

車車間通信の局所的情報による運転戦略の学習

Learning Deceleration Strategies via Utilizing Local Information of Vehicle's

石川翔太 *¹ 荒井幸代 *¹
Ishikawa Shota Arai Sachiyo

*¹千葉大学大学院工学研究科都市環境システムコース
Graduate School of Engineering, Chiba University, Division of Urban Environment Systems

The optimal driving strategy that intended to resolve traffic congestion is decided in a situation where the vehicle can get entire road condition through the road-to-vehicle communication. However, the road-to-vehicle communication is taken a lot of money to improve infrastructure. So the purpose of this study, in this paper, is to propose a method for acquiring the driving strategy when the intelligent vehicle decelerate to resolve traffic congestion effectively utilizing local information of the inter-vehicle communication that need no infrastructure. The driving strategy is acquired by reinforcement learning. We try to achieve a phase transition that shifts traffic flow from congestion phase to metastable phase by introducing the intelligent vehicle with learned driving strategy.

1. はじめに

文献 [辻野 13] によれば, 自動車の自動運転化がもたらす効果は, 人的事故減少, 渋滞解消・緩和, 環境負荷軽減, 高齢者等の運転補助に分類される. 本研究ではこのうち, 渋滞解消・緩和に向けた自動運転を実現するために必要な自動車間の情報共有に着目する.

渋滞解消・緩和の実現が期待できる自動運転技術として先行車との車間距離や速度を制御する Adaptive Cruise Control(以下 ACC と記す) が挙げられる [Arne 08, Florian 12, Kshitij 12]. 近年, ACC に先行車との通信機能を持たせた Cooperative Adaptive Cruise Control(以下 CACC と記す) の開発が進められ, 先行車の加減速情報を共有することによる精緻な制御が期待されている [Bart 06]. これらの制御と同様に筆者らは, インフラの拡張や整備によるのではなく, 自動車の走行ルールの修正によって自然渋滞を解消・緩和することを目指している.

筆者らは, これまで, Nagel-Schreckenberg(以下 NS と記す) モデル [Nagel 92] を一般化し, 情報共有できる先行車の台数を任意に設定できること, および, それらの情報をもとに速度を変更していくことを走行ルールに加えた Generalized Nagel-Schreckenberg(以下 GNS と記す) モデルを提案した. このモデルによる計算機実験の結果から, 先行車情報を共有して走行することが自然渋滞の抑制に有効であること, さらに, 共有台数の増加が渋滞からの復帰時間を早めることが示された. しかし, 先行車情報を共有して走行することができても, 車両間隔の均等性が保たれなければ, 均等性を保っているときよりも交通流量が小さくなってしまふことも示している.

そこで, 本研究では, 車両間隔が不均等な交通流を均等に修正していく運転戦略を考える. 文献 [許 13] では, 学習ペースカーを導入することによって, 自然渋滞の状況にある交通流をメタ安定相へ移行させるが, このペースカーは道路全体の交通状況を常に把握する必要がある. しかし, 走行ルールへ交通情報を導入することにおいて, 例えば, 道路全体の状況を把握するための路車間の通信が考えられるが, 通信インフラの拡張や設置をするコストがかかることなど課題がある. これらの理由から, 本研究では, 車車間通信によって得られる局所的情報を

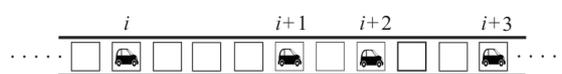


図 1: 自動車 i に対する先行車の表記法

用いた運転戦略の獲得を目的とする. また, 車車間通信によって共有できる先行車の情報が増えることが, 獲得する運転戦略に与える影響を示す.

以下, 2 章では, 対象問題として, 先行車情報を共有しながら走行ルールである GNS モデルを説明し, 車両間隔の均等性と交通流量の関係を示す. 3 章では, 提案手法として車車間通信の局所的情報による運転戦略の獲得法を説明する. 4 章では計算機実験の設定を示し, 5 章において本研究のまとめ.

2. 対象問題

2.1 準備

■**道路モデル** 本論文の道路モデルは図 1 に示す一次元 (単車線) で, 空間の両端を連続した空間として考える周期的境界条件の環境とする. 周期的境界条件は, セルオートマトンを用いた交通流の研究では一般的であり [Choudhury 00, Nagel 96, Nagel 98, Rajewsky 98], 各時刻 t において道路上の車両密度を一定に保持することができるため, 本論文でもこれを適用する. また, ある自動車 i に対して, 直前の先行車から順に前方の先行車を $i+1, i+2, i+3$ と添え字を 1 ずつ増やして表記する.

■**交通流の基本図 (Fundamental Diagram)** 車両密度 (横軸) と交通流 (縦軸) の関係を表したグラフは, 交通流の基本図と呼ばれ, 実際の交通流は大きく 3 つの相に分けられる. 1 つ目は, 自動車の流れがスムーズでほぼ一定の速度で走行している状況を示し, 車両密度の増加にしたがって交通流量が増加する特性をもつ「自由走行相」である. 2 つ目は, 渋滞している状況を示し, 車両密度の増加にしたがって交通流量が減少する特性をもつ「渋滞相」である. このとき, 自由走行相から渋滞相へと移行する境界の車両密度を「臨界密度」と呼ぶ. そして 3 つ目は, 臨界密度以上の車両密度に至っても, 交通流量が増加し続ける特性をもつ「メタ安定相」であるメタ安定相下では, 車が車間を一定の間隔で詰めて走行しており, 輸送効率の

GNS モデルの走行ルール

$$\text{GNS}(x_i(t), v_i(t), d_i, i, \text{size}^{\text{share}})$$

1. Decide speed

$$c \leftarrow i + \text{size}^{\text{share}}$$

$$v_i(t+1) \leftarrow \text{CHANGE SPEED}(v_i(t), d_i, i, c)$$

2. Stochastic deceleration

$$v_i(t+1) \leftarrow \max(v_i(t+1) - 1, 0) \text{ with probability } p$$

3. Movement

$$x_i(t+1) \leftarrow x_i(t) + v_i(t+1)$$

$$\text{CHANGE SPEED}(v_i(t), d_i, i, c)$$

I. Acceleration

$$v_i(t+1) \leftarrow \min(v_i(t) + 1, v_{\text{limit}})$$

$$\text{if } (v_i(t+1) \leq d_i)$$

$$\text{return } v_i(t+1)$$

$$\text{end if}$$

II-a. Cooperative

$$\text{if } (i+1 \leq c)$$

$$v_{i+1}^{\text{share}} \leftarrow \text{CHANGE SPEED}(v_{i+1}(t), d_{i+1}, i+1, c)$$

$$\text{if } (v_{i+1}^{\text{share}} > v_{i+1}(t))$$

$$v_{i+1}^{\text{pred}} \leftarrow v_{i+1}(t)$$

$$\text{else}$$

$$v_{i+1}^{\text{pred}} \leftarrow \max(v_{i+1}^{\text{share}} - 1, 0)$$

$$\text{end if}$$

II-b. Non cooperative

$$\text{else}$$

$$v_{i+1}^{\text{pred}} \leftarrow \max(\min(d_{i+1} - 1, v_{i+1}(t), v_{\text{limit}} - 1), 0)$$

$$\text{end if}$$

$$\text{return } \min((v_i(t+1), v_{i+1}^{\text{pred}} + d_i)$$

- v_{limit} : 制限速度
- p : 減速確率
- $\text{size}^{\text{share}}$: 情報を共有できる先行車の台数
- c : 自動車 i が情報を共有できる最前方の先行車
- v_{i+1}^{share} : 自動車 i が共有した先行車 $i+1$ の速度

図 2 に示す GNS モデルの走行ルールを説明する。

GNS モデルの走行ルールは 1. 速度決定 (Decide speed), 2. 確率的減速 (Stochastic deceleration), 3. 移動 (Movement) の 3 つの手順から成る。はじめに、「速度決定」では自動車 i に対して、情報共有できる先行車の台数 $\text{size}^{\text{share}}$ を基に情報共有できる最前方の先行車 c を設定する。そして、CHANGE SPEED のルールに従って $v_i(t+1)$ が決定する。次に、「確率的減速」では確率 p により $v_i(t+1) \leftarrow \max(v_i(t+1) - 1, 0)$ の減速を行う。最後に「移動」により「速度決定」「確率的減速」で決められた速度 $v_i(t+1)$ だけ前方のセルへ移動する。

CHANGE SPEED は、自動車 i の時刻 $t \rightarrow t+1$ の変化に伴う速度 $v_i(t+1)$ を決定する。 $v_i(t+1)$ は I. 加速の手順に従って決められるが、加速後の速度に対して自動車 i が十分な車間距離をもたない場合は、情報共有の有無によって II-a. 協調あり, II-b. 協調なしの 2 つに分岐する。

I. 加速 (Acceleration)

速度 $v_i(t+1) \leftarrow v_i(t) + 1$ とし、 $v_i(t+1) \leq d_i(t)$ を満たせば $v_i(t+1)$ が速度となり、満たされなければ「協調」と「協調なし」に分岐する。先行車 $i+1$ が情報の共有可能な最前方の先行車 c の後方または c であれば、「協調」の手順を、 c よりも前方に存在すれば、「協調なし」の手順に従う。

II-a. 協調 (Cooperative)

先行車 $i+1$ に CHANGE SPEED を適用させ、自動車 i が先行車 $i+1$ と共有した速度 v_{i+1}^{share} を決定する。このとき、 $v_{i+1}^{\text{share}} > v_{i+1}(t)$ を満たすとき、先行車 $i+1$ は減速しても $v_{i+1}(t)$ 進むので、先行車の予測速度は $v_{i+1}^{\text{pred}} \leftarrow v_{i+1}(t)$ となる。 $v_{i+1}^{\text{share}} > v_{i+1}(t)$ を満たさなければ、先行車 $i+1$ の減速を考慮し、先行車の予測速度は $v_{i+1}^{\text{pred}} \leftarrow v_{i+1}^{\text{share}} - 1$ となる。

II-b. 協調なし (Non cooperative)

ExNS と同様に先行車の予測速度 v_{i+1}^{pred} を決定する。すなわち、時刻 $t \rightarrow t+1$ への変化に伴い、自動車 i は先行車 $i+1$ の減速を考慮して先行車の予測速度を $v_{i+1}^{\text{pred}} \leftarrow \max(\min(d_{i+1}(t) - 1, v_{i+1}(t), v_{\text{limit}} - 1), 0)$ とする。

「協調」または「協調なし」のいずれかの手順を経た場合、自動車 i の速度 $v_i(t+1)$ は $\min(v_i(t+1), v_{i+1}^{\text{pred}} + d_i(t))$ によって決定される。

2.3 車両間隔の均等性が交通流に与える影響

本節では、先行車情報を共有しながら走行するとき、車両間隔の均等性が交通流に与える影響を示す。図 3 は情報共有できる先行車の台数 $\text{size}^{\text{share}}=1$ の実験設定における、横軸が車両密度 ρ と縦軸が交通流量 q の関係を示している。初期配置をランダムに配置した GNS Random は、臨界密度の車両密度 $\rho = 0.2$ 以上の車両密度のとき、メタ安定相と渋滞相の両方のプロットを確認することができる。一方、初期配置を車両間隔 1 に等間隔に配置した GNS Homogeneous は、臨界密度の車両密度 $\rho = 0.2$ 以上の車両密度でもメタ安定相だけ示している。図 3 の結果が示す通り、先行車情報を共有しながら走行

図 2: GNS モデルの走行ルール

高い状態といえる [酒井 06].

2.2 GNS モデル

GNS モデルは、車車間通信によって先行車の「速度」と「車間距離」を共有する。得られた先行車の情報から先行車 $i+1$ の先読みをすることにより、時刻 $t \rightarrow t+1$ の変化に伴う先行車の予測速度 v_{i+1}^{pred} を決定する。また、GNS モデルは、「情報を共有できる先行車の台数」を任意に設定できるため、設定する台数によって異なる自動車の挙動を観察することができる。GNS モデルにおける用語を以下に定義する。

- $v_i(t)$: 時刻 t における自動車 i の速度
- $v_{i+1}(t)$: 時刻 t における先行車 $i+1$ の速度
- v_{i+1}^{pred} : 自動車 i による先行車 $i+1$ の予測速度
- $x_i(t)$: 時刻 t における自動車 i の座標
- d_i : 自動車 i と先行車 $i+1$ の車間距離
- d_{i+1} : 先行車 $i+1$ と先行車 $i+2$ の車間距離

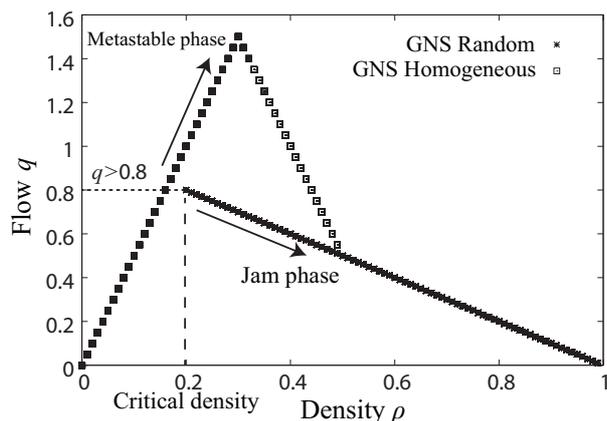


図 3: 初期配置が GNS モデルの交通流の基本図に与える影響 (情報共有できる先制車の台数 $size^{share}=1$, 減速確率 $p=0$)

しても、車両間隔が均等でなければメタ安定相へ移行しない場合が確認できる。

2.4 強化学習

強化学習 [Sutton 98] は、教師なし学習の一種であり、試行錯誤を通じて、環境から与えられる報酬値を最大にする行動を選択するための方策 (policy) を獲得する。本研究では、強化学習の一つである Q 学習を用いる。

Q 学習では、状態集合を $S = \{s_i | i = 1, 2, \dots, n\}$ 、行動集合を $A = \{a_j | j = 1, 2, \dots, m\}$ として定義する。学習主体は各時刻 t において状態 $s_t \in S$ を観測し、状態観測から行動出力へのマッピングである方策に基づき行動 $a_t \in A$ を決定し、出力する。行動出力により環境は $s' = s_{t+1}$ へ状態遷移し、その遷移に応じた報酬 r_t を学習主体に与える。報酬 r_t を獲得した学習主体は更新式 $Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha(r_t + \gamma \max_{a'} Q(s_{t+1}, a'))$ を用いて状態 s_t と行動 a_t の Q 値 $Q(s_t, a_t)$ を更新する。ここで、学習率 α は Q 値の更新の度合いを示すパラメータであり、割引率 γ は将来に得られると期待される報酬が現在においてどれだけの価値があるかを決定するパラメータである。

3. 提案手法

本論文では、周期的境界条件におけるセルオートマトンを用いた交通流モデルを考える。GNS モデルに従えば、減速確率 $p=0$ の車両群から構成される交通流は、一旦渋滞相に陥ったら、交通流量が変化せず、渋滞相となり続ける。一方、メタ安定相が観測できた交通流において、メタ安定相状態の交通流量を継続し続ける現象が観察されている。つまり、渋滞相となった交通流をメタ安定相へと遷移させることができれば交通流量を増加させることができる。本論文では、メタ安定相へと遷移させる方策 (運転戦略) を強化学習によって獲得する。この強化学習モデルを以下に定式化する。また、運転戦略の獲得アルゴリズムを図 4 に示す。

- 状態：車車間通信によって入手できる局所的情報は、「速度」と「車間距離」だけを用いて状態を表現する。道路側から道路全体の情報を入手できる場合は、「車両密度」を用いて状態を表現する。目標状態 s_{goal} はメタ安定相となった時点の状態とする。

Algorithm of Learning

```

Initialize  $Q(s, a)$  arbitrarily
Repeat (for each episode):
  Initialize vehicle placement
  if  $t = t_{init}$ 
    start episode
  if  $t \bmod T = 0$ 
    Choose  $a$  using policy derived from  $Q(s, a)$ 
    if  $s_t = s_{goal}$ , end episode
    if  $t = t_{max}$ , end episode
  else if  $t \bmod T = T - 1$ 
    Observe  $s_{t+1}$ 
    Observe reward  $r$ 
     $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_t + \gamma \max_{a'} Q(s_{t+1}, a) - Q(s_t, a_t))$ 
     $s_t \leftarrow s_{t+1}$ 

```

図 4: 強化学習による運転戦略の獲得アルゴリズム

- 行動：図 2 に示す GNS の走行ルールのなかには、確率 p で減速する「確率的減速」がある。この確率 $p=0$ とすることで、あえて減速を行う行動「戦略的減速」をルールの中に取り入れる。よって、学習主体は行動として「戦略的減速を行う」「減速を行わない」の 2 つの行動を行う。
- 報酬：メタ安定相時に獲得できる交通流量、すなわち、各車両密度 ρ において実現可能な最大の交通流量 q に達した時に報酬を与える。
- エピソード：車両はセル (道路) 上に車両密度 ρ となる台数分がランダムに配置され、それぞれ初期速度は 0 とし、GNS モデルの走行ルールに従い走行する。数ステップ動かして、交通流が一定になった時点をエピソードの開始時刻 $t=0$ とする。なお、エピソードの開始時刻の車両密度は渋滞相の車両密度とする。もし、メタ安定相の車両密度であった場合、ランダムの配置からやり直す。
- 最大ステップ数 t_{max} ：全てのエピソードで目標とする流量に達するわけではない。そこで、実験上、最大ステップ数を決めておき、これを越えて目標状態に至らなかった場合には、最大ステップ数が経過し時点でエピソードを終了する。
- 観測周期 T ：上記「ステップ」とは、セルオートマトンモデルにおいて各車両が時刻 t から $t+1$ に遷移する単位を示しているが、Q 学習の Q 値は毎ステップ更新するのではなく、複数のステップ数からなる周期ごとに更新する。なぜなら、ペースカーの行動の影響が後続車両に伝播するには、ある一定の時間を要する。そこで、ペースカーの観測周期を T として、車両 N 台が存在する場合は、 $T > N$ を満たす値を設定する。

4. 計算機実験

4.1 実験環境

計算機実験における道路と運転者の設定は以下である。

- 道路モデル：一次元周期的境界条件
- 道路長 L ：100 セル

- 制限速度 $v_{\text{limit}} : 5[\text{セル/ステップ}]$

道路モデルは、2.1 節に示した通り、一次元の周期的境界条件を適用する。このとき、道路の 1 セルに最大 1 台が存在すると考える。そこで、実際の車長 (軽自動車 3.4[m], 小型車が 4.7[m] 以下) に基づいて、道路長の 1 セルは 5[m] とする。また、1 ステップを 1 秒とし、 $v_{\text{limit}} = 5[\text{セル/ステップ}]$ は 90[km/h] に相当する。この値は実際の高速道路制限時速 (普通車・軽自動車 100[km/h], 大型貨物車が 80[km/h]) に相当する。

なお、セル数 $L \times$ 車両密度 ρ の数の車両は、全てランダムに初期配置される。

4.2 実験設定

道路上のある観察対象区間の交通流量 q , 車両密度 ρ , 平均速度 $\bar{V}(t)$ を定義する。交通流量 q は、単位時間内に道路上の一点を通過した自動車の台数で定義され、ある時刻 t から $t + \Delta t$ の間に道路上の一点を通過した自動車の台数を M とすると交通流量は $q = \frac{M}{\Delta t}$ となる。

車両密度 ρ は、道路上に存在する単位距離当たりの自動車の台数で定義され、道路長を L , 道路上に存在する自動車の台数を N とすると、車両密度は $\rho = \frac{N}{L}$ により与えられる。

平均速度 $\bar{V}(t)$ は、道路上に存在する全ての自動車の速度の平均値で定義され、ある時刻 t における自動車 i の速度を $v_i(t)$, 道路上に存在する自動車の台数を N とすると、平均速度は $\bar{V}(t) = \sum_i^N \frac{v_i(t)}{N}$ により与えられる。

これらの車両密度 ρ と平均速度 $\bar{V}(t)$ と交通流量 q との間には式 (1) の関係がある。

$$q = \rho \bar{V}(t) \quad (1)$$

4.3 局所的情報が運転戦略に与える影響の観察

文献 [許 13] では、交通流量を状態とすることにより、道路全体の状況を把握できていたが、本研究の学習主体の観測できる状態は「速度」と「車間距離」の局所的情報に限る。計算機実験を通して、道路全体の情報を用いて学習する場合と、局所的情報を用いて学習する場合を比較し、学習主体が獲得できる運転戦略に与える影響を観察する。また、情報を共有できる先行車の台数が増加すれば、学習主体が観測できる局所的情報も増加するので、局所的情報の大きさが運転戦略に与える影響も観察する。

4.4 メタ安定相へ転移までに生じる損失の観察

渋滞相にある交通流をメタ安定相へ転移させる場合、渋滞相からメタ安定相になるまでに要した時間の渋滞解消時間と、損失した流量が計算できる。渋滞相の交通流量を $flow_{step}^{jam}$, メタ安定相へ転移させるまでに遷移した各ステップの交通流量を $flow_{step}$, 学習主体が行動し始めたステップ数を $step = start$, メタ安定相へ転移した時のステップ数を $step = end$ とすると、損失した流量は $\sum_{step=start}^{end} (flow_{step} - flow_{step}^{jam})$ の式を用いて計算できる。道路全体の状況を把握した場合、局所的情報を用いた場合それぞれの運転戦略における、渋滞解消時間と損失した流量の関係を示す。

5. まとめ

本研究は、先行車情報を共有しながら走行したとしても、車両間隔の均等性が交通流に大きな影響を及ぼすことに着目し、車両間隔が不均等な交通流を均等に修正していく運転戦略を獲得することを目的としている。運転戦略を獲得する際、道路側

からの路車間通信による道路全体の情報を用いずに、車車間通信による局所的情報だけを用いる状況を想定している。運転戦略の獲得には強化学習を用いる。道路全体の情報を用いて獲得した運転戦略と比較して、局所的情報を用いた運転戦略が交通流に与える影響、および、メタ安定相へ転移までに生じる損失の観点から計算機実験を用いて検証する。

参考文献

- [Arne 08] Arne Kesting, Martin Treiber, Martin Schönhof, Dirk Helbing: Adaptive cruise control design for active congestion avoidance, *Transportation Research Part C, Emerging Technologies*, Vol.16, No.6, pp.668–683(2008)
- [Bart 06] Bart van Arem, Cornelia J. G. van Driel, Ruben Visser: The Impact of Cooperative Adaptive Cruise Control on Traffic-Flow Characteristics, *Intelligent Transportation Systems*, *IEEE Transactions*, Vol.7, No.4, pp.429–436(2006)
- [Choudhury 00] D. Choudhury, L. Santen, and A. Schadschneider: Statistical physics of vehicular traffic and some related systems, *Physics Reports*, Vol.329, No.4-6, pp.199–329(2000)
- [Florian 12] Florian Knorr, Michael Schreckenberg: Influence of inter-vehicle communication on peak hour traffic flow, *Physica A, Statistical Mechanics and its Applications*, Vol.391, No.6, pp.2225–2231(2012)
- [Kshitij 12] Kshitij Jerath, Sean N. Brennan: Analytical Prediction of Self-organized Traffic Jams as a Function of Increasing ACC Penetration, *Intelligent Transportation Systems*, *IEEE Transactions*, Vol.14, No.4, pp.1782–1791(2012)
- [許 13] 許海遲, 荒井幸代: 学習ベースカーによるメタ安定相への遷移の実現, *電気学会論文誌 C*, Vol.133, No.9, pp.1709–1716(2013)
- [Sutton 98] R. S. Sutton and A. G. Barto: Reinforcement Learning: A Introduction, *The MIT Press*(1998)
- [Nagel 92] K. Nagel, M. Schreckenberg: A cellular automaton model for freeway traffic, *J.Phys.IFrance*, Vol.2, No.12, pp.2221–2229(1992)
- [Nagel 96] K. Nagel: Particle hopping models and traffic flow theory, *Phys. Rev. E*, Vol.53, No.5, pp.4655–4672(1996)
- [Rajewsky 98] N. Rajewsky, L. Santen, A. Schadschneider, and M. Schreckenberg: The asymmetric exclusion process: Comparison of update procedures, *J. Statistical Physics*, Vol.92, No.1-2, pp.151–194(1998)
- [酒井 06] 酒井聡士, 西成活裕, 飯田晋司: 新しい確率交通流セルオートマトンモデルが示す渋滞相転移, *日本応用数理学会論文誌*, Vol.16, No.4, pp.371–384(2006)
- [辻野 13] 辻野照久, 坪谷剛: 自動運転自動車の研究開発動向と実現への課題, *科学技術動向* 1・2月号, pp.9–16(2013)