な相転移が見られる。低い突然変異率では,個体集団はクラスタを形成しネットワークに沿って移動する。しかし,突然変異率が高くなるにつれ,いくつかの個体は現在の中立ネットワークから適応度の低いものへと落ち始める。そして突然変異率があるしきい値を越えると,選択の効果では個体を現在の中立ネットワークに保持できなくなり,集団は適応度の低い方へ完全に落ちてしまう。このしきい値(q_{err})は中立ネットワークごとに異なることが報告されている^{5, 6, 7)}。

突然変異率が一定の場合,すべての中立ネットワークに対し q_{err} を越えないように突然変異率をあらかじめかなり低く設定する必要がある。しかしながら,適応度景観に凸凹(ruggedness)を同時に含む問題においては,従来議論されてきたように低い突然変異率では個体集団は局所解に陥ってしまうと予想される。このような凸凹と中立性を含む適応度景観の形状および問題の構造はどのようになっているのであろうか。

この研究領域でしばしば使用される中立性と凸凹の度合いを調節可能なテスト問題である拡張型NKランドスケープに関する研究^{6, 7)}において,中立性を増加させると局所解の数は減少するものの,凸凹の度合いは変化しないということが報告されている。これらの報告は統計的手法に基づいて行われており,中立性と凸凹を同時に含む適応度景観が実際にいかなる形状をしているのかは明らかにされていない。ここで,中立ネットワークが形成されていても凸凹の影響が現れる場合,ネットワークの連なりが凸凹を示すか,あるいは,ある一連のネットワーク群が他のネットワークに対して孤立しているような状況になるのではないかと予想される。

そこで,本研究では中立性と凸凹を同時に含む適応度景観の形状を明らかにすることを目的と

Analysis on Topologies of Fitness Landscapes with Both Ruggedness and Neutrality Based on Neutral Networks

[†] Yoshiaki Katada(katada@ele.setsunan.ac.jp)

Department of Electrical and Electronic Engineering, Setsunan University (†)

する．まず，Harveyが定義した中立ネットワークの概念を集合として定式化する．これに基づき凸凹と中立性を同時に含む問題の解空間をクラスタリングし，その問題の適応度景観の形状，つまり，中立ネットワークの構造および中立ネットワーク間の構造(つながり)を調べる．以下，第2節では，中立ネットワークを集合として定式化する．第3節では，第2節で定式化した中立ネットワークに基づいた，各ネットワークおよびネットワーク間の構造の解析法を示す．第4節では，非常に簡単なテスト問題を例に構造解析法について具体的に示し，第5節で凸凹と中立性を同時に含むテスト問題において構造解析を行う．

2 中立ネットワーク

本節では，文献⁴⁾の記法に従って，中立ネットワークを集合として定式化する．本研究で扱う遺伝子型は，固定長(長さ N)でバイナリコーディングされていると仮定する． Φ_g は遺伝子型の集合，つまり遺伝子型空間を表す．また， Φ_p は表現型空間を表す．第1節で述べたHarveyの定義より，遺伝子型-表現型写像が冗長であることから形成される中立ネットワークは，

$$\Phi_g^* = \{x^g \in \Phi_g | \exists y^g \in \Phi_g \text{ s.t. } f_g(y^g) = f_g(x^g), \\ 0 < H(x^g, y^g) \leq d\}, \quad (1)$$

と表される．ただし， f_g は Φ_g から Φ_p への写像を表し，全射であって単射ではない．また， $H(x^g, y^g)$ は2つの遺伝子 x^g, y^g の間のハミング距離であり， d は十分に小さいものとする．

Rothlaufら⁴⁾は冗長な遺伝子型-表現型写像を採用する進化型計算が有効である必要条件として，同義的冗長表現(synonymously redundant representation)を提案した．この条件は，とり得る全ての表現型に対して，同じ表現型に写像される遺伝子型すべての遺伝的な距離の総和が，十分に小さいというものである．これを満たす遺伝子型の集合は式(1)にほぼ一致すると考えられる．

これまでの議論を表現型-適応度間の冗長写像に拡張する． Φ_f は適応度の集合とする．説明の簡略化のため， f_g は全単射(つまり，冗長性を含まない)であると仮定すると，Harveyの定義より，表現型-適応度写像が冗長であることから形成される中立ネットワークは，

$$\Phi_g^* = \{x^g \in \Phi_g | \exists y^g \in \Phi_g \text{ s.t. } (f_p \circ f_g)(y^g) = \\ (f_p \circ f_g)(x^g), 0 < H(x^g, y^g) \leq d\}, \quad (2)$$

と表される．ただし， f_p は Φ_p から Φ_f への写像を表し，全射であって単射ではない．

一般に，写像 f_g は進化型計算の設計者によって問題ごとに決定される．一方， f_p は設計者が設定する適応度関数を介して決定されるが，問題によって決まると考えられる．そこで，本稿では表現型-適応度写像が冗長であることから形成される中立ネットワークを扱う．

Harveyの定義の“single (neutral) mutations”という表現は多義的である．これは式(1)(2)中の $0 < H(x^g, y^g) \leq d$ に対応するが，上限値である d についての考え方は研究者によって様々である．一般的に進化型計算の実行においては， $y^g = E(x^g)$ で d の値が定まる(ただし， $E(\cdot)$ は遺伝的操作を表す)．そこで本研究では $H(x^g, y^g) = d_{min}$ を考える．バイナリコーディングを用いる場合， $d_{min} = 1$ となり，1遺伝子座の値(1ビット)だけが異なる．このとき，中立ネットワーク Φ_g^* は以下のように表される．

$$\Phi_g^* = \{x^g \in \Phi_g | \exists y^g \in \Phi_g \text{ s.t. } (f_p \circ f_g)(y^g) = \\ (f_p \circ f_g)(x^g), H(x^g, y^g) = 1\}. \quad (3)$$

以降，記法の簡略化のため， Φ_g^* で表される異なる集合を Φ_i ($i = 0, 1, \dots, n$. ただし， n はネットワークの総数)と表す．

3 中立ネットワークに基づく構造解析法

式(3)に基づき，遺伝子型全体を各々の中立ネットワークに分類し，各ネットワークおよびそれらネットワーク間の構造を調べる．個々のネットワークについてはその大きさ($|\Phi_i|$)・適応度($f(\Phi_i)$)を，ネットワーク間については，次の $\min H(\Phi_i, \Phi_j)$ ， $p(\Phi_i, \Phi_j)$ を調べる； $\min H(\Phi_i, \Phi_j)$: 2つの中立ネットワーク Φ_i, Φ_j 間の最小ハミング距離 ($\min H(x^g, y^g), x^g \in \Phi_i, y^g \in \Phi_j$). $p(\Phi_i, \Phi_j)$: Φ_i, Φ_j 間のポータルの数 ($\min H(x^g, y^g) = 1$ となる (x^g, y^g) の組合わせの数)．

ここで，図1のような構造を考えると，(i) $\min H(\Phi_1, \Phi_2) = 1$ かつ $f(\Phi_1) \neq f(\Phi_2)$ ，つまり， Φ_1 と Φ_2 は連結している(図1では $f(\Phi_1) < f(\Phi_2) < f(\Phi_3)$ としている)，(ii) $\min H(\Phi_2, \Phi_3) \gg 1$ かつ $f(\Phi_2) \neq f(\Phi_3)$ ，つまり， Φ_2 と Φ_3 は連結していない，という2つの状況が存在すると想定される．(ii)では中立ネットワーク沿い($\Phi_1 \rightarrow \Phi_2$)に集団が移動していくとき Φ_3 にたどり着けない，進化型計算にとって難しい状況と考えられる．また， $\min H(\Phi_i, \Phi_j) = 1$

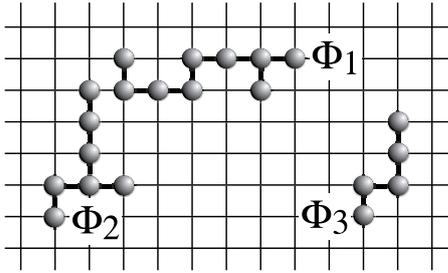


Fig. 1 ネットワーク間の関係

であっても， $p(\Phi_i, \Phi_j)$ が極端に小さい状況(図1の例では $p(\Phi_1, \Phi_2) = 1$)も進化型計算にとって難しい状況と考えられる．

4 ロイヤル・ロード関数の構造解析

本節では，その適応度景観に中立性が含まれることが知られているロイヤル・ロード関数(以下，RRF)⁵⁾について非常に簡単な景観パラメータ設定にした場合を例として，第3節で述べた中立ネットワークに基づく適応度景観の構造解析法について具体的に示す．

ロイヤル・ロード関数 ($N = 2, K = 2$) では，遺伝子長は4であり， $|\Phi_g| = 2^4 = 16$ となる．式(3)に基づき，遺伝子型を各々の中立ネットワークに分類すると以下ようになる． $\Phi_0 = \{0000, 0001, 0010, 0100, 0101, 0110, 1000, 1001, 1010\}$ ， $\Phi_1 = \{0011, 0111, 1011\}$ ， $\Phi_2 = \{1100, 1101, 1110\}$ ， $\Phi_3 = \{1111\}$ ．まず，個々のネットワークについて調べる．各ネットワークの大きさは， $|\Phi_0| = 9$ ， $|\Phi_1| = 3$ ， $|\Phi_2| = 3$ ， $|\Phi_3| = 1$ となる．また，各ネットワークが示す適応度は次のようになる： $f(\Phi_0) = 0$ ， $f(\Phi_1) = f(\Phi_2) = 1$ ， $f(\Phi_3) = 2$ ．

次にネットワーク間について調べる． $\min H(\Phi_i, \Phi_j)$ を表1に示す． $\min H(\Phi_0, \Phi_3) = 2$ となっており， Φ_0 から Φ_3 へは1ビットでは移動できないことがわかる．また，同じ適応度を示す Φ_1 と Φ_2 についても同様である． $p(\Phi_i, \Phi_j)$ を表2に示す． $p(\Phi_0, \Phi_1) = p(\Phi_0, \Phi_2) = 6$ となっており， Φ_0 からより適応度が高い Φ_1, Φ_2 へ個体集団が移りやすいと予想される．

これらのネットワーク構造を図2に表す．ここで，ノードの大きさは $|\Phi_i|$ を，エッジの種類は $p(\Phi_i, \Phi_j)$ を表し実線に近いものほど値は大きい．二重線で囲まれたノードは最大適応度を示したネットワークを表し，同じ形状で示すノードは同じ適応度を示すことを表す．表1,2が示すもの

Table 1 RRFにおける中立ネットワーク間の最小ハミング距離 ($\min H(\Phi_i, \Phi_j)$)

$\Phi_i \backslash \Phi_j$	Φ_0	Φ_1	Φ_2	Φ_3
Φ_0	0	1	1	2
Φ_1	-	0	2	1
Φ_2	-	-	0	1
Φ_3	-	-	-	0

Table 2 RRFにおける中立ネットワーク間のポータル数 ($p(\Phi_i, \Phi_j)$)

$\Phi_i \backslash \Phi_j$	Φ_0	Φ_1	Φ_2	Φ_3
Φ_0	0	6	6	0
Φ_1	-	0	0	2
Φ_2	-	-	0	2
Φ_3	-	-	-	0

と同様に Φ_1 または Φ_2 から Φ_3 へ移りにくいことがわかる．

5 Terraced NK ランドスケープの構造解析

5.1 テスト関数

本節ではNewmanらによって凸凹および中立性の二つの特徴量を調節できるテスト問題として提案されているTerraced NK ランドスケープ(以下TNK)⁷⁾を用いる．TNKは遺伝子長 N ，遺伝子間のエピスタシス数 $K (< N)$ ，中立性の度合いを調節する値 F を景観パラメータとしてもつ． K が大きくなると凸凹が増し， $K = N - 1$ のとき，凸凹が最大になる．また， $F = 2$ のとき，中立性の度合いは最大となり， $F \rightarrow \infty$ のとき，中立性の度合いは理論上0となる．

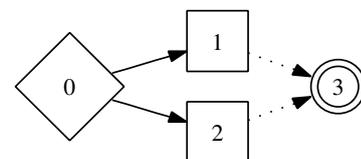


Fig. 2 RRFにおける中立ネットワークの構造

Table 3 TNK($K = 0$)におけるネットワークサイズとその個数

$F = 2$		$F = 8$	
size	Number	size	Number
16	64	2	512

5.2 計算機実験設定

Φ_g の全探索を行うことでTNKの構造解析を行う。全探索および中立ネットワークのクラスタリングに計算コストがかかることから、一般性を損なわないと考えられる範囲で小さな遺伝子型空間を扱う。そこで、TNKの景観パラメータを $N = 10$, $K = \{0, 9\}$, $F = \{2, 8\}$ に設定し、各 (K, F) の組合わせに関し10試行を行い、任意の1試行について結果を示す。なお、 $|\Phi_g| = 2^{10} = 1024$ となる。

5.3 計算機実験結果

各 (K, F) の組合わせに関し、各ネットワークのサイズおよびそのサイズのネットワークがいくつあるかを表3, 4に示す。 $K = 0$ では、すべての遺伝子型はいずれかの中立ネットワークに属し、そのネットワークのサイズは $F = 2$ のとき16, $F = 8$ のとき2と均一になっている。この結果から、 $K = 0$ のとき、 F が小さい、つまり、中立性の度合いが大きくなるとネットワークのサイズは大きくなるのがわかる。 $K = 9$ では、多くの遺伝子型が中立ネットワークを形成せず(表4のサイズが1のもの)、単独で存在する($F = 2$ のとき全遺伝子型の20%, $F = 8$ のとき全遺伝子型の66%)。またネットワークのサイズは不均一である。 $F = 2$ では $F = 8$ よりもネットワークの個数は少なくなっているが、サイズが大きいものが含まれている。

これらのネットワーク構造をグラフ作成ソフトGraphviz⁸⁾を用いて図3-6に示す。後述する各 (K, F) の組合わせによって形成される構造を比較するため、最大適応度を示すネットワークが1つとなる試行について結果を示している。ネットワークの総数が非常に多く(表3, 4)、本稿に全体を掲載することは難しい。このため最大適応度を示すネットワークを含み、高い方から上位4番目までの適応度を示すネットワークおよびそれらをつなぐエッジのみを示す。ノードの形状が

Table 4 TNK($K = 9$)におけるネットワークサイズとその個数

$F = 2$		$F = 8$	
size	Number	size	Number
1	210	1	680
2	39	2	92
3	10	3	19
4	11	4	15
5	6	5	4
6	2	6	2
7	1	-	-
8	2	-	-
9	1	-	-
11	1	11	1
14	1	-	-
15	1	-	-
16	1	-	-
22	1	-	-
72	1	-	-
212	1	-	-
226	1	-	-

○, ◇, □, ◎の順で適応度が高くなり、◎が最大適応度を表す。

$K = 0$ のとき、○で表されるノードから出たエッジは、途中で◇または□もしくはその両方を通り、すべて◎で表されるノードに到達している(図3,5)。ここで第1節の議論から、個体の移動は1ビットずつであり、中立ネットワーク内では個体集団は自由に移動していくことが可能でどのポータルに到達するかは等確率であると仮定すると、 $K = 0$ ではどのネットワークを通っても最終的に最大適応度を示すネットワークに到達できることが構造からわかる。一方、 $K = 9$ のとき、多くのノードは孤立している。 $(K, F) = (9, 2)$ のとき、□から◎に向かうエッジは存在しない。また、◇が端点となっている場合が多い(図6)。これはどのネットワークを通っていくかによって最終的に到達するネットワークが異なることを示している。 Φ_{max} は最大適応度を示すネットワークを表すとすると、 $p(i, max) = 0$ の場合、 Φ_i から Φ_{max} に個体集団が移動することは難しく、最大適応度を示すネットワークが孤立している、もしくは、“だ

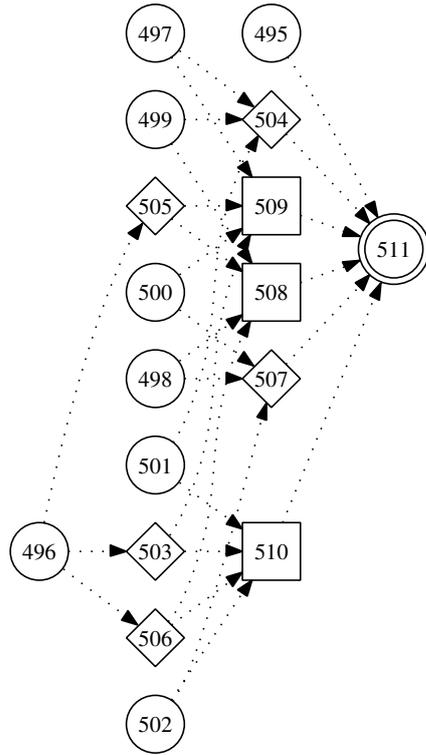


Fig. 3 TNK((K, F) = (0, 8))における中立ネットワークの構造

まし”のような状況になっていると考えられる。(K, F) = (9, 8)のとき, 上位4番目までの適応度を示すネットワーク間にエッジはなく, それぞれのネットワークは孤立している. この状態で個体集団が Φ_{max} に到達することは難しいと考えられる(図4). 最大適応度を示すネットワークが2つ以上ある試行においても, $K = 9$ のときの傾向は上述のものと同様であることを確認している.

6 おわりに

遺伝子型-表現型写像もしくは表現型-適応度写像が冗長であることから現れる中立性および従来議論されてきた凸凹を同時に含む適応度景観が実際にいかなる形状をしているのかは明らかにされてこなかった. 本研究では, まず, Harveyが定義した中立ネットワークの概念を集合として定式化した. そして, これに基づき凸凹と中立性を同時に含むテスト問題の解空間をクラスタリングし, その問題の適応度景観の形状, 中立ネットワークの構造および中立ネットワーク間の構造を調べた. 結果を以下にまとめる. (1)中立性が大きくなると, ネットワークのサイズが大きくなる. (2)凸凹が最も小さい問題では, どのネットワークを通っても最大適応度を示すネット

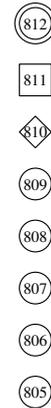


Fig. 4 TNK((K, F) = (9, 8))における中立ネットワークの構造

ワークに到達できる構造になっている. (3)凸凹が最も大きい問題では, 高い適応度を示すネットワークは, より高い適応度を示すネットワークへ移るパスを持たないものが多く, “だまし”のような状況になっていると考えられる. 今後, 本研究で用いた手法を実問題に適用する予定である.

参考文献

- 1) T. Smith, P. Husbands and M. O'Shea: Neutral networks in an evolutionary robotics search space; *Proceedings of the 2001 IEEE Congress on Evolutionary Computation: CEC2001*, pp. 136–145, IEEE Press (2001)
- 2) A. Thompson: An evolved circuit, intrinsic in silicon, entwined with physics; *Proceedings of the First International Conference on Evolvable Systems: From Biology to Hardware*, pp. 390–405 (1996)
- 3) I. Harvey and A. Thompson: Through the labyrinth evolution finds a way: A silicon ridge; *Proceedings of the First International Conference on Evolvable Systems: From Biology to Hardware*, pp. 406–422 (1996)
- 4) F. Rothlauf and D. Goldberg: Redundant representations in evolutionary computation; *Evolutionary Computation*, Vol. 11, No. 4, pp. 381–415 (2003)
- 5) E. Nimwegen, J. Crutchfield and M. Mitchell: Statistical dynamics of the royal road genetic algorithm; *Theoretical Computer Science*, Vol. 229, No. 1, pp. 41–102 (1999)
- 6) L. Barnett: Tangled Webs: Evolutionary Dynamics on Fitness Landscapes with Neutrality; MSc. dissertation, School of Cognitive and Computing Sciences, Sussex University (1997)
- 7) M. Newman and R. Engelhardt: Effect of neutral selection on the evolution of molecular species; *Proceedings of the Royal Society of London B*, 256, pp. 1333–1338, Morgan Kaufmann (1998)
- 8) <http://www.graphviz.org/>

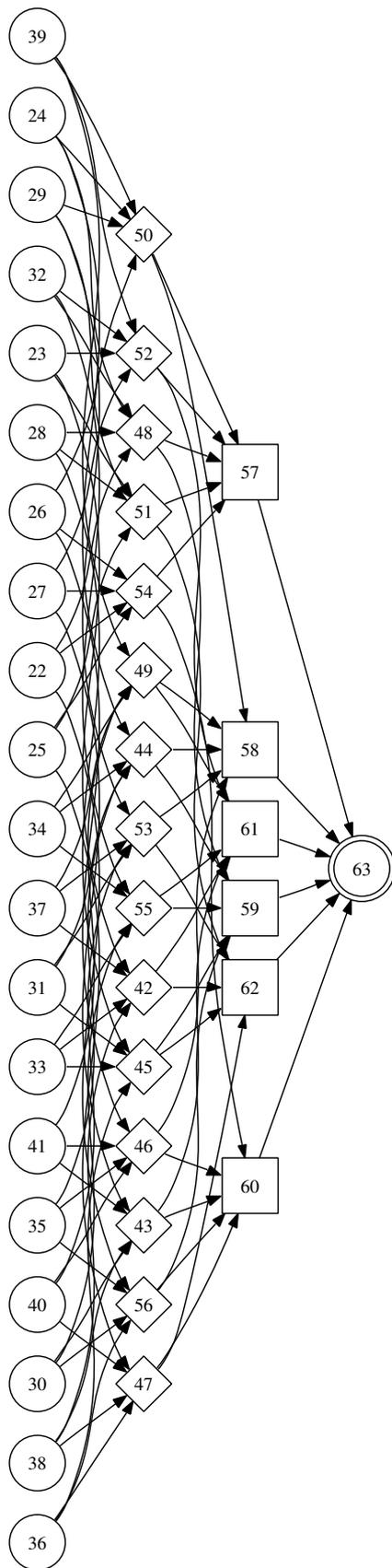


Fig. 5 $TNK((K, F) = (0, 2))$ における中立ネットワークの構造

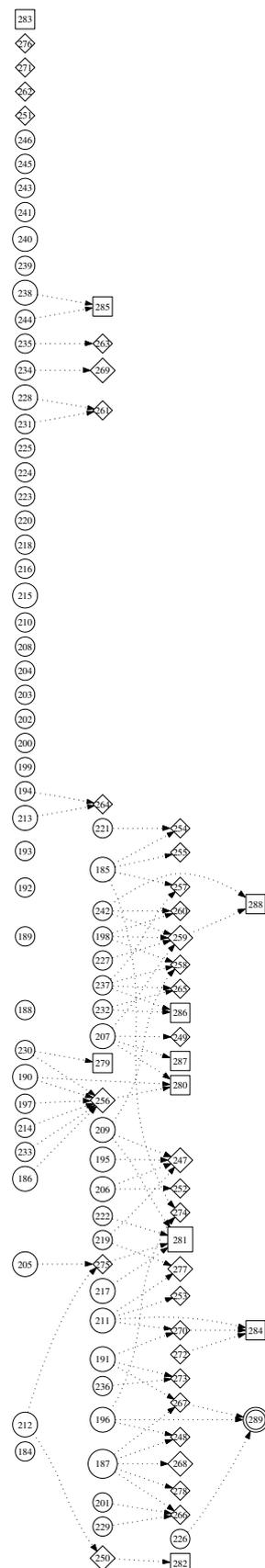


Fig. 6 $TNK((K, F) = (9, 2))$ における中立ネットワークの構造