

【2】

氏名(本籍)	李 福 瑞 (中 国)
学位の種類	学 術 博 士
学位記番号	博 甲 第 586 号
学位授与年月日	昭 和 63 年 5 月 31 日
学位授与の要件	学 位 規 則 第 5 条 第 1 項 該 当
審 査 研 究 科	社 会 工 学 研 究 科
学 位 論 文 題 目	A Statistical Database Model : Its Uniqueness and the Design Procedure
主 査	筑波大学教授 高 橋 磐 部
副 査	筑波大学教授 経済学博士 穂 鷹 良 介
副 査	筑波大学教授 理学博士 藤 原 讓
副 査	筑波大学教授 工学博士 藤 重 悟
副 査	筑波大学助教授 工学博士 山 本 芳 嗣

論 文 の 要 旨

現実の社会統計表は様々な機関によって、別々に作られているため、これらを整理統合して標準化された、情報の重複や欠落のない、統計表を作ることが極めて重要である。この論文はそのような標準化された統計表を作る為の一定の方式を示したものである。

現実の統計表は、たとえば、性別、年齢別の収入とか、地域別、産業クラス別の従業員数とか、いくつかの属性別に統計量が与えられている。しかし統計量として男性従業員数とか、管理者数といったものが含まれていることがある。これらは属性と統計量とが複合した概念であって、こうした混乱を、この論文では、category 属性と summary 属性という2種の属性に明確に分離することによって防ぐことを提案している。たとえば男性従業員数というのは、一つの統計量ではなく、男性という category 属性と従業員数という summary 属性の結合と考えるべきだと主張する。この立場に立ってまず summary 属性のもつ purity という概念を論理的に明確に定義している。(§3・2)。

また category 属性として男性管理者、女性管理者、男性技術者、女性技術者といった分類が用いられることがある。これは性別と職種という2つの category 属性の組合わせとして表現できるもので、これに関連して2つの以上の category の組合わせとして表現できない、primitive な category 属性という概念を論理的に明確にした (§3・3)。

summary 属性の purity と category 属性の primitiveness がこの論文の基調になっているが、これによって今まであいまいであった概念が整理され統計表の標準化が簡単化される。またこの2つの

概念に関連するいくつかの命題が第3章 Statistical Database Model に述べられている。これらは大まかに言って、より広い場で pure (あるいは primitive) ならより狭い場でも pure (あるいは primitive) であるという特徴をもっている点でも、基本的概念であると言える。

以上の概念や命題を形式論理的に表現するために、第2章 Fundamental Data Model において、entity, entity type, 属性, level 等の基本概念が説明されている。

第4章 Design Procedure of Statistical Databases においては、与えられたいくつかの統計表を標準化された統計表に変換する手続きが示されている。その方法はまず、与えられた統計表をすべて summary 属性が1つになるように分解し、各 summary 属性が、カテゴリ属性を新たに加えるなどして、pure となるように操作する。これが purification という操作である。つぎに category 属性の中に primitive でないものがあれば、これを primitive ないくつかの category 属性の組合わせで表現する。これを orthogonal decomposition と呼んでいる。

つぎにこうして標準化された統計表の中に category 属性と summary 属性は同じだが、category 属性の分割の違うものがあれば、これらに共通の分割を導入して、2つの表を1つにまとめる。これが union という操作である。最後に category 属性が同一であるいくつかの表を一つにまとめ、その summary 属性を列挙するという操作を行う。これを join と呼んでいる。

こうしてできた統計表の集りは summary 属性がすべて pure であり、category 属性がすべて primitive であるような標準化された統計表となる。そしてまたある条件があれば、この操作によってえられる統計表は unique であることが示されている。

第5章 Problems on Statistical Database and Conclusions ではこれを現実に適用する際の様々な問題点を指摘し、附録において、ある産業統計表のモデル例を以上の手続きで標準化する例が挙げられている。

審 査 の 要 旨

一般に統計データベースの標準化はやりにくいという通念があるが、この論文では category 属性と summary 属性という概念を基調として、論理的にきちんとまとめられている、という点で、データベース関連の学会においても十分評価されうると思われる内容をもっている。

統計表の一意性が言えたと言っても、それはある前提があつての話であり、たとえば2次元の点を表現するのに直交座標系と極座標系のいずれをとるかという前提まで一意にできるという話ではなく、これを実際に適用する段となると、一種の ISO 規格のような標準化が行きわたらないと難しいと思われる。

また論理構造だけを考えると、一つ一つ概念や命題は自明であつて、何ら新しい発見がないような印象を受ける。しかしこれは統計データベースという現実の世界と論理モデルとの橋渡しとして価値あるものとみるべきであろう。

よって、著書は学術博士の学位を受けるに十分な資格を有するものと認める。