グラフ上の変形操作を用いた経路関数従属性のカバーリング導出一手法

3V－3

山田 光博

NTT情報通信研究所

1．まえがき

Wedell等によって複合オブジェクトモデルが提唱されている[1,2,3]．特に文献(1)では，関係モデルにおける関数従属性に対応する経路関数従属性を提案し，その健全かつ完全な導出律を示している．文献(1)では更に，これらの導出律に関する閉包を求める手法を提案している．しかし経路関数従属性の左辺から冗長な属性を排除する手法は提案されていない．

本稿では，経路関数従属性の集合Fが与えられた場合に，各経路関数従属性の左辺から冗長な属性を排除し，更に冗長な経路関数従属性をFから排除したカバーリングを求める手法を提案する．提案手法の特徴は，クラススキーマおよびその上の経路関数従属性を，文献(4)と同様の超グラフに類似のモデルで表現し，このモデル上の変形操作を用いてカバーリングを求める点である．

2．基本的事項

Sをクラスの集合とし，Pijをクラスciからクラスcjへの一価関数とする．このときクラススキーマを$\cup_{Ci\in S}$ {Pi1:C1, …Pij:Cj, …, Pin:Cn} で表現する[1]．また，経路関数（Path Function(pf)）を以下で定義する．

［経路関数］[1]

以下を満たす最小の集合PFを経路関数集合とする．

(1) Id∈PF, ∀i∀jに対してPij∈PF

(2) pf1, pf2∈PFのときpf1の終点＝pf2の始点⇒pf1とpf2の合成関数pf2(pf1)=pf1 o pf2∈PF. 以降クラスCを始点とする経路関数の集合をPF(C)で表す．

［クラススキーマのインスタンス］[1]

以下を満たす有向グラフG=G(V,E)をクラススキーマのインスタンスと呼ぶ．(1)各クラスのインスタンスに対して，頂点が唯一に対応し(2)クラス間経路関数に始点および終点のクラスのインスタンスに対応する頂点間の有向枝が対応する．

［経路関数従属性（PFD）］[1]

クラスCを始点とする経路関数pf1, …, pfnに対して経路関数従属性C（pf1…pfm→pfm+1…pfn）の成立を以下で定義する．

u, vをその所属するクラスがCであるG上の頂点とする．u.pfi＝v.pfi（1≦i≦m）（但し，u.pfi, v.pfiは有向枝pfiの始点をそれぞれu, vとした場合の終点．）のとき，u.pfj＝v.pfj（m+1≦j≦n）が成立する．

［経路関数従属性の導出規則］[1]

X,Y,Z等で経路関数の集合を表し，Fで経路関数従属性の集合を表すものとする．

PF1 Y⊆X⇒C（X→Y）.

PF2 C（X→Y），C（X→Z）∈F ⇒ C（X→YZ）∈F.

PF3 C（X→Y），C（Y→Z）∈F ⇒ C（X→Z）∈F.

PF4 C∈S, P∈PF(C)のとき，C（Id→P）∈F.

PF5 pf∈PF(C1)かつpfの終点はC2とする．更に，C2（X→Y）∈Fが成り立つとき，C1（pf·X→pf·Y）

PFD集合Fに対して，PF1～PF5に関する閉包をF^+で表現する．更に，PFDのカバーリングを定義するために以下の諸概念を定義する．以降，PFDの右辺は単一の経路関数からなるものとする．PF2によりこの仮定は一般性を失わない．

［冗長な経路関数従属性］

C（X→Y）∈［F－C（X→Y）］$^+$なるC（X→Y）を冗長な経路関数従属性とよぶ．

［経路関数従属性の左辺の冗長な経路関数（属性）］

C（X→Y）に対して，pf∈XでX-pf→X'∈F^+が成り立つときpfを冗長な経路関数と呼ぶ．但し，X'は∃pf1, …, pfnに対してPf1

［経路関数従属性集合Fのカバーリング］

経路関数従属性集合Fに対して，各経路関数従属性に対して左辺から冗長な経路関数を削除し，冗長な経路関数従属性を削除したものをFのカバーリングと呼ぶ．

3．クラススキーマの超グラフ表現

Pijは始点のクラス名·終点のクラス名からなるラベルを持ち，pf＝Pij oP'jk o …o P"mnは，Pijの始点のクラス名·P'jkの始点のクラス名·…·P"mnの始点クラス名·P"mnの終点のクラス名からなるラベルを持つものとする．ここでクラススキーマ$\cup_{Ci\in S}$ {Pi1:C1, …Pij:Cj, …, Pin:Cn} およびその上の経路関数従属性集合Fを超グラフH=H(V, E1,HE1, HE2,E2)を用いて表現する．

(i)各pf（Pijをも対象とする）に対し，v∈Vを唯一つ対応づける

(ii)ラベルの先頭クラス名が同一のCであるvからなる集合を超枝he2(C)∈HE2とする．以降これをスキーマ表現超枝とよぶ．

(iii)PFDの左辺が複数の経路関数の集合 {pf1, …, pfn} からなるとき，これに対応するVの元の集合を超枝he1∈HE1とする．

(iv)he1を始点とし，v∈he1を終点とする有向枝およびhe2（∈HE2）を始点とし v∈HE2を終点とする有向枝をE1"の元とし，ラベルを1'とする．

(v)PFDの各元に対し，その左辺に対応する頂点または超枝から右辺に対応する頂点への有向枝を対応付け，E1'の元とする．但し，E1'の元のラベルは1とする．

(vi)E=E1'∪E1"とする．

(vii)∃v∈he2, ∃v'∈he2', vのラベル中最終のクラス名＝v'のラベル中先頭のクラス名のとき，有向枝（he2, he2'）∈E2とし，そのラベルはvのラベルの最終のクラス名とする．

4．超グラフ上変形規則

本節では，PFDの導出律をH上で表現するための変形規則について述べる．規則IはPF3に規則IIはPF2に規則IIIはPF5に規則IVはPF4にほぼ対応する．

規則I：Vx, Vy∈V∪HE1∪HE2, Va∈V∪HE2に対し，（Vx, Vy）∈E1,（Vy, Va）∈E1'の場合には，（Vx, Va）をE1'に加える．（Vx, Va）のラベル付けは以下の通りに行う．

(1) （Vx, Va）∈E1'（Vx, Va）のラベルαのとき，（Vx, Va）（∈E1'）のラベルに1"を加える．但し1"＋1"=1"が成り立つものとする．

(2) （Vx, Va）∈E1'ではないとき，（Vx, Va）をE1'に加え，そのラベルは1"とする．

A Method for Derivation of Covering of Path Functional Dependencies

Mitsuhiro Yamada

NTT Information and Communication Systems Labs

規則II：Vx∈V∪HE1, Vy1…yn∈HE1∪HE2（Vyi∈V）に対し，（Vx, Vyi）∈E1（i=1,…,n）かつ∃iに対して（Vx, Vyi）∈E1'のとき（Vx, Vy1…yn）（ラベル1"とする）をE1'に，（Vx, Vyi）∈E1"（i=1,…,n）のとき（Vx, Vy1…yn）（ラベル1"とする）をE1"に加える．

規則III：he(Ai)∈HE2をラベルの先頭が，Aiである頂点の集合である超枝とする．

① he(Ai), he(Aj)∈HE2, (he(Ai), he(Aj))∈E2, Va1, Va2∈he(Aj),（Va1, Va2）∈E1のとき，he(Ai)内にVa1, Va2を加え（それぞれラベルは（Label(Va1')−Label(he(Ai), he(Aj)))·Label(Va1),（Label(Va2')−Label(he(Ai), he(Aj)))·Label(Va2)とする．但し，Va1'およびVa2'はラベルの最終クラス名がLabel(he(Ai), he(Aj)と一致するhe(Ai)の元），（Va1, Va2）をE1に加える．（ラベルは（Va1, Va2）と同一とする．）

② he(Ai), he(Aj)∈HE2, (he(Ai), he(Aj))∈E2, he1={Va1,…,Van}⊆he(Ai), Van+1∈he(Ai), (he1, Van+1)∈E1のとき，he(Ai)内にVai（i=1,…,n+1）と超枝{Va1,…,Van}を加える．頂点Vaiのラベルは（Label(Vai')−Label(he(Ai), he(Aj)))·Label(Vai)とする．但しVai'はラベルの最終クラス名がLabel(he(Ai), he(Aj))と一致するhe(Ai)の元．有向枝（{Va1,…,Van}, Van+1）にはhe(Aj)における有向枝（{Va1,…,Van}, Van+1）のラベルと同一のラベル付けを行うものとする．

③ he(Ai), he(Aj)∈HE2, (he(Ai), he(Aj))∈E2,（he(Aj), Va）∈E1"のときhe(Ai)に頂点Va'を，有向枝（Vbi, Va'）をE1に加える．但し，Vbiはそのラベルが Ai·Label(he(Ai), he(Aj))なるVの元．Va',（Vbi, Va）のラベルはそれぞれ，Ai·Va（∈he(Aj)）のラベル，1"とする．

規則IV he2(C)に対して，v∈ he2(C), he1⊆he2(C)なるv, he1へ有向枝(he2(C), v)および(he(C), he1)を加え，ラベルは1'とする

5．閉包およびカバーリングの導出手法

本節では，PF集合に非閉路性を仮定する．

5-1閉包の導出手法

H上の規則I〜規則IVに関する閉包$^{\tau}$Hを$\cup_{\gamma\in\Gamma}\gamma(T_I, T_{II}, T_{III}, T_{IV})$ H）（但しγ (x, y, z, w)はx,y,z,wの単項式とし，Tiは規則iによる変形を施す作用素とする．）で定義する．このときX∈(lhs(F))', pf∈(rhs(F))'に対してX→pf∈F^{+}⇔(Vx, Vpf)∈$^{\tau}$H．但し，(lhs(F))'= {lhs(f)|∃pf1, …, pfnに対して，f＝pf1o…o pfn of, f∈F}，(rhs(F))'= {rhs(f)|∃pf1, …, pfnに対して，f=pf1 o…o pfo f, f∈F}．冗長な経路関数従属性および経路関数従属性の左辺上の冗長な経路関数を判別するためにはHの閉包$^{\tau}$Hを求めることが必要となる．以下にH上の閉包を求めるための手法を示す．

[手法1:Hの閉包を求めるための手法]

step0 クラススキーマC(S)に対して，スキーマ表現グラフを作成する．

step1 G=G(HE2, E3)上でHE2の各元に対して，位相的順序を付与する．

step2 HE2の各元he2に対して位相的順序の逆順に従って以下を行う．

step2-1 Hのhe2への制限H|he2=(V∩he2, E1|he2, HE1|he2, he2, ϕ)における規則I, 規則IIに関する閉包$\cup_{\gamma\in\Gamma}\gamma(T_I, T_{II})$(H|he2)を求める．

step2-2 (hei, he2)∈E2なるheiに対して，規則IIIを適用し，heiにそれにより導かれる頂点と有向枝を加える．

step2-3 step2-2で頂点或は有向枝を加えた超枝he2∈HE2に対して，規則IVを適用する．

5-2 カバーリング導出手法

カバーリング導出には，冗長な経路関数従属性および冗長な経路関数の削除が必要であるが，それぞれ超グラフH上では以下の通り特徴づけされる．

経路関数従属性fがF上で冗長⇔f∈［F−{f}］$^{+}$
⇔ e∈$^{\tau}$［H−{e}］但しe=（$V_{lhs(f)}$, $V_{rhs(f)}$）
⇔ e∈［H−{e}]e' 但し［H−{e}]e'はeが含まれるスキーマ表現超枝およびそれより位相的順序が大きいスキーマ表現超枝に対して手法1のstep2を適用したもの．

経路関数従属性fの左辺の経路関数pfが冗長⇔lhs(f)−pf→pf∈F^{+}
⇔e∈$^{\tau}$［H−{e}］但しe=（$V_{lhs(f)}$, V_{pf}）∈E1",
⇔e∈［H−{e}]e' 但し［H−{e}]e'はeが含まれるスキーマ表現超枝およびそれより位相的順序が大きいスキーマ表現超枝に対して手法1のstep2を適用したもの．

以上により，冗長な経路関数従属性および冗長な経路関数の削除を行うためにはそれらに対応するH上の有向枝が含まれるスキーマ表現超枝以上の位相的順序を持つスキーマ表現超枝上に手法1のstep2を適用することが必要である．これにより以下の手法2を得る．

手法2：カバーリングの導出手法

step0 クラススキーマC(S)に対して，スキーマ表現グラフを作成する．

step1 G=G(HE2, E3)上でHE2の各元に対して，位相的順序を付与する．

step2 HE2の各元he2に対して位相的順序の逆順に従って以下を行う．

step2-1 Hのhe2への制限H|he2=(V∩he2, E1|he2, HE1|he2, he2, ϕ)における規則I, 規則IIに関する閉包$^{\tau}$(H|he2)=$\cup_{\gamma\in\Gamma}\gamma(T_I, T_{III})$(H|he2)を求める．

step2-2 (hei, he2)∈E2なるheiに対して，規則IIIを適用し，heiにそれにより導かれる頂点と有向枝を加える．

step2-3 step2-2で頂点或は有向枝を加えた超枝he2∈HE2に対して，規則IVを適用する．

step2-4 e∈$^{\tau}$［H|he2−{e}］なるe（∈E1|he2）を削除する．

step3 step2終了時点のHとstep 0でのHの共通部分を求める

6．むすび

本稿では，経路関数集合が閉路を持たないという仮定のもとでクラススキーマ上に与えられた経路関数従属性集合から，そのカバーリングを求める手法を提案した．

文献

(1)Weddell, G. E.: Reasoning about Functional Dependencies Generalized for Semantic Data Model, ACM TODS Vol 17,1 (1992)

(2)Weddell, G.E.:A Theory of Specialization Constraints for Complex Objects, proc. of 3rd ICDT (1990).

(3)伊藤, 中西:複合オブジェクトモデルにおける型制約の導出について, 信学技報, COMP92-37 (1992).

(4)山田, 中川:関数従属性と包含従属性の相互作用をも対象とした場合の正規形データベーススキーム作成手法, 情処研究会資料, 93-FI-32-2 (1993).