

【公報種別】特許法第 17 条の 2 の規定による補正の掲載
 【部門区分】第 6 部門第 3 区分
 【発行日】平成20年2月14日 (2008.2.14)

【公開番号】特開2006-4411 (P2006-4411A)
 【公開日】平成18年1月5日 (2006.1.5)
 【年通号数】公開・登録公報2006-001
 【出願番号】特願2005-139882 (P2005-139882)
 【国際特許分類】

G 0 6 F 17/30 (2006.01)

G 0 6 F 12/00 (2006.01)

【 F I 】

G 0 6 F 17/30 2 3 0 Z

G 0 6 F 17/30 3 5 0 C

G 0 6 F 12/00 5 4 6 T

【手続補正書】
 【提出日】平成19年12月21日 (2007.12.21)
 【手続補正 1】
 【補正対象書類名】特許請求の範囲
 【補正対象項目名】全文
 【補正方法】変更
 【補正の内容】
 【特許請求の範囲】
 【請求項 1】

コンピュータシステムにおいて生起キューブを生成する方法において、
データベースのドメインのそれぞれのグローバル属性に関して、前記データベースのそれぞれのインタフェース属性に対して、前記データベースにクエリをサブミットするステップであって、各クエリは、前記データベースのインタフェース属性の値が前記データベースのドメインのグローバル属性のグローバル属性値にセットされる、クエリをサブミットするステップと、

それぞれのサブミットされたクエリの結果に対して、前記グローバル属性の前記値が前記結果の各結果属性に現れる回数をカウントするステップと、

グローバル属性、インタフェース属性、および結果属性の各組合せに対して、前記グローバル属性のグローバル属性値にセットされた前記インタフェース属性を有してサブミットされるクエリからの結果としての前記グローバル属性の前記値が各結果属性に現れる回数をカウントする蓄積を、前記生起キューブの要素として、記憶するステップと
を含み、

前記記憶された要素は、生起キューブを形成することを特徴とする生起キューブを生成する方法。

【請求項 2】

前記生起キューブから、グローバル属性およびインタフェース属性に関連する生起行列を生成するステップを含むことを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記生起キューブから、グローバル属性および結果属性に関連する生起行列を生成するステップを含むことを特徴とする請求項 1 に記載の方法。

【請求項 4】

前記生起キューブから、インタフェース属性および結果属性に関連する生起行列を生成するステップを含むことを特徴とする請求項 1 に記載の方法。

【請求項 5】

グローバル属性値およびインタフェース属性の各組合せごとにクエリがサブミットされることを特徴とする請求項 1 に記載の方法。

【請求項 6】

前記生起キューブが、グローバル属性、インタフェース属性、および結果属性の各組合せに対するカウントを含むことを特徴とする請求項 1 に記載の方法。

【請求項 7】

コンピュータシステムにおいてドメイン内のデータベースの属性を識別する方法において、

前記ドメインのグローバルスキーマのグローバル属性ならびに前記データベースのインタフェーススキーマのインタフェース属性および結果スキーマの結果属性に関連する生起カウントを提供するステップであって、カウントのそれぞれは、グローバル属性、インタフェース属性、および結果属性の各組合せに対して、グローバル属性に関するグローバル属性値が、前記グローバル属性値にセットされた前記インタフェース属性を有して前記データベースにサブミットされるクエリの結果の前記結果属性の値として現れる、生起の数を表す、生起カウントを提供するステップと、

提供されたカウントに基づいてスキーマの対の相互情報量を推定するステップと、
推定された相互情報量からどの属性がマッチするかを識別するステップと
を含むことを特徴とするデータベースの属性を識別する方法。

【請求項 8】

前記カウントを提供するステップが、グローバル属性、インタフェース属性、および結果属性に関連する生起カウントを提供する生起キューブを、スキーマの対に関連する行列に射影するステップを含むことを特徴とする請求項 7 に記載の方法。

【請求項 9】

前記インタフェース属性の値が前記グローバル属性のグローバル属性値にセットされたクエリを前記データベースにサブミットすることによって前記生起キューブを生成するステップを含むことを特徴とする請求項 8 に記載の方法。

【請求項 10】

前記生起キューブ内の前記生起カウントは、クエリにおけるインタフェース属性の値として用いられるグローバル属性のグローバル属性値が該クエリの結果の結果属性に現れる回数を表すことを特徴とする請求項 9 に記載の方法。

【請求項 11】

前記インタフェース属性が、HTML の入力関連要素に基づいて識別されることを特徴とする請求項 7 に記載の方法。

【請求項 12】

前記結果属性が、正規表現ラップを用いて識別されることを特徴とする請求項 7 に記載の方法。

【請求項 13】

インタフェース属性の値が前記グローバル属性のグローバル属性値にセットされたクエリを前記データベースにサブミットすることによって、前記生起カウントが提供されることを特徴とする請求項 7 に記載の方法。

【請求項 14】

前記相互情報量が次式

【数 1】

$$EMI(S_{1i}, S_{2j}) = \frac{m_{ij}}{M} \log \frac{\frac{m_{ij}}{M}}{\frac{m_{i+}}{M} * \frac{m_{+j}}{M}}$$

によって推定されることを特徴とする請求項 7 に記載の方法。

【請求項 15】

スキーマの対において、一方のスキーマの属性であって、他方のスキーマの属性について最高の推定相互情報量を有する属性が、他方のスキーマの別の属性についてさらに高い推定相互情報量を有しない場合に、スキーマの対における属性間のマッチが識別されることを特徴とする請求項 7 に記載の方法。