

問 1 データベースの基礎理論に関する次の記述を読んで、設問 1~3 に答えよ。

M 社は、Web 上の SNS、ブログなど（以下、Web サービスという）のアクセスログデータを分析するサービスを提供している。M 社では、契約している Web サービスのサービスプロバイダ（以下、プロバイダという）に登録された利用者を対象として、Web サービスのアクセスログをとり、様々な観点から分析する情報システム（以下、本システムという）を、新たに構築することになった。

本システムは、Web サービスのリソース（SNS、ブログなどのページ）にアクセスした利用者の情報を収集する。具体的には、利用者のプロフィール情報の収集、利用時の位置情報の収集などである。M 社では、これらの機能によって、利用者の行動傾向などを時間的・空間的に分析することを目指している。本システムを構築するに当たって、具体例を用いて検討しながら、関係スキーマを設計することにした。本システムのデータモデルで検討した関係スキーマは、図 1 のとおりである。

図 3~5 は、図 2 の関数従属性の表記法に従って、属性間の関数従属性を表したものである。図 1、図 3~5 の属性とその意味及び制約を、表 1 に示す。

名寄せ（プロバイダ ID, 利用者 ID, 名寄せ ID）
利用者（プロバイダ ID, 利用者 ID, 生年月日, 性別, 郵便番号, 特性 ID, 特性名, 特性タイプ, 特性値）
組織（組織 ID, 組織名称, 郵便番号, 組織登録日付）
所属（プロバイダ ID, 利用者 ID, 組織 ID, 職種, 所属登録日付）
Web サービス（プロバイダ ID, サービス機能 ID, サービス名称, 管理者名）
サービス提供リソース（プロバイダ ID, サービス機能 ID, 利用者 ID, 利用者登録日付, リソース ID, URI, リソース種別, リソース名称, 配信スキーマ）
アクセスログ（プロバイダ ID, 利用者 ID, リソース ID, 配信値, CRUD, タイムスタンプ, 位置情報）

図 1 本システムのデータモデルで検討した関係スキーマ

表1 属性とその意味及び制約

属性名	意味及び制約
プロバイダ ID	プロバイダを一意に識別する ID
利用者 ID	利用者 ID を発行したプロバイダ内で、利用者を一意に識別する ID
名寄せ ID	プロバイダごとに登録されている利用者が同一人物であることを特定して、本システム内で利用者を一意に識別する ID
生年月日	利用者の生年月日
性別	利用者の性別
郵便番号	利用者又は組織の住所の郵便番号
特性 ID	プロバイダごとに個別に定義された利用者の特性（趣味、保有資格など）を、本システム内で一意に識別する ID。利用者ごとに複数の特性が管理される。
特性名	特性 ID に対応する特性の名称
特性タイプ	特性 ID に対応する値のデータ型
特性値	特性 ID に対応する利用者ごとの値
組織 ID	利用者の所属する組織を本システム内で一意に識別する ID。利用者が所属する組織は、同時期には一つだけ登録できる。
組織名称	組織 ID に対応する組織の名称
組織登録日付	組織情報を登録、変更した日付。組織情報（組織名称、郵便番号）は、履歴が管理される。
職種	利用者の所属する組織での職種。例えば、事務、営業など。利用者が所属する組織での職種は、同時期には一つだけ登録できる。
所属登録日付	所属情報を登録、変更した日付。利用者の所属する組織及び職種は、履歴が管理される。
サービス機能 ID	Web サービスの機能を本システム内で一意に識別する ID
サービス名称	プロバイダごとの Web サービスの名称。例えば、“知り合い検索”など。
管理者名	プロバイダごとの Web サービスの管理者の氏名
利用者登録日付	利用者を、プロバイダごと Web サービスごとに本システムに登録した日付
リソース ID	リソースを本システム内で一意に識別する ID。リソースは、それぞれの利用者が所有する SNS、ブログなどのページを指す。複数の Web サービスが同じリソースを利用することがある。
URI	インターネット上のリソースの存在場所を指示する文字列（URI : Uniform Resource Identifier）
リソース種別	リソースの種別。例えば、SNS、ブログなど。
リソース名称	リソースにつけられた名称。例えば、“○○の日記”など。
配信スキーマ	SNS、ブログなどのリソースで管理されるデータの最新情報の概要を配信するためのスキーマ
配信値	配信スキーマの規定に基づいて配信する値
CRUD	配信値に対する操作を識別する記号（C : 生成、R : 参照、U : 更新、D : 削除）
タイムスタンプ	利用者が Web サービスのリソースの配信値を操作した日時
位置情報	Web サービスにアクセスした時点の利用者の位置（GPS で計測した緯度及び経度）。テキスト形式で与える。

凡例			
意味	$A \rightarrow B$	$\{A, B\} \rightarrow C$	$C \rightarrow \{A, B\}$ $C \rightarrow A$ $C \rightarrow B$

図2 関数従属性の表記法

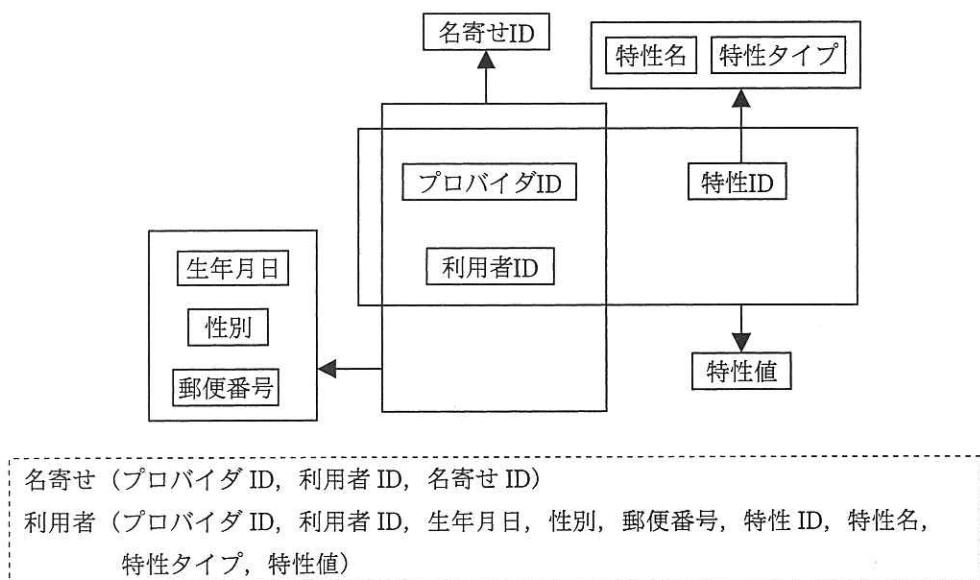


図3 関係“名寄せ”，“利用者”的属性間の関数従属性

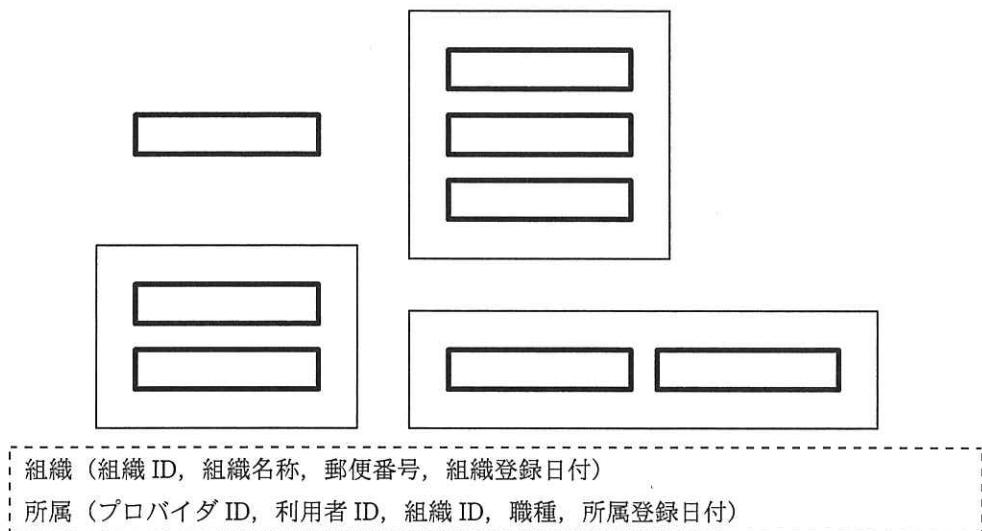


図 4 関係“組織”，“所属” の属性間の関数従属性（未完成）

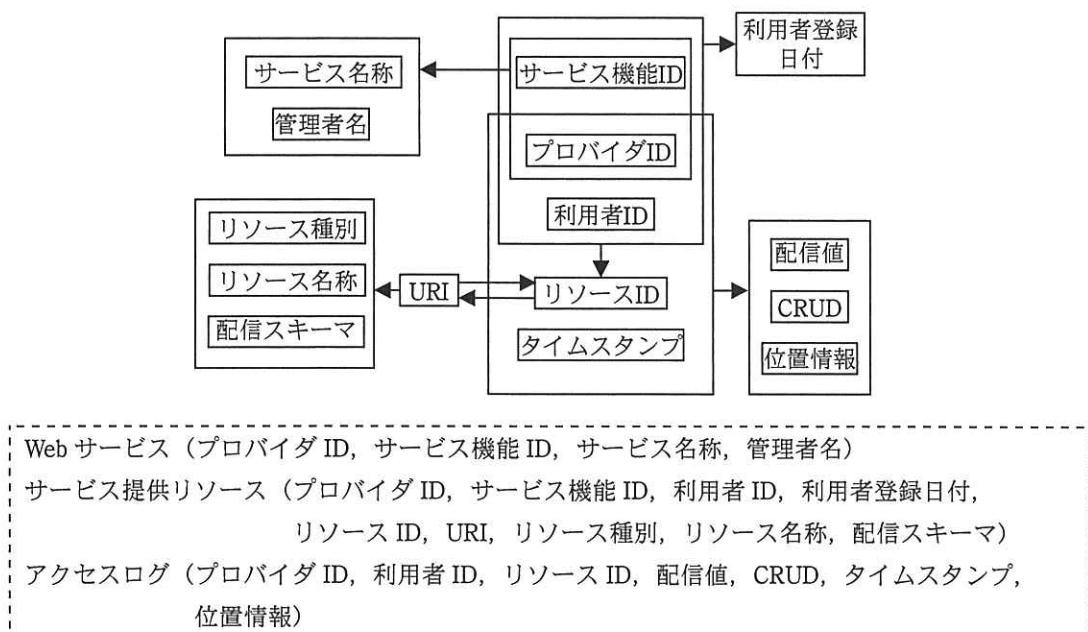


図 5 関係“Web サービス”，“サービス提供リソース”，“アクセスログ” の属性間の  
関数従属性

[利用者の行動傾向分析]

利用者の行動傾向分析を行うために、関係“所属”及び関係“アクセスログ”に対して内自然結合演算及び射影演算を行い、関係“利用実績”を作成した。表2、3は、関係“所属”及び関係“アクセスログ”的具体例である。表4は、表2及び表3に対する演算結果の関係“利用実績”的具体例である。

表2 関係“所属”的具体例

プロバイダID	利用者ID	組織ID	職種	所属登録日付
A	1	#1	事務	2013-01-10
A	1	#1	営業	2013-04-15
A	2	#2	開発	2013-03-10
B	1	#1	事務	2013-02-10
B	2	#2	事務	2013-04-01

表3 関係“アクセスログ”的具体例

プロバイダID	利用者ID	リソースID	配信値	CRUD	タイムスタンプ	位置情報
A	1	S1	V1	C	2013-04-10 12:00:00	P1
A	2	S1	V2	R	2013-04-03 14:00:00	P2
B	1	S2	V3	U	2013-04-17 16:00:00	P3
B	1	S3	V4	D	2013-04-18 18:00:00	P4

注記 属性の値は、一部簡略化のため記号で表している。

表4 関係“利用実績”的具体例

	プロバイダID	利用者ID	組織ID	職種	配信値	タイムスタンプ	位置情報
①	A	1	#1	事務	V1	2013-04-10 12:00:00	P1
②	A	1	#1	営業	V1	2013-04-10 12:00:00	P1
③	A	2	#2	開発	V2	2013-04-03 14:00:00	P2
④	B	1	#1	事務	V3	2013-04-17 16:00:00	P3
⑤	B	1	#1	事務	V4	2013-04-18 18:00:00	P4

注記 ①～⑤は、各タプルを識別するために付与した番号である。

**設問 1** 関係“名寄せ”，“利用者”及び図3，4について，(1)～(3)に答えよ。

- (1) 図3の関係“名寄せ”的候補キーを全て答えよ。
- (2) 図3の関係“利用者”は，タプル挿入に関してどのような問題があるか。その内容を，35字以内で具体的に述べよ。また，第3正規形に分解した関係スキーマを示せ。

なお，分解した関係スキーマの関係名は任意とし，主キーは実線の下線で示すこと。

- (3) 図4の関数従属性を，□には属性名を記入し，図2中の凡例の欄に示した表記法に従って完成させよ。

**設問 2** 関係“サービス提供リソース”及び図5について，(1)，(2)に答えよ。

- (1) 図5の関係“サービス提供リソース”的候補キーを全て答えよ。また，部分関数従属性，推移的関数従属性の有無を，“あり”又は“なし”で答えよ。“あり”的場合は，その関数従属性の具体例を，図2中の意味の欄に示した表記法に従って示せ。
- (2) 図5の関係“サービス提供リソース”は，第1正規形，第2正規形，第3正規形のうち，どこまで正規化されているかを答えよ。また，第3正規形でない場合は，第3正規形に分解した関係スキーマを示せ。

なお，分解した関係スキーマの関係名は任意とし，主キーは実線の下線で示すこと。

**設問 3** 表2～4について，(1)～(3)に答えよ。

- (1) 表4の関係“利用実績”には，アクセスした時点の利用者の所属を前提にした場合に現れてはならないタプルがある。そのタプルの番号を全て答えよ。
- (2) 表4の関係“利用実績”的タプルに現れない利用者がある。どのような条件に当てはまる利用者か。25字内で述べよ。
- (3) (2)の現れない利用者も表4の関係“利用実績”的タプルに現れるようにするためには，内自然結合演算をどのような演算に変更すればよいか。30字内で述べよ。