

1H-5

直列化可能性の新しいモデル

中津 楯男

愛知教育大学情報科学コース

1. はじめに

データベースの並行制御において、最も基本となる問題は、ある1つの実行順がデータベースに矛盾を引き起こすかどうか(直列化可能性)を判定することである。この判定問題は一般的にはNP完全であることが証明され¹⁾、その後、多項式時間でその判定が行えるできるだけ大きな部分クラスの発見に焦点が移った¹⁾。この報告では、直列化可能性を判定するための新しいモデルの提案と、それにもとづき、これまでより広い多項式時間認識可能なクラスを発見したので報告する。また従来の結果の多くは、このモデルから簡単に導くことができる。

2. 基本定義とこれまでの結果

データベースは実体(entity)の集合とし、 $D = \{x, y, \dots\}$ と書く。

[定義1] トランザクションはアクションと呼ばれるシンボルの有限列である。アクションは $R(x)$ または $W(x)$ ($x \in D$)であり、それぞれ実体 x を読む、実体 x に書き込むことを意味する。 $R(x)$ を伴わない $W(x)$ を特に盲目的書き込み(blind write)という。 x に書き込む値は $W(x)$ に先行する読みだし命令で読み出された値すべてに依存するものとする。□

複数のトランザクションが並行して実行される場合、各アクションが入り交じりながら実行される。その場合の各アクションの実行順序を履歴(history)という。

履歴を考える場合、仮想的なトランザクション T_0 と T_r を考えると便利なおことが多い。 T_0 は履歴 h 内で読み出されるす

べての実体に初期値を書き込み、 T_r は h 内で書き込まれた実体をすべて読み出すトランザクションである。これ以後履歴は T_0 の実行後に h が実行され、最後に T_r が実行されるものとする。

直列化可能な履歴の集合をSRとしたとき、与えられた履歴がSRに含まれるかどうかの判定は一般にはNP完全であるため、次の研究の興味は直列化可能な部分クラスで、多項式時間で判定可能なクラスを求めることになる。Papadimitriouは競合直列化可能(conflict serializable, DSR)というクラスを考えた²⁾。

茨木らはより広いクラスとしてWRWクラスを見つけた。またDSRが茨木らの定義したクラスWWと一致することも証明している。これらの包含関係を図1に示す¹⁾。

3. 直列化可能性の新しいモデル

[定義2] 与えられたスケジュール h において、とくに実体 x にアクセスするアクションだけを考える。トランザクションに対して1つの節点を対応させ、 h が $W_1(x) \cdots R_j(x)$ を部分系列として含む(ただし、その間に $W(x)$ は無い)時、 j read-from i といい、 i より j へ枝を結ぶ。1つのトランザクションを1つの節点对応させれば、このread-from関係をグラフの形で表現できる(これを x に対するread-fromグラフといい、 $G(x)$ と書く)。 $G(x)$ によれば、 x にアクセスするトランザクション集合は $G(x)$ の連結成分に対応させて C_0, C_1, \dots, C_r と分割される。ただし、 C_0 は仮想的な初期トランザクション T_0 、 C_r は仮想的な最終トランザクション T_r

A new model for the database serializability

Narao NAKATSU

Aichi University of Education

を含む成分である。□

このとき次の補題が成立する。

[補題 1] h が直列化可能ならば、任意の x の $G(x)$ の連結成分 C_i について、 C_i は木であり、 C_i の同一レベルの節点の内、 x に write する節点は高々 1 つである。従ってどのレベルにおいても子孫を持つ節点は高々 1 つにかぎる。特に C_r においては、葉に write する節点があつてはならない。

[補題 2] 履歴 h が直列化可能とする。任意の $G(x)$ とその 2 つの連結成分を考える。 h と等価な任意の直列スケジュールにおいて、 $\exists s, s \in C_1, \exists t, t \in C_2, s < t$ ならば $\forall u \in C_1, \forall v \in C_2, u < v$ である。 □

read-from の関係のほかには何らかの制約 c をもうけてトランザクションの間に順序を規定することもできる。多項式時間判定可能な十分条件として次の定理を示すことができる。

[定理 1] 任意のデータ項目について、各連結成分の順序が 1 意であるならば、スケジュールの直列化可能性は多項式時間で判定できる。特に任意の実体 x について、 $G(x)$ が 1 つの連結成分からなるばあい、スケジュールの直列化可能性は多項式時間で判定できる。 □

定理 1 を用いて次のことが示せる (証明は [4] を参照のこと)。

[系 1] ¹⁾ クラス WW は多項式時間で認識できる。

[系 2] ^{1), 3)} 盲目的書き込みが許されない場合、 $SR=WRW=DSR$ となる。

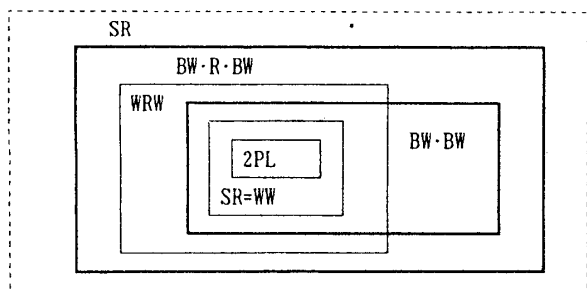


図1 クラスの包含関係 (細線は従来知られているクラス、太線は本稿での提案、点線はNP完全であるクラス)

定理 1 を満足させるための制約をいろいろ考えることでさまざまなクラスを定義できる。

[系 3] 任意の x について、盲目的書き込みを行うトランザクション間の順序を出現順にあわせて規定する。この順序を read-from グラフに加えても閉路持たない時、そのクラスを $BW \cdot BW$ とする。 $BW \cdot BW$ は多項式時間で判定可能な直列化可能なスケジュールのクラスである。しかも $BW \cdot BW \supset WW$ である。

[系 4] 任意の x について、盲目的書き込みを行うトランザクションと x を読み出すトランザクション間の順序を出現順にあわせて規定する。この順序を read-from グラフに加えても閉路持たない時、そのクラスを $BW \cdot R \cdot BW$ とする。 $BW \cdot R \cdot BW$ は多項式時間で判定可能な直列化可能なスケジュールのクラスである。しかも $BW \cdot R \cdot BW \supset WRW$ である。

4. 先読みスケジューラ²⁾

トランザクションの形が前もってわかっている場合、直列化可能な履歴を生成するスケジューラを先読みスケジューラという。これまで多項式時間で実行可能な先読みスケジューラとして、クラス WW に入る履歴を生成するものが知られていた。より広いクラスに入る履歴を生成する多項式時間実行可能なスケジューラについて次のことがわかる。

[定理 3] クラス $BW \cdot BW$ について、多項式時間実行可能な先読みスケジューラが存在する。

[参考文献] [1] Ibaraki 他, "Serializability with Constraints", ACM TODS, 12, 3, 1987.

[2] Ibaraki 他, "Cautious Transaction Schedulers for Database Concurrency Control", IEEE SE, 14, 7, 1988.

[3] Papadimitriou, C. H., The Theory of Database Concurrency Control, Computer Science Press, 1986.

[4] 中津, 「直列化可能性のための新しいモデル」、信学研資、COMP91-78, 1991.