

【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 23 年 9 月 22 日 (2011.9.22)

【公開番号】特開 2010-86332 (P2010-86332A)
 【公開日】平成 22 年 4 月 15 日 (2010.4.15)
 【年通号数】公開・登録公報 2010-015
 【出願番号】特願 2008-255501 (P2008-255501)
 【国際特許分類】

G 0 6 F 17/30 (2006.01)

G 0 6 F 3/048 (2006.01)

【F I】

G 0 6 F 17/30 3 3 0 Z

G 0 6 F 3/048 6 5 4 B

G 0 6 F 17/30 3 4 0 Z

【手続補正書】
 【提出日】平成 23 年 8 月 4 日 (2011.8.4)
 【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】

検索対象のデータ集合を検索する検索装置における検索方法であって、
 複数の検索式を組み合わせる検索式セットを生成する検索式セット生成工程と、
 前記検索式セット生成工程で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率又はヒット数に基づいて検索式セットの提示順位スコアを算出する検索式セットスコア算出工程と、
 前記検索式セット生成工程で生成された検索式セットを前記検索式セットスコア算出工程で算出された提示順位スコア順に提示する検索式セット提示工程と、
前記検索式セット提示工程で提示された検索式セットからユーザに選択された検索式に基づいて前記データ集合に対して検索を実行する検索式実行工程と、
 を含むことを特徴とする検索方法。

【請求項 2】

前記検索式セット提示工程では、前記検索式セット生成工程で生成された検索式セットとともに、検索式セットに含まれる各検索式に対応する前記ヒット率又はヒット数を前記検索式セットスコア算出工程で算出された提示順位スコア順に提示することを特徴とする請求項 1 に記載の検索方法。

【請求項 3】

前記検索式セットスコア算出工程では、前記検索式セット生成工程で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率又はヒット数に基づいて、前記ヒット率又はヒット数が均等に近い検索式セットが優先的に提示されるよう提示順位スコアを算出することを特徴とする請求項 1 又は 2 に記載の検索方法。

【請求項 4】

前記検索式セットスコア算出工程では、前記検索式セット生成工程で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率に基づいて、前記データ集合内での重複度が少ない検索式セットが優先的に提示されるよう提示順位スコアを算出することを特徴とする請求項 1 又は 2 に記載の検索方法。

【請求項 5】

検索対象のデータ集合を検索する検索装置であって、

複数の検索式を組み合わせる検索式セットを生成する検索式セット生成手段と、

前記検索式セット生成手段で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率又はヒット数に基づいて検索式セットの提示順位スコアを算出する検索式セットスコア算出手段と、

前記検索式セット生成手段で生成された検索式セットを前記検索式セットスコア算出手段で算出された提示順位スコア順に提示する検索式セット提示手段と、

前記検索式セット提示手段で提示された検索式セットからユーザに選択された検索式に基づいて前記データ集合に対して検索を実行する検索式実行手段と、

を含むことを特徴とする検索装置。

【請求項 6】

前記検索式セット提示手段は、前記検索式セット生成手段で生成された検索式セットとともに、検索式セットに含まれる各検索式に対応する前記ヒット率又はヒット数を前記検索式セットスコア算出手段で算出された提示順位スコア順に提示することを特徴とする請求項 5 に記載の検索装置。

【請求項 7】

前記検索式セットスコア算出手段は、前記検索式セット生成手段で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率又はヒット数に基づいて、前記ヒット率が均等に近い検索式セットが優先的に提示されるよう提示順位スコアを算出することを特徴とする請求項 5 又は 6 に記載の検索装置。

【請求項 8】

前記検索式セットスコア算出手段は、前記検索式セット生成手段で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率又はヒット数に基づいて、前記データ集合内での重複度が少ない検索式セットが優先的に提示されるよう提示順位スコアを算出することを特徴とする請求項 5 又は 6 に記載の検索装置。

【請求項 9】

検索対象のデータ集合を検索するコンピュータを、

複数の検索式を組み合わせる検索式セットを生成する検索式セット生成手段と、

前記検索式セット生成手段で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率又はヒット数に基づいて検索式セットの提示順位スコアを算出する検索式セットスコア算出手段と、

前記検索式セット生成手段で生成された検索式セットを前記検索式セットスコア算出手段で算出された提示順位スコア順に提示する検索式セット提示手段と、

前記検索式セット提示手段で提示された検索式セットからユーザに選択された検索式に基づいて前記データ集合に対して検索を実行する検索式実行手段と、

して機能させることを特徴とするプログラム。

【請求項 10】

請求項 9 に記載のプログラムを記憶したコンピュータにより読み取り可能な記憶媒体。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0007

【補正方法】変更

【補正の内容】

【0007】

そこで、本発明は、検索対象のデータ集合を検索する検索装置における検索方法であって、複数の検索式を組み合わせる検索式セットを生成する検索式セット生成工程と、前記検索式セット生成工程で生成された前記検索式セットに含まれる各検索式の前記データ集合内でのヒット率又はヒット数に基づいて検索式セットの提示順位スコアを算出する検索式セットスコア算出工程と、前記検索式セット生成工程で生成された検索式セットを前記

検索式セットスコア算出工程で算出された提示順位スコア順に提示する検索式セット提示工程と、前記検索式セット提示工程で提示された検索式セットからユーザに選択された検索式に基づいて前記データ集合に対して検索を実行する検索式実行工程と、を含むことを特徴とする。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0022

【補正方法】変更

【補正の内容】

【0022】

図9は、ステップS8-4における検索式セット候補設定の処理詳細を示したフローチャートである。ステップS9-1において、CPUは、必要な変数を初期設定する。ステップS9-2において、CPUは、指定文書集合において指定属性の値として出現するものを候補としてリストアップする。ステップS9-3において、CPUは、候補のうちから値を一つ取り出す。ステップS9-4において、CPUは、指定文書集合内で「属性＝値」となる件数を求め、文書集合内ヒット率を算出する。ステップS9-5において、CPUは、求めた値と文書集合内ヒット率とを検索式セット候補テーブル内の該当検索式の然るべきフィールドに設定する。ステップS9-6において、CPUは、全ての値候補を処理し終えたかを確認し、未処理の値があるときはステップS9-7において次の値を選択し、ステップS9-3にループする。ステップS9-6において、全ての値の処理が終了しているときは、CPUは、ステップS9-8において、この検索式セットの提示順位スコアを図6に示した算出式に基づいて算出（検索式セットスコア算出）し、検索式セット候補テーブルに格納する。その後リターンする。ここで、図6通りの算出式で提示順位スコアを設定すればヒット率が均等に近い検索式セットが優先的に提示されることになる。或いは、データ集合内（文書集合内）での重複度が少ない検索式セットを優先的に提示したければ、提示順位スコアとして各文書集合内ヒット率の平均を採用することもできる。ヒット率の平均は、各検索式候補でヒットするデータの重複が全く無い場合に（全体集合数）／（検索式候補数）に一致するが、データの重複がある場合には、その分だけ値が大きくなるので、データ重複がない検索式セットの提示順位スコアが小さく算出されることになる。更に、ヒット率の均等化とデータの重複度との両方を加味して優先するには、提示順位スコアとして、図6の算出式とヒット率の平均を、それぞれの重複具合に応じて適切な重みをつけて加算した値を採用することにより行なうことができる。