

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-531808

(P2004-531808A)

(43) 公表日 平成16年10月14日(2004.10.14)

(51) Int.Cl.⁷

G06F 17/30

G06F 12/00

F I

G06F 17/30

G06F 12/00

310A

513D

テーマコード (参考)

5B075

5B082

審査請求 未請求 予備審査請求 未請求 (全 53 頁)

(21) 出願番号 特願2002-578179 (P2002-578179)
 (86) (22) 出願日 平成14年3月12日 (2002.3.12)
 (85) 翻訳文提出日 平成15年9月26日 (2003.9.26)
 (86) 国際出願番号 PCT/GB2002/001107
 (87) 国際公開番号 W02002/080025
 (87) 国際公開日 平成14年10月10日 (2002.10.10)
 (31) 優先権主張番号 01302892.3
 (32) 優先日 平成13年3月28日 (2001.3.28)
 (33) 優先権主張国 欧州特許庁 (EP)

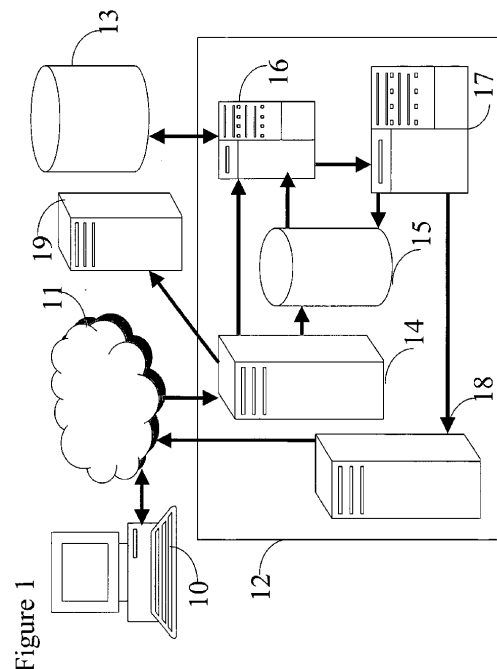
(71) 出願人 390028587
 ブリティッシュ・テレコミュニケーションズ・パブリック・リミテッド・カンパニー
 BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY
 イギリス国、イーシー1エー・7エー ジェイ、ロンドン、ニューゲート・ストリート 81
 (74) 代理人 100058479
 弁理士 鈴江 武彦
 (74) 代理人 100091351
 弁理士 河野 哲
 (74) 代理人 100088683
 弁理士 中村 誠

最終頁に続く

(54) 【発明の名称】 データ検索システム

(57) 【要約】

オンラインのデータベースサーチングシステムでは、プロダクト選択プロセスを使用して、進化サーチ方式を実行し、このとき、ユーザは選択圧力 (selective pressure) の役割を果たし、最近の選択に基づく "変化 (mutation)" を使用する。特徴の新しい組を定めるために選択される項目の特徴を変えることによって、選択を変化させ、新しく選択された特徴の組の別の項目を、ユーザによる検討のために選択する。特徴の多次元の範囲において選択することによって、このようなプロセスから、'サーチ空間' の予想外の探査、すなわち既存のカタログベースシステムの一般的な分岐形サーチ方式ではなく、小売店で使用されているブラウジングプロセスにより類似した探査を生成することができる。



【特許請求の範囲】**【請求項 1】**

データベースから、表示する項目を選択するための装置であって、各項目ごとに、その項目とデータベース内の他の項目との類似度を示すデータを記憶するためのデータ記憶手段と、
前記データベース内の第 1 の項目を識別するユーザ入力を受信するための入力手段と、
第 1 の項目に対する所定の類似度を有する進化基準を特定するための進化プロセッサ手段と、
データベースから、進化基準を満たす異なる項目を識別することと、
異なる項目の中から第 2 の項目を選択するための選択手段と、
選択された第 2 の項目を識別する出力を表示するための出力手段とが構成されている装置。

10

【請求項 2】

複数の項目を表示するための表示手段が構成されていて、入力手段が、表示されている項目の中の、第 1 の項目になる 1 つを選択するための手段を有する請求項 1 記載の装置。

【請求項 3】

出力手段が、表示手段を制御して、複数の項目の中の 1 つを、選択された第 2 の項目へ置き換える手段を有する請求項 2 記載の装置。

【請求項 4】

選択されずに最も長い間表示されていた項目を、選択された第 2 の項目へ置き換える請求項 3 記載の装置。

20

【請求項 5】

選択された項目の表示を生成するための手段、最初はゼロに設定された経過期間値を、各表示された項目に割り当てるための手段、各表示された項目の経過期間値を定期的にインクリメントするための手段、この項目を識別する入力手段が入力を受信したことに応答して、所与の表示された項目の経過期間値をゼロへリセットするための手段と、所定値よりも大きい経過期間値を有する表示項目から削除するための手段とがさらに構成されている請求項 4 記載の装置。

【請求項 6】

データ記憶手段には、1 組の属性の各々に各項目の特定値を割り当てるための手段が構成されていて、2 つの項目間の類似度が、2 つの項目が共通の値を有する属性数によって識別される請求項 1 ないし 5 の何れか 1 項記載の装置。

30

【請求項 7】

第 1 の項目の属性値の組から所定度分異なる 1 組の属性値を生成するための属性進化生成手段と、生成された組に対応する属性値を有する第 2 の項目をデータベースから選択するための選択手段とがさらに構成されている請求項 6 記載の装置。

【請求項 8】

データ記憶手段が、各項目の特定値を、各他の項目との類似度を定める 1 組の属性の各々に割り当てるための手段がさらに構成されている請求項 1 ないし 5 の何れか 1 項記載の装置。

40

【請求項 9】

進化基準手段には、第 1 の項目に対する所定の類似度を特定する進化基準を定めるための手段が構成されている請求項 8 記載の装置。

【請求項 10】

各特定値が、2 つの項目間の関係付けが存在するか、しないかを定める請求項 8 または 9 記載の装置。

【請求項 11】

属性値の生成された値が、2 つ以上の前の入力の属性にしたがって判断される請求項 7 または 8 記載の装置。

【請求項 12】

50

装置の動作サイクルにおいて行われた選択を記録するための記録手段がさらに構成されていて、属性生成手段が、所定回数前からの装置の動作サイクルの中で選択された項目であって、記録された項目の属性に関係する 1 組の属性を生成するように配置されている請求項 11 記載の装置。

【請求項 13】

各属性と重み付け値とを関係付けるための手段と、第 1 の項目に関係する入力を受信したときに第 1 の項目と関係付けられた属性の重み付け値を増加するための手段とが構成されていて、より大きい重みを割当てられた属性を有する項目が、より小さく重み付けされた属性よりも選択の可能性がより大きくなるように、第 2 の項目を検索するための手段が動作する請求項 6 ないし 11 のいずれか 1 項記載の装置。

10

【請求項 14】

装置の各動作サイクルにおいて行われた選択を記録するための記録手段が構成されていて、選択手段が、第 2 の項目の選択を制約して、所定回数前からの装置の動作サイクルの中で既に選択された項目であって、記録手段によって記録された項目の選択を防ぐように配置されている請求項 1 ないし 13 のいずれか 1 項記載の装置。

【請求項 15】

データベースから、表示するための項目を選択する方法であって、各項目とデータベース内の他の項目との類似度を示すデータを生成するステップと、前記データベース内の第 1 の項目を識別する入力を受信するステップと、進化基準を生成し、第 1 の項目に対する所定の類似度を識別するステップと、データベース内の、進化基準を満たす項目を選択するステップと、選択された第 2 の項目を表示するステップとが含まれる方法。

20

【請求項 16】

複数の項目が表示され、第 1 の項目が、表示される項目の中の 1 つである請求項 15 記載の方法。

【請求項 17】

反復プロセスであって、選択された第 2 の項目を表示部へ加え、表示された項目の中の 1 つに関係する別の入力を受信して、表示の別の項目を選択する請求項 16 記載の方法。

【請求項 18】

複数の表示された項目の中の 1 つを、選択された第 2 の項目へ置き換える請求項 16 または 17 記載の方法。

30

【請求項 19】

各表示された項目が、経過期間値を割り当てられ、経過期間値は、最初はゼロに設定され、定期的にインクリメントされ、所与の表示された項目が選択されたときに、その項目の経過期間値がゼロにリセットされ、経過期間値を使用して、表示部から削除する項目を識別する請求項 18 記載の方法。

【請求項 20】

選択されずに最も長い間表示されていた項目を、選択された第 2 の項目へ置き換える請求項 19 記載の方法。

【請求項 21】

第 2 の項目の選択を制約して、プロセスを所定回数、反復する中で同じ項目の選択を防ぐ請求項 18 ないし 20 のいずれか 1 項記載の方法。

40

【請求項 22】

各項目が、特定値を、1 組の属性の各々に割り当てて、項目間の類似度を、これらが共通の値を有する属性数によって識別される請求項 15 ないし 21 のいずれか 1 項記載の方法。

【請求項 23】

進化基準が、第 1 の項目の属性値の組から所定度分異なっている 1 組の属性値を生成することによって選択され、第 2 の項目が、データベースから、属性値の生成された組に対応する属性値を有する項目から選択される請求項 22 記載の方法。

50

【請求項 2 4】

進化基準が、2つ以上の前の入力 of 属性にしたがって判断される請求項 2 3 記載の方法。

【請求項 2 5】

各項目において、属性の組が定められ、属性の組の項目が、その項目と各他の項目との類似度を表わしている請求項 1 5 ないし 2 1 のいずれか 1 項記載の方法。

【請求項 2 6】

進化基準が、第 1 の項目に関係する各属性の組の項目の所定の値を特定することによって生成される請求項 2 5 記載の方法。

【請求項 2 7】

各特定値が、2つの項目間の関係付けがあるか、ないかを定める請求項 2 5 または 2 6 記載の方法。 10

【請求項 2 8】

各属性が、重み付け値と関係付けられ、第 1 の項目に関係する入力を受信すると、第 1 の項目と関係付けられている属性の重み付け値が増加し、より大きい重みを割り当てられた属性を有する項目が、より小さく重み付けされた属性を有する項目よりも、選択される可能性がより大きいといった進化基準が生成される請求項 2 1 ないし 2 5 のいずれか 1 項記載の方法。

【請求項 2 9】

請求項 1 5 ないし 2 8 の何れか 1 項記載のステップを実行するためのコンピュータプログラム。 20

【請求項 3 0】

コンピュータの内部メモリへ直接にロード可能なコンピュータプログラムプロダクトであって、前記プロダクトがコンピュータ上で実行されるときに、請求項 1 5 ないし 2 8 の何れか 1 項記載のステップを実行するためのソフトウェアコード部分が構成されているコンピュータプログラムプロダクト。

【請求項 3 1】

コンピュータの使用可能な媒体上に記憶されているコンピュータプログラムプロダクトであって、

コンピュータに、各項目と、データベース内の他の項目との類似度を示すデータを生成させるコンピュータ読み出し可能プログラム手段、 30

コンピュータに、前記データベース内の第 1 の項目を識別する入力を受信させるコンピュータ読み出し可能プログラム手段、

第 1 の項目に対する所定の類似度を特定する進化基準を生成するためのコンピュータ読み出し可能プログラム手段、

コンピュータに、データベース内の、進化基準を満たす項目を選択させるコンピュータ読み出し可能プログラム手段、

コンピュータに、選択された第 2 の項目を表示させるコンピュータ読み出し可能プログラム手段が構成されているコンピュータプログラムプロダクト。

【発明の詳細な説明】**【技術分野】**

40

【0001】

本発明は、データ検索システム、とくに、ユーザが広範な有効項目から選択するのに援助するシステムに関する。このシステムは、検討すべき多数の変数があるサーチ可能なデータベースに適用され、ユーザは、ユーザ自身の優先基準にしたがってサーチする自由を必要とするが、完全に構造化されたサーチが可能になるには、データベースが大き過ぎるか、またはユーザの基準に対する規定が不十分過ぎる。

【背景技術】**【0002】**

サーチャーは、一般に、サーチ可能なデータベースにおいて、分岐決定“ツリー”に沿って、サーチャーが希望すると期待される宛先へ向かってナビゲートされる。これは、既知 50

の目標へ向かってサーチするための良好な方法である。しかしながら、異なる宛先に到達するには、経路を後退しなければならないので、目標の定義が比較的に不明確であるか、または幾つかの目標を調べる必要があるときは、このようなシステムは、比較的に不十分に構造化されたサーチング（“ブラウジング”）には、あまり適していない。サーチャーは、データベースのカテゴリ分類に完全に依存し、発見し難いか、または入手可能なものの全体的な印象を形成し、自分が選んだものを注文する（例えば、衣服を購入するときの通常の手順と同じである）。

【発明の開示】

【発明が解決しようとする課題】

【0003】

オンラインサーチングの欠点は、ユーザとデータベースとの間のリンクの帯域幅が狭いときは、さらに一層大きくなる。モデムを介してオンラインデータベースを‘ブラウズ’する試みは、例えば、典型的には、ホームページにロードする一方で休止し、サーチャーはデータベース内の1セクションを比較的に迅速に選択し、そのセクションのページにロードする一方で再び休止し、項目のカテゴリを迅速に選択し、再び休止し、など、などを含む。インターネットへのモバイルアクセスは、人々が固定リンクに広帯域幅の接続を取入れても、比較的狭帯域幅のオンラインサーチングが増加しつづける可能性が高くなることを意味する。

【課題を解決するための手段】

【0004】

本発明にしたがって、データベースから、表示する項目を選択する装置であって、各項目ごとに、その項目とデータベース内の他の項目との類似度を示すデータを記憶するためのデータ記憶手段と、

前記データベース内の第1の項目を識別するユーザ入力を受信するための入力手段と、

第1の項目に対する所定の類似度を有する進化基準を特定するための進化プロセッサ手段と、

データベースから、進化基準を満たす異なる項目を識別することと、

異なる項目の中から第2の項目を選択するための選択手段と、

選択された第2の項目を識別する出力を表示するための出力手段とが構成されている装置を提供する。

【0005】

本発明は、データベースから、表示する項目を選択する方法であって、

各項目とデータベース内の他の項目との類似度を示すデータを生成するステップと、

前記データベース内の第1の項目を識別する入力を受信するステップと、

進化基準を生成し、第1の項目に対する所定の類似度を識別するステップと、

データベース内の、進化基準を満たす項目を選択するステップと、

選択された第2の項目を表示するステップとが含まれる方法にも拡張される。

【0006】

本発明は、本発明の方法を実行するためのコンピュータプログラム、およびコンピュータの内部メモリ内へ直接にロード可能なコンピュータプログラムプロダクトであって、コンピュータ上でプロダクトが実行されているときに方法のステップを実行するためのソフトウェアコード部分が構成されているコンピュータプログラムプロダクトへも拡張される。

【0007】

本発明は、コンピュータの使用可能媒体上に記憶されているコンピュータプログラムプロダクトであって、

コンピュータに、各項目と、データベース内の他の項目との類似度を示すデータを生成させるコンピュータ読み出し可能プログラム手段、

コンピュータに、前記データベース内の第1の項目を識別する入力を受信させるコンピュータ読み出し可能プログラム手段、

第1の項目に対する所定の類似度を特定する進化基準を生成するコンピュータ読み出し可

10

20

30

40

50

能プログラム手段、

コンピュータに、データベース内の、進化基準を満たす項目を選択させるコンピュータ読み出し可能プログラム手段、

コンピュータに、選択された第2の項目を表示させるコンピュータ読み出し可能プログラム手段が構成されているコンピュータプログラムプロダクトへも拡張される。

【0008】

複数の項目が表示され、入力手段が、表示された項目の中の、第1の項目となる1つを選択する手段を有することが好ましい。好ましい構成では、出力手段は、表示手段を制御して、複数の項目の中の1つ、好ましくは、選択されずに最も長い間表示された項目を、選択された第2の項目によって置き換えるための手段を有する。これは、反復プロセスであって、選択された第2の項目を表示部へ加え、表示されている項目の中の1つに関する別の入力を受信し、表示の別の項目を選択する。何れの項目が置き換えられるかを識別するために、各表示されている項目に、経過期間値 (age value) を割り当てる。経過期間値は、最初はゼロに設定され、定期的にインクリメントされる。所与の表示されている項目の経過期間値は、その項目が選択されるとき、ゼロにリセットされる。経過期間値は、表示部から削除する項目を識別するのに使用される。第2の項目の選択に制約を加えて、プロセスを所定回数、反復する中で、同じ項目を選択するのを防ぐ。

10

【0009】

プロセスは、それ自体が反復的に繰返されるとき、プロダクト選択プロセスにおいて、最近の選択に基づく“変化 (mutation)”を使用して、ユーザが選択圧力 (selective pressure) として働く進化サーチ方式を実行する。このようなプロセスにより、“サーチ空間”の予想していなかった探索、すなわち小売店や図書館で使用されている普通のブラウジングプロセスにより類似した探索ができる。

20

【0010】

n次元のサーチ空間における値の最良の組を発見することが目的であるとき、既存の進化サーチ方式は最適化プロセスとして考えることができる。例えば、 $n = 1$ 、すなわち、1次元‘x’であるときは、関数 $f(x)$ が最大となるようなxの値を見付けることが目的である。xに対する異なる値を使用して、多数の‘単位体 (individual)’を生成する。 $f(x)$ が大きい値である単位体は、 $f(x)$ が小さい値である単位体よりも、より多くの“子”で‘報酬’を与えられる。単位体の子は、その単位体のxの値と同一ではないが、ほぼ同じ“x”の値を有する。全ての子は、 $f(x)$ の値にしたがって評価され、それにしたがって子供で報酬を与えられる。このプロセスは、終了条件が満たされるまで反復される。

30

【0011】

子の値が親の値とどのくらい異なるかを定める‘変化’演算子が適切に生成されることを条件として、このやり方で、連続的に、またはディスクリートに変化する次元をサーチすることができる。例えば、xが1ないし100の整数値であるときは、(適切な‘境界条件’を使用して、1より小さいか、または100より大きい値を避けて) 明白な変化演算子は、親の値から1を加算または減算して、子の値を与える。xが、1ないし100の実数値であるときは、(ここでも境界条件に応じて) 変化演算子は親の値付近の確率分布にしたがって子の値を摂動する。

40

【0012】

本発明では、データベース内の最良の項目を見付けることを目的とする。これは、データベース内の項目数が有限であるために、本質的にディスクリートなサーチである。ブラウジングとの関連において、“最良”を定める基準は、最初は定められていないことに注意すべきである。

【0013】

本発明の1つの態様では、これらの項目を、n個の属性にしたがって記載し、進化サーチにおいて、これらの属性はn次元である。各項目は、属性の組の各々において特定の値を割り当てられる(2つの項目間の類似度は、2つの項目の値が共通である属性数によって

50

識別することができる)。項目はユーザに表示され、目的関数にしたがって各項目を評価するのではなく、対象のユーザからの報酬によって、各項目の‘最適性’または‘適合性’を判断する。より多くの報酬を与えられた項目は、次の世代へ寄与する見込みが大きい。進化サーチにおけるように、親の属性を‘変化’させることによって、子が生成される。したがって、進化基準は、第1の項目の属性値の組から所定の度合い分だけ異なる1組の属性値を生成することによって選択され、第2の項目は、データベースから、生成された組の属性値に対応する属性値を有する項目から選択される。しかしながら、n次元のサーチ空間がデータベースに精密に整合しない可能性がある。n個の属性の中の幾つかの値は、データベース内の項目に対応しない。対照的に、n個の値から成る幾つかの組には、2個以上の項目がある。したがって、データベース内には実際には存在しない名目上の‘子’か、またはデータベース内の幾つかの実際の項目に対応する子を生成することになる。表示する次の項目を選択するプロセスにより、これらの両者が可能となる。

10

【0014】

進化基準は、2つ以上前の入力属性にしたがって判断され、したがって、装置には、装置の各動作サイクルごとに行われる選択を記録する記録手段が構成されていて、所定回数前からの装置の動作サイクルの中で選択された項目であって、記録された項目の属性に関係する1組の属性を生成するために、属性生成手段が配置されている。

【0015】

特別な属性の組が各項目ごとに定められ、組の中の各属性は、色、などのような特性ではなく、その項目と幾つかの他の項目との類似度を表わしている。サーチ空間内の各項目ごとに、これらの属性の中の1つがあてられていて、サーチ空間内の項目は互いに関係している。各特定値は、単に、2つの項目間に関係性があるか、ないかを定める。この実施形態では、サーチ空間は、連結グラフとして構成されると考えられる。言い換えると、同じ組の属性を使用して、サーチ空間を構成(または‘生成’)するのではなく、それをナビゲートするために、空間のナビゲーションは、一連の連結点にしたがう。進化基準は、第1の項目に関係する各属性の組の項目に対する所定の値を特定することによって生成することができる。

20

【0016】

これを実現するために、サーチ空間のシステムオペレータは、まず、いずれの項目が互いに“隣り合っている”かを判断しなければならない。これは、人間のオペレータが行うか、または各項目に、1組の属性の値を与え、次に属性と同じ次元数のサーチ空間内に置くといった半自動プロセスにおいて行われる。各項目は、1組の近隣、すなわちこの空間内の特定の距離内に位置する項目として定義されるものを有する。ここでも、プロセスを連続的に反復することにより、この距離を縮めることができる。ナビゲーションは、選択された項目から、1組の近隣の中の1つへ移動することにより行われる。これが、“変化”を構成している。

30

【0017】

別の実施形態では、一連の連結点にしたがうナビゲーションを可能にするのではなく、空間を連結グラフとして構成し直して、表示するための次の項目を、データベース内のどこからでも取入れられるようにしている。各属性は、重み付け値と関係付けられていて、第1の項目に関係する入力を受信すると、第1の項目に関係する属性の重み付け値が増加し、より大きい重みを割り当てられた属性を有する項目が、より小さく重みを付けられた属性よりも、選択される可能性がより大きいといった進化基準が生成される。1つの項目をクリックするとき、その項目に連結された全ての項目に報酬が与えられる。表示するための次の項目を選択すると、この方法では、サーチの途中で項目によって累積される報酬の数にしたがって、選択にバイアスをかける。

40

【0018】

生成された組の属性値を2つ以上前の入力属性にしたがって判断し、装置の各動作サイクルにおいて行われた選択を記録し、第2の項目の選択を制約して、所定回数前からの装置の動作サイクル内で既に選択された項目であって、記録手段によって記録された項目の

50

選択を防ぐ。

【0019】

開示されている実施形態は、オンラインの小売り品目、とくに衣服の購入に使用されるが、多くの他の適用も可能である。例えば、本発明は、衣服材料の‘購入者’が、購入者が希望する色、パターン、生地をブラウズするのに使用される。さらに加えて、不動産業者（不動産屋）および旅行代理店の業界では、売り出されている商品が本質的に表示し難いために、オンラインのブラウジングがとくに適しており、競売にかけられた住宅は、商品のカタログが大きく、商品がユニークであるために、幅広い客層に表示することが物理的に容易でない。本発明は、大きいデータベース（例えば、プレゼンテーションスライドのようなグラフィックディスプレイに取入れられる“クリップアート（clip-art）”画像）から他の項目を選択するのににも使用される。このデータベースの多くは、カタログに載せる基準の範囲が広いので、ブラウズするのが困難である。本発明は、既に選択された項目に類似する内容のニュース情報のポップアップウィンドウを準備して、オンラインでニュースを供給するのににも使用される。本発明は、実在の人々のデータベースをサーチすることによってか、または目撃者の証言から顔を作成することによって、犯罪者または行方不明者を識別するためのIdentikitまたはe-fitシステムへも適用される。本発明は、通常の意味でのオンラインではなく、店内のキオスク（In-store Kiosk）に適用され、実在の小売り店内の端末を使用して希望の商品を見付けてから、それを‘グッズアウト（goods out）’から取ってくることができるようにするものである。

10

【0020】

ここで、本発明の実施形態を、例示的に、添付の図面を参照して記載する。

20

【発明を実施するための最良の形態】

【0021】

図1は、ユーザ端末10が、狭帯域幅の電話接続のような通信ネットワーク11を介して、サーバ12へ接続されていることを示している。サーバはデータベース13へアクセスし、サーバ自体には、通常はソフトウェアによって実行される多数のサブシステムが構成されている。これらのサブシステムには、受信ポート14、セッション記録データベース15、進化プロセッサ16、選択プロセッサ17、および出力ポート18が構成されている。注文処理サーバ19もシステムに関係付けられている。

【0022】

素子の分布は変更できることが分かるであろう。例えば、ネットワーク11と主サーバ12との間にはクライアントサーバが置かれていて、記載されている実施形態では、端末10によって実行される機能の幾つかを実行する。その代りに、ユーザ端末10上でオンラインデータベース13からデータに直接にアクセスするプロセスを実行してもよい。

30

【0023】

使用の際に、システムは、局所的に、あるいはサービスプロバイダによって保持されるサーチ空間または‘ガーデン（garden）’をサーチャーに与える。これらは、サーチャーが、1つ、または多数のセッション中に、ユーザが関心をもっている項目を収集する領域である。サービスプロバイダは、‘サーチ空間’、すなわち全有効項目を含む多次元空間を生成する。簡素化した例は、図9に示されている。項目の近隣は、その項目と1つ以上の特徴が類似している項目で占められている。

40

【0024】

データベース13は、検査に使用でき、かつ多数の属性によって分類されている全ての項目のカタログを記憶している。例えば、衣服は、種類（靴、帽子、スカート、など）、色、パターン、スタイル、デザイナー、値段、などによって分類される。データベースは、手入力によって、またはカタログからデータを抽出することによって設定することができる。スプレッドシートに提示されている項目は、カテゴリーとして異なる列のエントリを使用して、データベースをコンパイルするのにとくに適している。

【0025】

図2には、本発明の第1の実施形態によって実行されるプロセスが示されている。サーチ

50

ングセッションは、次に記載するように実行される。端末10のユーザは、記述子を使用して、（新しいまたは既存の）サーチ空間または“ガーデン”を開く。サーチ空間または“ガーデン”は、全般的（例えば、‘衣服’）であっても、またはより特定の（例えば、“ズボン”）であってもよい。ある特定の他の制限を加えて、表示できる種々の項目の種類を制限してもよい。とくに、ユーザは衣服のサイズを特定して、ユーザ自身のサイズに対して無効の項目の表示を避けることができる。このような所定の制限にしたがって、選択プロセッサ17は、最初に、データベース13からランダムに項目を選択し、出力ポート18へ送り、ユーザへ順方向に送る（ステップ20）。大抵のホームユーザの装置において使用可能な狭帯域幅を最良に使用するために、出力ポート18には、バッファ記憶装置が構成されていて、データベース13からの項目をユーザ端末10へ連続的に供給することができる。その後で、新しい項目が（適切であれば、記述と画像とで）表示部内に到達し始める。これらの項目は、最初に、課されている最初の制限にしたがって、図9に示されている‘サーチ空間’内の全ての項目からランダムに選択される。ユーザ端末10は、新しい項目を受信すると、それに“経過期間”値（初期値はゼロ）を割り当てる（ステップ21）。この特徴は、古い順に、または別の項目が表示部へ加えられたときにインクリメントされ、所定の経過期間に到達した項目は、ガーデンから削除される。

10

20

30

40

50

【0026】

図4ないし10に示されている簡潔に示された例では、文字A、B、C、D．．．．．Zによって識別される多数の項目を表示することができる。これらの項目は、それぞれ多数の属性と関係付けられて、データベース13上に記憶されている。この例のために、図9では、3つのみの属性のカテゴリ、すなわち、衣服の種類、色、およびパターンが識別される。図8には、このような属性が他にも幾つか示されているが、この例では使用されていない。実際には、多数の異なる属性、例えば値段、デザイナー、年齢層、素材、サイズ、などが使用される。各項目は、これらのカテゴリの各々の属性（例えば、デザイナー Paul、ジャケット、グレイ、無地、大人用、ウール素材、胸囲96センチメートル、60ポンド）ごとにエントリされる。属性は、多次元の“サーチ空間”90における位置を定めると考えられ、“サーチ空間”90内では、属性を共有する項目は、関連する次元において互いに隣り合っていて、図9では、3つの例示的な属性について示されている。（図9では、イタリックで記載された項目B、E、F、O、P、U、Wは、他の項目を含む面とは異なる面内の位置を表わしている）。

【0027】

表示されている各項目は、各カテゴリごとに値を選択し、次に、それらの全ての値に整合する項目を選ぶことによって選択される。最初は、この選択にはバイアスはかけられない。例えば、11人の異なるデザイナーと、8つの異なる衣服の種類があるとき、11回の機会の中の1回は、‘デザイナー Paul’が、表示される最初の項目として選択されるデザイナーになり、8回の中の1回は、衣服がジャケットになる。ユーザは、希望するならば、表示部へ入力する項目を受動的に観察することができる。ユーザは、いつでも、自分にとって関心のある項目を識別することができる。このような項目は、さらに検討する価値のある類であるとして、ユーザを惹きつけるものであり、例えば、図4の表示部内の項目“J”である。

【0028】

項目が選択されると、その項目の経過期間値はゼロにリセットされ（ステップ22）、通信リンク11上で信号が受信ポート14へ伝送され（ステップ23）、データベース15を記録するセッション中に項目の識別子を記憶する。進化プロセッサ16は、進化サーチ空間技術を、受信したデータへ適用する。このときに、新しい項目はランダムなサンプルから外れて到来する。

【0029】

図2の実施形態では、進化プロセッサ16は、最初に、データベース13内に記憶されている選択された項目Jの属性を検索する（ステップ24）。進化プロセッサ16は、選択された項目のこれらの属性を使用して、表示する次の項目の属性の組を選択するランダムなプロセ

スにバイアスをかける。このバイアスと共に、属性を選択し、生成された‘進化した’属性を選択プロセッサ17へ送る（ステップ25）。好ましい構成では、進化プロセッサ16は、現在の選択ではなく、セッション記録データベース15から検索した最後の幾つかの選択の履歴を使用して、いずれの属性が、属性の選択にバイアスをかけることに影響を与えるかを判断する。したがって、新しい選択は2つ以上の“親”をもつ。

【0030】

式で表現すると、

$$P_{c_i} = (1 + n_{c_i}) / (N_c + n_{total})$$

なお、 P_{c_i} は、表示する次の画像を選択するときの確率であり、値 i は、カテゴリー c から選択され、

n_{c_i} は、サーチャーが、カテゴリー c において値 i の項目をクリックした回数であり、

N_c は、カテゴリー c の異なる値の数であり、

n_{total} は、このセッションにおけるサーチャーによるクリックの合計回数である。

【0031】

例えば、11人の異なるデザイナーと、8つの異なる衣服の種類とがあるとき、最初は、‘デザイナー Paul’ が、表示する第1の項目として選択されるデザイナーになる見込みは、11回中の1回であり、衣服がジャケットになる確率は8回中の1回である。サーチャーがデザイナー Paul のジャケットを選択すると、重み付けが変更され、表示される第2の項目が、同じデザイナーによる別の項目になる確率は、 $1/11$ (0.091) から $2/12$ (0.167) へ上がり、別のジャケットを得る見込みは重み付けされ、 $1/8$ (0.125) ないし $2/9$ (0.222) へ上がる。同じデザイナーから別のジャケットを得る確率は、依然として非常に低く、 $1/27 = 0.037$ であるが、これは、 $1/88 = 0.011$ のランダムな確率よりも高いことが容易に分かるであろう。特定のカテゴリー値を選択する確率へのインクリメントは、ユーザセッション中に累積し続ける。最初の重み付けは、各カテゴリー内の使用可能な項目数を考慮せず - 各デザイナーが、最初に選択される見込みは等しいが、個々の項目の多くはデータベース内にあることに注意すべきである。

【0032】

ランダムな選択プロセスにおいて、次の項目に対するカテゴリー値が決定されると、項目のデータベース(13)は、選択プロセッサ(17)によってサーチされ、これらの値に整合する項目を識別する。2つ以上の項目がこれらの基準を満たすときは、これらの項目の一方がランダムに等しい確率で選択される。

【0033】

セッション記録データベース15を調べて、可能であれば、同じ基準で別の選択を行って（ステップ26）、既に提案された項目が反復されないことを保証する（ステップ27）。（これらの基準を有する項目が全て既に選択されているために）このような項目を選択する所定回数の試行が失敗するときか、またはデータベース内に、カテゴリー値の組の項目がないときは、カウンタ（システム271）はタイムアウトし、新しい組のカテゴリー値が生成される（ステップ25へ戻る）。

【0034】

次に、選択プロセッサ17は、選択された項目を出力ポート18へ送り、順方向へ伝送する。ユーザ端末では、システムによって与えられた各提案が表示部に加えられ、最高の“経過期間”を有する項目が置き換えられる（ステップ28）。

この方法は、簡単で、計算上効率的であり、別途記載するように、多数のユーザがいる状態に容易に適用することができる。 $1/11$ (0.091) から $2/12$ (0.167) への変化が、セッションの後半において、 $33/217$ (0.152) から $34/218$ (0.156) への変化よりも大きくなる相当に前に、カテゴリー値が報酬を与えられないとき、報酬から得られる確率の割合の変化は最大であるので、同様に、迅速にサーチに集中する傾向がある。これは、集中対サーチの良好な混合に依存して、長所であるときと、長所でないときとがある。必要であれば、確率に対する選択に関係する異なる関数を使

10

20

30

40

50

用することができる。

【0035】

しかしながら、サーチャーの項目と通常でないカテゴリー値との組み合わせとを示すことに対して、サーチにバイアスがかけられる。例えば、デザイナー P a u l による大人用のグレイのジャケットは幾つかあるが、デザイナー P a u l による大人用の赤のズボン
は1つしかない。さらに加えて、可能なカテゴリー値の組合せの空間が、実際の項目で十分に満たされていないとき、すなわち、最もランダムに生成されたカテゴリー値の組が、データベース内の項目に対応せず、項目に整合する値の組合せに到達する前に、何度も“サイコロを振りなおす”ことを要求されるアルゴリズムになるときは、効率に悪影響を及ぼすことになる。

10

【0036】

図3の実施形態のために、リンクは、図10に示されているように、ある特定の項目間で定められている。これらのリンクは、項目をカテゴリー分類する個々の属性に関係しているか、またはサーチャーの好みを示すリサーチデータによって経験的に判断される。実際には、このような関係を判断する両方の方法を使用して、図10に示されているリンクを定める。

【0037】

図3のプロセス(ステップ30、31、32、および33)は、採用されるサーチ方式が異なるために、サーチ空間の進化がランダムから外れるときまで、図2(ステップ20、21、22、23)のプロセスの手続きに類似している。図3の実施形態では、所定の近隣のリストは、図
8の最右列に示されているように、各項目ごとに生成され、図10の項目間のリンクによって示されている。リンクは、2つの項目間に存在する接続に関係することに注意すべきである。例えば、所与の商品の購入者が別の商品と一緒に購入することを示す市場のリサーチデータを使用して、他の点では明らかに関係のない項目間のこのようなリンクを生成する。近隣を選択する確率が等しい場合は、リンクは全て単位値であり、近隣を選択する確率が、リンクの値か、または‘重み付け’に比例する場合は、0と1との間の実数値をとる。

20

【0038】

この実施形態では、項目は、以前に選択された項目の近隣リストから選択され(ステップ36)、表示される(ステップ38)。このやり方では、表示部は、全てがユーザによって選
択されるか、またはこのような項目へリンクされる項目のグループへ“進化”する。図2の実施形態では、重複を避けるために検査を行い(ステップ37)、可能であれば、近隣リストから再び選択する(ステップ371、再びステップ36)か、またはこのような項目が使用可能でないときは、ランダムな選択を行う(ステップ361)。その後で、選択プロセッサ17は、(ステップ36または361において)選択された項目を出力ポート18へ送り、順方向に伝送する。ユーザ端末では、システムによって与えられた各候補(suggestion)を表示部へ加え、最高“経過期間”を有する項目と置き換える(ステップ38)。

30

【0039】

図11に示されている第3の実施形態では、サーチ空間は、図3の実施形態における項目間のリンクを使用して構成される。しかしながら、この場合に、表示する次の項目は、図
2の実施形態のように、データベース内の全ての項目から確率的に選択される。リンクを使用することにより、‘報酬’、すなわち1つの項目をクリックし、隣接項目に拡張、それらの項目が表示されるために、バイアスをかけられたランダムセレクトによって、選択される確率を高めるといった‘報酬’が得られる。

40

【0040】

各反復では、表示するための1つの項目をデータベースからランダムに選択する(ステップ46)。1回目の反復を除いて、全ての項目が等しく表示される可能性が高いとき、個々の項目が表示するために選択される確率は、前の反復の結果にしたがって重み付けされる。最初に検査をして(ステップ47)、項目が以前に表示されていなかったこと(この場合は、新しい選択を行うこと)を保証し、新しく選択された項目を表示部へ加え、最も古い

50

ものを外す（ステップ48）。表示部上の項目の経過期間はインクリメントされる。ユーザは、表示部から項目を選択する（ステップ42）。このような選択を行うと、選択された項目にリンクされたデータベース内の各項目の重み付けが増加され（ステップ45）、次の反復では、ユーザが既に関心を示した項目へ選択にバイアスをかける。さらに加えて、ユーザは、他の実施形態におけるように、選択した項目を購入するオプションを与えられる（ステップ49）。

【0041】

次に、調節された重み付けを使用して、データベースから別の項目を選択する（ステップ46）。所定の間隔を置いた後で、ユーザが選択を行わないときは（ステップ42）、既存の重み付けに基づいて、選択を行う（ステップ46）。

10

前の実施形態と同じ例を採用し、この場合は、 $M \times M$ （ M は項目の合計数である）の行列を生成する。各行は、1項目に対応し、その行内の各エントリは、その項目と別の項目との間の関係の強さを示す数値である。

【0042】

M 個の項についてのベクトルもあり、サーチングセッションが進むときに更新される。ベクトルの各項 p_n は、対応する項目 n を選択する確率を表わしている。最初に、全ての項 p_n は、等しく設定される。

次に示す項目は、確率計数 p_n を考慮して、ランダムに選択される。

【0043】

サーチャーのクリックに応答して、次に示すように、ベクトルを更新する：

20

- i) サーチャーは項目をクリックする；
- ii) その項目に対応する $M \times M$ の行列内の行を見付ける；
- iii) その行内の項目の値をベクトルに加える；
- iv) ベクトルを正規化して、項目の和を M に等しくする。

【0044】

このプロセスは、項目間の関係の精細な細分性を与える。これは、カテゴリーに結び付ける必要はない（1デザイナーのグレーのジャケットを、別のデザイナーのオリーブ色のスカートと強く関係付けたいときは、可能である）。関係性は対称的である必要はない（ジャケットが好きな人にスカートを示してもよく、スカートが好きな人にジャケットを示してもよい）。この方法がサーチに集中する速さは、行列の更新を示す関数をパラメータ化することによって設定できる。したがって、システムのオペレータ、またはサーチャーでさえも、アルゴリズムの「探索対探索」を変更することができる。

30

【0045】

サーチャーのクリック動作に関する情報を失なうのを避けるために（このような情報の損失は、この方法において単一の更新されたベクトルを使用するときに発生する）、クリックの履歴を使用して、（例えば、赤いものを探すといった）サーチャーのサーチを主に駆動するものに関して推論する。これは、この情報に依存し、項目どうしをリンクする方法によって、失った情報を暗黙的に回復するよりも、より迅速に集中できる。この変形を実行するには、3つのステップがある。

【0046】

40

最初に、1組の「仮」ベクトルであって、各ベクトルが M 個の二値項を含むものを生成する。 M 番目の位置における「1」は、「 M 番目の項目がこの仮説にしたがう」ことを意味する。例えば、仮説は「赤の項目」であり、ベクトルは、単に、赤の項目に対応する各位置に1のみを有する。

【0047】

次に、サーチャーが項目 M をクリックすると、 M 番目の項目がインクリメントされるときに、履歴ベクトルが維持される。

履歴ベクトルは、仮ベクトルと比較され、これが、ユーザの動作に関する最も可能性の高い説明であることを示唆する。

このプロセスにより、特定の顧客が赤色が好きだといった、広範な情報を抽出することが

50

できる。十分なデータ（長い単一のセッション、または多数のユーザの短いセッション）を使用して、新しい仮説、すなわち説明的な意味を保持し、かつ小売り業者に有益な仮説をたてることができるようにする。

【0048】

ユーザ間の相関関係を使用する多数の方法において、これらの実施形態を展開することができる。ここでは、2つのアプローチを記載する。一方は、購入された商品にリンクするアプローチであり、他方はナビゲーションの動作のみに依存するアプローチである。

【0049】

先ず、ある期間にわたって、多くのユーザのサーチングセッションからデータを集める。データベース内の各項目について、最も人気のある購入される項目に関係する情報を検索することができる。ユーザセッションの履歴における時間数に関して、セッション中の幾つかの時点で、1つの特定の項目の選択が、他の特定の項目の最終的な購入に対応するといった情報も記憶することができる（例えば、1000個のユーザセッションを記録し、デザイナー Paul のグレーのジャケットを最終的に購入したユーザが、デザイナー Peter の緑色のセーターを与えられたことが15回があったことが分かった。このとき、セーターはジャケットと15の重みでリンクされる）。

10

【0050】

新しいセッションが開始されると、情報を使用して、‘一番売れている’項目を選択的に表示する。これは、セッション中にランダムに行われるか、またはとくに、しばらくの間ユーザに項目の報酬を与えられないときに行われる。現在のセッション中の選択の履歴をみることによって、表示されている一番売れている項目を選択する。選択された各項目は、それを一番売れているものの各々に関係付ける一定の強さ（恐らくは、ゼロ）のリンクを有する。各一番売れているものに対して、報酬を与えられた項目のリンク強度が加算される。表示される項目は、これらの最も売れているものから、リンク強度の和に比例する確率で選択される。

20

【0051】

代わりのアプローチは、項目どうしのリンクを使用してナビゲートされるサーチ空間に依存し、図3および12の実施形態に適用することができる。これは、ユーザの選択動作を使用して、項目間のリンク行列内の値を直接に変更する。項目Cに報酬を与え、次に項目Dに報酬を与えることにより、 (c, d) の行列値がインクリメントされる。行列が対称的であることが望ましいときは、 (d, c) も、もちろんインクリメントできる。次に与えられる項目がEであるときは、 (d, e) がインクリメントされる。 (c, e) も比較的に小さくインクリメントされる。

30

【0052】

図12において、Dを選択する前に、項目A、B、およびCを選択すると仮定すると、ユーザセッション中に、一連の4つの選択が既に行われていることになる。

図12は、5番目の報酬結果、すなわち項目Eをクリックした結果を示している。第1のインクリメント関数 F_1 は、新しく報酬を与えられた項目（E）を、既に報酬を与えられた項目（D）へ関係付ける行列値に適用される。第2、第3、などの関数 F_2 、 F_3 、 F_4 、などは、連鎖をさらに戻った項目と、新しく選択された項目との間の直接的なリンクをインクリメントする。もちろん、B、C、およびDをクリックする間に、A、B、C、およびD間のリンクは、既にインクリメントされている。

40

【0053】

ユーザセッション中に行列に対して行われる変更は、中心に保持されている行列へ取入れられ、次のユーザのシステムの動作に影響を与える。

実際の項目において、行列値に適用される関数の各々のインクリメントは小さく、単一のユーザの動作は、サーチ空間をあまり歪めないことが予想される。

【0054】

システムのオペレータは、同じプロセスを使用して、項目間に新しいリンクを生成するか、または既存のリンクの強度を変更する。管理者は、リンクされる項目を探すサーチ空間

50

をブラウズする。例えば、管理者は、スクリーン上に３つの赤い項目を見ると、各項目をクリックして、項目どうしをリンクさせる（または、リンクを強化する）。

【 0 0 5 5 】

実施形態の幾つかのシステムによって生成される表示部は、ユーザと同様に生成され、一般的な表示部は、図 4 ないし 7 に示されている。各項目と関係付けられている“経過期間値”も、図 4 ないし 7 に示されているが、これは通常は表示されない。図 4 に示されている表示部から、ユーザは項目“J”を選択し、既に記載したように、X、Kを追加する。図 4 および 5 と比較することによって分かるように、最も高い“経過期間”値を有する項目 F、P から、X、K へ置き換わる。項目“J”の経過期間値は、ゼロにリセットされる。続いて、プロセスにおいて、ユーザは、例えば項目“R”（図 5 参照）を選択し、項目“S”および“L”を表示する。項目“S”および“L”は、（図 2 のプロセスを実行するときは）項目“R”と属性を共有するか、または（図 3 のプロセスにおいて）各々が項目“R”の近隣リストにある。現在最も高い経過期間値を有する項目“C”および“W”を項目“S”および“L”へ置き換え、項目“R”の経過期間をゼロにリセットする。次に項目“L”を選択して（図 6 参照）、２つの別の項目“N”および“M”を表示し（図 7 参照）、現在最も高い経過期間値の項目“J”および“X”を置換する。

10

【 0 0 5 6 】

セッションが進むと、表示部は、ユーザによって最近選択されたか、またはこのような項目から派生した項目を次第に取入れる。項目が選択されたときにその経過期間をゼロにリセットするか、または最初に選択したものを、表示スクリーン上の別のウィンドウ上へ移動することによって、選択した項目を継続的に提示することができる。したがって、進化したガーデンには項目の組合せ、すなわちユーザが関心をもっていることが分かっている項目と幾つかの属性を共有しているが、恐らくは、セッションの開始時はサーチャーが意図していなかった少なくとも幾つかのものを含んだ組み合わせが含まれている。各新しい項目は、ユーザが以前に選択した幾つかの項目にしたがって選択されるので、表示部は、ユーザが関心をもっている可能性の高い項目を示すように次第に進化する。経過期間値がインクリメントするとき、ユーザがそれらを選択することを促さないという提案は、システムから消え、同様の提案は行われない。

20

【 0 0 5 7 】

ユーザは、注文したい項目を見ると、以前のように、現在は、その項目を選択するが、ユーザがそれを購入したいことを示す信号 29（39，49）を注文処理サーバ 19 へ送る。当業者には、分かるように、本発明を実行するのに使用されるソフトウェアの一部または全ては、種々の伝送媒体、または記憶媒体、あるいはこの両者上に含まれ、したがってプログラムを汎用コンピュータ上にロードすることができるか、または適切な伝送媒体を使用して、コンピュータネットワーク上でダウンロードすることができる。コンピュータプログラムは、適切なコンピュータ入力デバイス、例えば CD-ROM、光学読出しマスク、磁気媒体、パンクチャされたカードまたはテープ、あるいは電磁信号または光信号によって読み出し可能な適切な担体上で実現できる。

30

【図面の簡単な説明】

【 0 0 5 8 】

40

【図 1】本発明を実行するために協働する種々の要素間の相互関係を模式的に示す図。

【図 2】本発明の第 1 の実施形態によって実行されるプロセスを示すフローチャート。

【図 3】本発明の第 2 の実施形態によって実行されるプロセスを示すフローチャート。

【図 4】プロセスを例示的に実行する間に現れる表示部を示す図。

【図 5】プロセスを例示的に実行する間に現れる表示部を示す図。

【図 6】プロセスを例示的に実行する間に現れる表示部を示す図。

【図 7】プロセスを例示的に実行する間に現れる表示部を示す図。

【図 8】図 4 ないし 7 に示されているプロセスを支援するのに使用されるデータベースを示す図。

【図 9】図 8 に示されているデータをグラフで示した図。

50

【図 10】異なるグラフ形式で同じデータを示す図。

【図 11】本発明の別の実施形態を示す図。

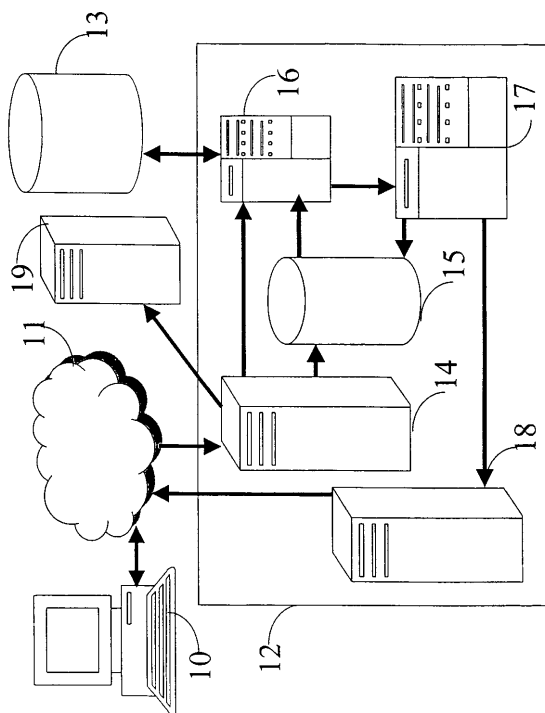
【図 12】図 3 ないし 11 の実施形態の多数のユーザの変形に使用される追加のステップを示す図。

【符号の説明】

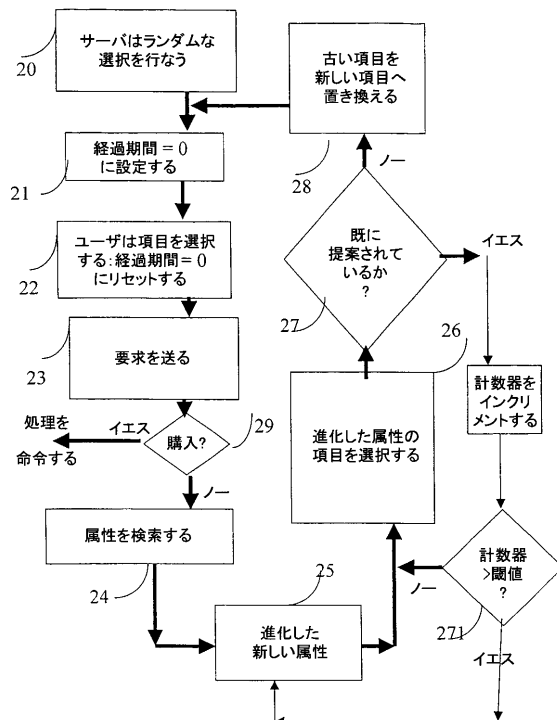
【0059】

10 ユーザ端末、11 通信ネットワーク、12 サーバ、13 データベース、14 受信ポート、15 セッション記録データベース、16 進化プロセッサ、17 選択プロセッサ、18 出力ポート、19 注文処理サーバ。

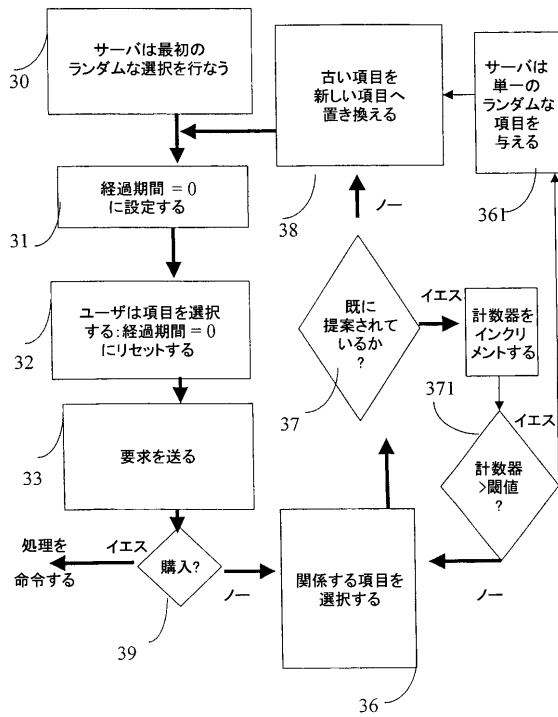
【図 1】



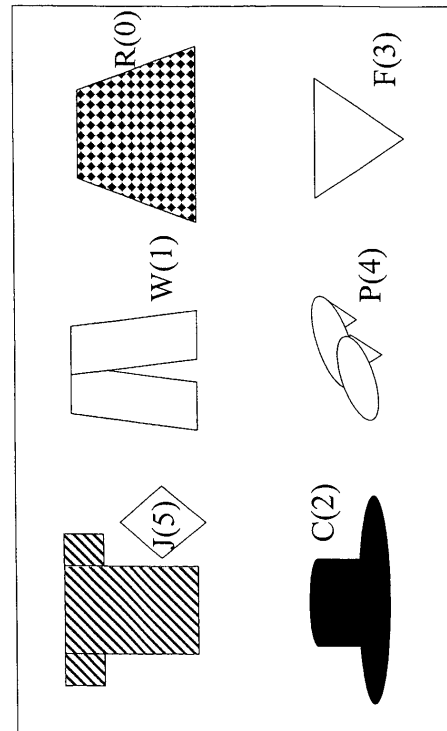
【図 2】



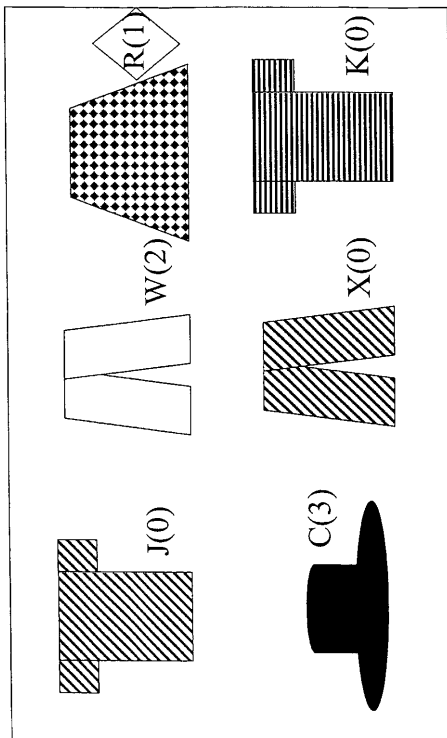
【 図 3 】



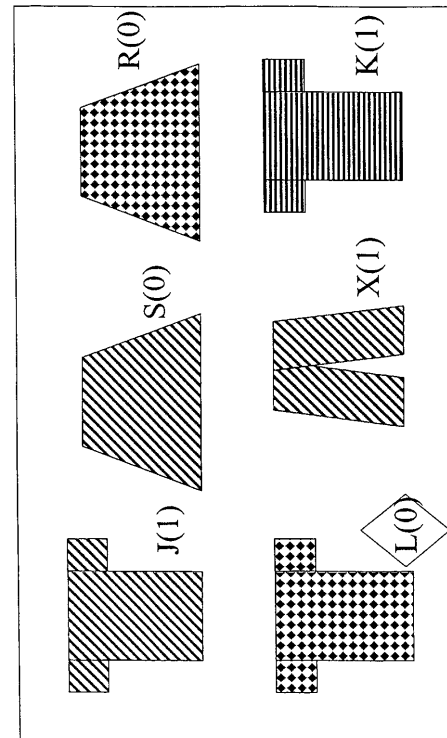
【 図 4 】



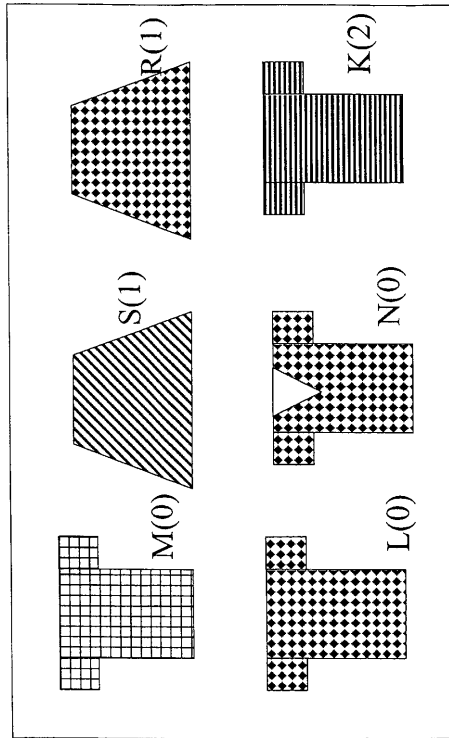
【 図 5 】



【 図 6 】



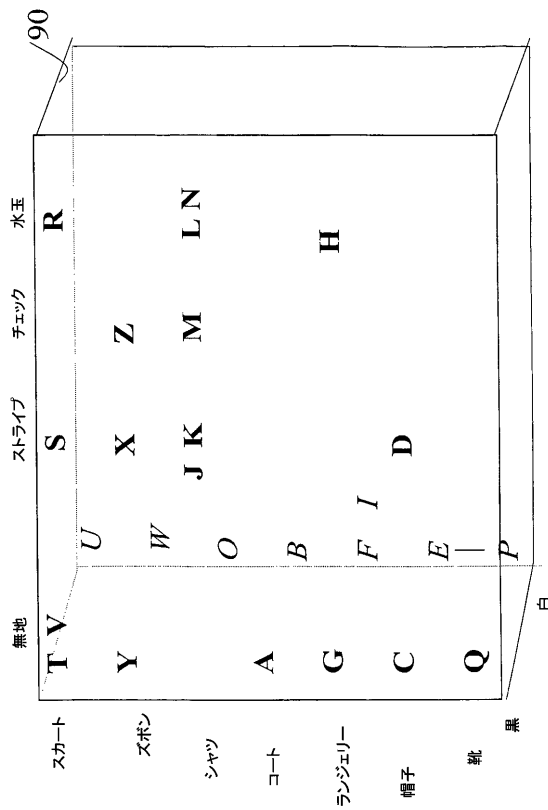
【図 7】



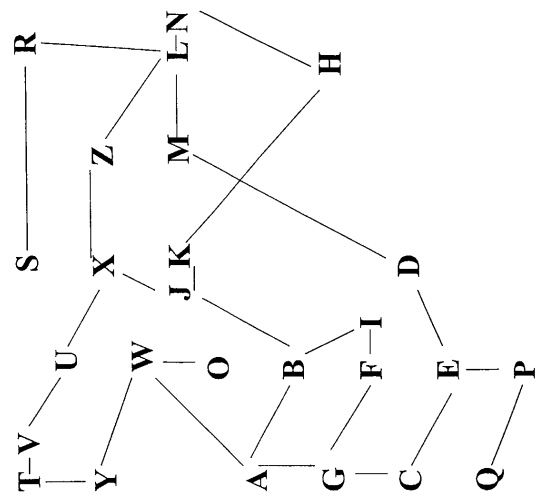
【図 8】

項目	種類	柄	色	値段	デザイナー	近隣
A	コート	無地	黒			B G W
B	コート	無地	白			A I J
C	帽子	無地	黒			E F G
D	帽子	ストライプ				E M
E	帽子	無地	白			C D P
F	ランジェリー	無地	白			G I
G	ランジェリー	無地	黒			A C F
H	ランジェリー	水玉				K N
I	ランジェリー	無地	白			B F
J	シャツ	ストライプ				B X
K	シャツ	ストライプ				J H
L	シャツ	水玉				M N R Z
M	シャツ	チェック				D L
N	シャツ	水玉				L H
O	シャツ	無地	白			W
P	靴	無地	白			C E
Q	靴	無地	黒			P
R	スカート	水玉				L S
S	スカート	ストライプ				R
T	スカート	無地	黒			V Y
U	スカート	無地	白			V X
V	スカート	無地	黒			T U
W	ズボン	無地	白			A O Y
X	ズボン	ストライプ				U J Z
Y	ズボン	無地	黒			T W
Z	ズボン	チェック				L X

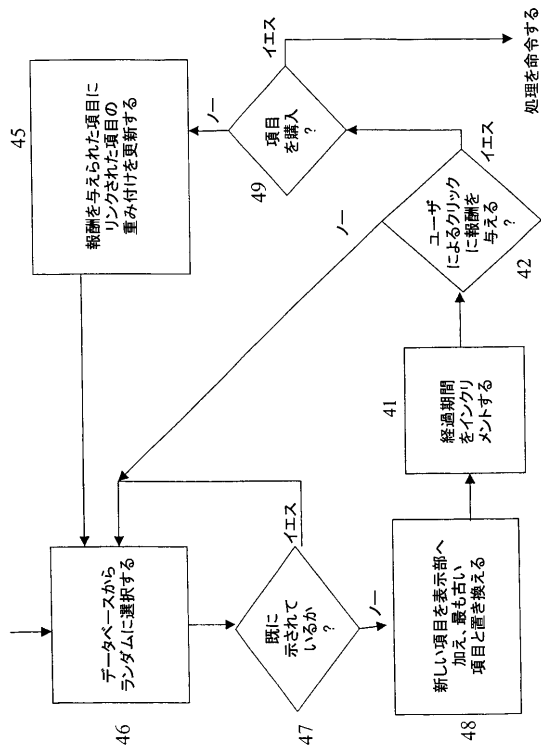
【図 9】



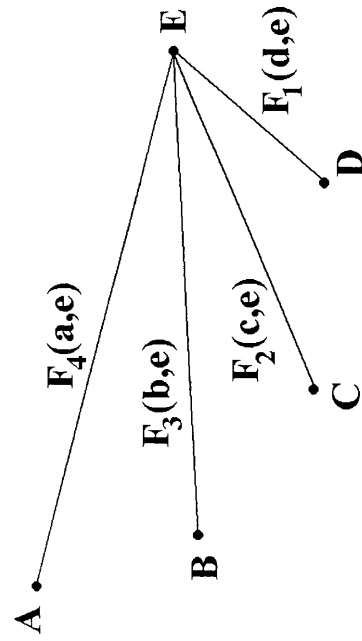
【図 10】



【図 1 1】



【図 1 2】



【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
10 October 2002 (10.10.2002)

PCT

(10) International Publication Number
WO 02/080025 A2

- (51) International Patent Classification: **G06F 17/30**
- (21) International Application Number: PCT/GB02/01107
- (22) International Filing Date: 12 March 2002 (12.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
01302892.3 28 March 2001 (28.03.2001) EP
- (71) Applicant (for all designated States except US): **BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY** [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).
- (72) Inventors; and
(75) Inventors/Applicants (for US only): **TATESON, Jane, Elizabeth** [GB/GB]; 145 High Street, Wickham Market, Suffolk IP13 0RD (GB). **TATESON, Richard, Edward** [GB/GB]; 145 High Street, Wickham Market, Suffolk IP13 0RD (GB).
- (74) Agent: **LIDBETTER, Timothy, Guy, Edwin**; BT Group Legal Services, Intellectual Property Department, 8th floor, Holborn Centre, 120 Holborn, London EC1N 2TE (GB).
- (81) Designated States (national): AF, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LI, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SI, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— with declaration under Article 17(2)(a); without abstract;
title not checked by the International Searching Authority
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/080025 A2

(54) Title: DATA RETRIEVAL SYSTEM

(57) Abstract:

WO 02/080025

PCT/GB02/01107

1

DATA RETREIVAL SYSTEM

This invention relates to data retrieval systems, and in particular to systems for assisting users making a selection from a large range of available items. It has application in searchable databases in which there are a large number of variables to consider and the user needs freedom to search according to his own preferred criteria, but the database is too large, or the user's criteria too poorly defined, for a fully structured search to be possible.

In searchable databases a searcher is generally forced to navigate along a branching decision 'tree' towards a destination that will hopefully be what he wants. This is a good method for searching towards a known objective. However, because paths must be retraced to arrive at different destinations such a system is not so good for less structured searching ("browsing") where the objective is less clearly defined, or where several objectives may need to be inspected. The searcher is entirely at the mercy of the database's categorisation and will be unlikely to make chance finds, or form a general impression of what is available and thus direct his choices (a common strategy when shopping for clothes for example).

The shortcomings of online searching are magnified still further when the bandwidth of the link between the user and the database is low. An attempt to 'browse' an online database via a modem, for example, typically consists of a pause while the homepage loads, a relatively rapid selection by the searcher of a section within the database, another pause while the section page loads, rapid selection of a category of items, a further pause, etc. etc. Mobile access to the internet will mean that relatively low bandwidth online searching is likely to continue to grow even as people adopt high bandwidth connections for their fixed links.

According to the invention there is provided an apparatus for selecting items from a database for display, comprising a data-storage means for storing, for each item, data indicative of the degree of similarity between that item and other items in the database;

input means for receiving a user input identifying a first item in said database

evolution processor means for specifying an evolved specification having a predetermined degree of similarity to the first item,

WO 02/080025

PCT/GB02/01107

2

identifying from the database one or more variant items meeting the evolved specification,

selection means for selecting a second item from amongst the variant items

and output means for displaying an output identifying the selected second

5 item.

The invention also extends to a method of selecting items from a database for display, comprising the steps of:

generating data indicative of the similarity between each item and other items in the database;

10 receiving an input identifying a first item in said database;

generating an evolved specification, identifying a predetermined degree of similarity to the first item,

selecting an item in the database meeting the evolved specification,

displaying the selected second item.

15 The invention also extends to a computer program for performing the method of the invention, and to a computer program product directly loadable into the internal memory of a computer, comprising software code portions for performing the steps of the method when the product is run on a computer.

The invention also extends to a computer program product stored on a computer usable medium, comprising:

20 computer-readable program means for causing a computer to generate data indicative of the similarity between each item and other items in a database

computer-readable program means for causing the computer to receive an input identifying a first item in said database

25 computer-readable program means for generating an evolved specification, specifying a predetermined degree of similarity to the first item,

computer-readable program means for causing the computer to select an item in the database meeting the evolved specification,

30 computer-readable program means for causing the computer to generate a display of the selected second item.

Preferably a plurality of items are displayed, and the input means has means for selecting one of the displayed items to be the first item. In a preferred arrangement the output means has means for controlling the display means to replace one of the

WO 02/080025

PCT/GB02/01107

3

plurality of items, preferably the item which has been displayed for longest without having been selected, by the selected second item. This can be an iterative process wherein the selected second item is added to the display, a further input being received relating to one of the displayed items, for selection of a further item for display. In order to identify which item is to be replaced, each displayed item may be allocated an age value, the age value being initially set at zero and incremented periodically, the age value of a given displayed item being reset to zero if that item is selected, and the age values being used to identify items for deletion from the display. Selection of the second item may be constrained to prevent selection of the same item within a predetermined number of iterations of the process.

The process, when allowed to repeat itself iteratively, allows the product selection process to perform an evolutionary search strategy with the user acting as the selective pressure, using "mutations" based on the most recent selection or selections. Such a process can create a serendipitous exploration of 'search space', more akin to the natural browsing process used in a shop or library.

Existing evolutionary search strategies can be thought of as optimisation processes where the goal is to find the best set of values for the n dimensions of the search space. For example if $n = 1$, i.e. there is a single dimension 'x', the goal is to find the value of x for which some function $f(x)$ is maximal. A number of 'individuals' are created with different values for x . The individuals that have high values for $f(x)$ are 'rewarded' with more "children" than individuals giving low values for $f(x)$. The children of an individual have values for "x" which are similar to, but not identical to, the value of x for that individual. All the children are then evaluated according to their values for $f(x)$, and rewarded with children accordingly. This process iterates until some termination condition is met.

Continuously or discretely varying dimensions can be searched in this way, provided the 'mutation' operators which dictate how a child's value may vary from its parent's are constructed appropriately. For example, if x can take integer values from 1 to 100, an obvious mutation operator would be to add or subtract 1 from the parent's value to give the child's value (with suitable 'boundary conditions' to avoid less than 1 or more than 100). If x can take any real value from 1 to 100, the mutation operator might be to perturb the child's value according to some probability distribution around the parent's value (with boundary conditions again).

WO 02/080025

PCT/GB02/01107

4

In the invention the goal is to find the best item in a database. This is an inherently discrete search – as the number of items in the database is finite. It should be noted that in the context of browsing the criteria defining “best” are not defined initially.

5 In one version of the invention these items are described according to n attributes, and those attributes are the n dimensions for an evolutionary search. Each item may be allocated specified values for each of the set of attributes, (the degree of similarity between any two items being identifiable by the number of attributes for which the two items have values in common). Items are displayed to the user, and
10 rather than evaluating each item according to some objective function, the ‘optimality’ or ‘fitness’ of each item is determined by rewards from the subjective user. The more rewarded items have a greater chance of contributing to the next generation. As in evolutionary search, a child is generated by ‘mutating’ the attributes of the parent. The evolved specification is therefore selected by generating
15 a set of attribute values differing by a predetermined degree from the set of attribute values of the first item, the second item being selected from the database from those having attribute values corresponding to the generated set of attribute values. However, it is possible that the n dimensional search space will not precisely match the database. There may be some values of the n attributes, which do not
20 correspond to any item in the database. Conversely for some sets of n values there may be more than one item. Therefore we could create a notional ‘child’, which does not actually exist in the database, or a child that corresponds to several real items in the database. The process of choosing the next item to display allows for both these possibilities.

25 The evolved specification may be determined according to the attributes of two or more previous inputs, and to this end the apparatus may comprise recording means for recording the selections made on each cycle of operation of the apparatus, wherein the attribute generation means is arranged to generate a set of attributes related to attributes of items recorded as having been selected within a
30 predetermined number of previous cycles of operation of the apparatus

A special attribute set may be defined for each item, each attribute of the set representing, not properties such as colour, etc, but the degree of similarity between that item and some other item. For each item in the search space, there would be

WO 02/080025

PCT/GB02/01107

5

one of these attributes relating to each other item in the search space. Each specified value may simply define the presence or absence of an association between the two items. In this embodiment the search space may be considered to be organised as a connected graph. In other words, rather than using the same set of attributes to
5 organise (or 'create') search space, and to navigate it, navigation of the space follows a series of connected points. The evolved specification can then be generated by specifying a predetermined value for the term of each attribute set relating to the first item.

In order to achieve this the system operator of the search space must first
10 determine what items should be "adjacent" to each other. This may be done by a human operator, or in a semi-automated process in which every item is given values for a set of attributes and then placed in a search space that has the same number of dimensions as there are attributes. Every item then has a set of neighbours, defined as being those items located in this space within a specified distance. Again, this
15 distance may reduce on successive iterations of the process. Navigation is carried out by moving from the selected items to one of its set of neighbours. This constitutes a "mutation".

In a further embodiment the space is once again organised as a connected graph, but rather than allowing navigation to follow a series of connected points, the
20 next item to display may be taken from anywhere in the database. Each attribute may be associated with a weighting value, such that on receipt of an input relating to a first item, the weighting values of the attributes associated with the first item are increased, and the evolved specification is generated such that items having attributes allocated higher weightings have a greater probability of selection than
25 those with lower-weighted attributes. Clicking on an item rewards all items connected to that item. When choosing the next item to display, this method biases the choice according to the number of rewards accumulated by items over the course of the search.

The generated set of attribute values may be determined according to the
30 attributes of two or more previous inputs, and the selections made on each cycle of operation of the apparatus may be recorded, and the selection of the second item constrained to prevent selection of an item recorded by the recording means as

WO 02/080025

PCT/GB02/01107

6

having already been selected within a predetermined number of previous cycles of operation of the apparatus.

Although the embodiments to be described are used for on-line retail shopping, and specifically for clothes, many other applications are possible. For example, the invention may be used for fashion material 'buyers' to browse towards colours, patterns, textures they like. On-line browsing is also particularly suited to fields of estate agency (real estate) and travel agency, where the products on sale are inherently difficult to display, and auction houses, which have big catalogues of items that, because they are unique, cannot readily be physically displayed to a wide audience. The invention may also be used for selecting other items from a large database, such as "clip-art" images for incorporation in graphic displays such as presentation slides, many databases of which are difficult to browse because of the wide range of criteria under which they might be catalogued. The invention may also be used for on-line news feeds, arranging for pop-up windows with news information having content similar to items previously selected. The invention may also be applied to Identikit or e-fit systems for identifying criminals or missing persons, either by searching through a database of real people, or by generating a face from a witness's description. Rather than being on-line in the usual sense, the invention may be applied to an In-store Kiosk, for finding a desired item using a terminal in a real shop before collecting it from 'goods out'.

Embodiments of the invention will now be described, by way of example only and with reference to the drawings, in which:

Figure 1 illustrates schematically the inter-relationships between the various elements that co-operate to perform the invention;

Figure 2 is a flow chart illustrating the process performed by a first embodiment of the invention;

Figure 3 is a flow chart illustrating the process performed by a second embodiment of the invention;

Figures 4, 5, 6 and 7 illustrate displays that may appear during an illustrative run of the process.

Figure 8 illustrates the database used to support the processes illustrated in Figure 4 to 7.

Figure 9 illustrates in graphical form the data depicted in Figure 8.

WO 02/080025

PCT/GB02/01107

7

Figure 10 illustrates the same data in a different graphical form.

Figure 11 illustrates a further embodiment of the invention

Figure 12 illustrates an additional step used in a multiple user variant of the embodiments of Figures 3 and 11

5 Figure 1 illustrates a user terminal 10 connected through a communications network 11 such as a low-bandwidth telephone connection to a server 12. The server has access to a database 13, and itself comprises a number of subsystems, which will typically be implemented by software. These subsystems include a receive port 14, a session recording database 15, an evolution processor 16, a selection
10 processor 17, and an output port 18. An order-processing server 19 is also associated with the system.

It should be understood that the distribution of the elements may be varied. For example a client server, interposed between the network 11 and main server 12, may perform some of the functions performed by the terminal 10 in the described
15 embodiment. Alternatively, the process could be run on the user terminal 10, accessing the data directly from an online database 13.

In use, the system offers the searcher one or more search spaces or 'gardens', which can either be held locally or by the service provider. These are the areas within which the searcher, over the course of one or many sessions, will
20 cultivate a collection of items of interest to the user. The service provider creates a 'search space' which is a multidimensional space consisting of all items available. A simplified example is shown in Figure 9. The neighbourhood of an item is populated with items that, in one or more characteristics, are similar to that item.

The database 13 stores a catalogue of all the items available for inspection,
25 classified by a large number of attributes. For example, clothes may be classified by type (shoes, hats, shirts, etc), colour, pattern, style, designer, price and so on.

The database can be set up by manual entry or by extracting data from a catalogue. Items presented in spreadsheet form are particularly suited for compiling the database, using the various column entries as the categories.

30 The process performed by a first embodiment of the invention is represented in Figure 2. A searching session operates as follows. The user of the terminal 10 opens a search space or "garden" (new or pre-existing) with a descriptor, which may be general (e.g. 'clothes') or more specific (e.g. "trousers"). Certain other limitations

WO 02/080025

PCT/GB02/01107

8

may be added to limit the variety of items available for display: in particular the user may specify clothes sizes, to avoid the display of items not available in the user's own size. Subject to any such predetermined limitations, the selection processor 17 selects items, initially at random, from the database 13 and passes them to the
5 output port 18 for onward transmission to the user (step 20). To make the best use of the narrow bandwidth available on most home user's equipment, the output port 18 includes a buffer store so that it can continuously provide the user terminal 10 with items from the database 13. New items then start arriving in the display (description plus picture wherever appropriate). Initially these items are randomly
10 selected from all items within the 'search space' shown in Figure 9, subject to any initial limitations imposed. When the user terminal 10 receives a new item it allocates it an "age" value, which is initially zero (step 21). This characteristic is incremented either in accordance with chronological time or when further items are added to the display, and items achieving a predetermined age are deleted from the
15 garden.

In the simplified illustrated examples shown in Figures 4 to 10, a number of items, identified by the characters A, B, C, D.....Z are available for display. These are stored on the database 13 each with a number of associated attributes. For the purposes of this example only three attribute categories are identified in Figure 9,
20 namely garment type, colour and pattern. Several further such attributes are listed in Figure 8, although not used in this example. In practice many different attributes such as price, designer, age group, material, size, etc, would be used. Each item would then have an entry for an attribute in each of these categories (e.g. Designer Paul, Jacket, Grey, plain, adult, wool mixture, 96cm chest, £60). The attributes may
25 be considered as defining a position in a multidimensional "search space" 90, in which items sharing an attribute would be adjacent to each other in the relevant dimension, as shown for the three illustrative attributes in Figure 9. (In Figure 9 the italicisation of items B, E, F, O, P, U, W represents their location in a different plane from that containing the other items).

30 Each item to be displayed is selected by choosing values for each category and then picking an item that matches all those values. Initially this selection is unbiased. For example, if there are eleven different designers and eight different garment types, there would be a 1 in 11 chance that 'Designer Paul' would be the

WO 02/080025

PCT/GB02/01107

9

designer chosen for the first item to display and a 1 in 8 chance that the garment would be a jacket. The user can passively observe items entering the display as long as he likes. At any time the user may identify an item of interest to him. Such an item would be one that attracts the user as being of a kind worthy of further consideration, for example the item "J" in the display of Figure 4.

When an item is selected the age value of that item is reset to zero (step 22) and a signal is transmitted over the communications link 11 (step 23) to the receive port 14, causing the product identifier to be stored in the session recording database 15. The evolution processor 16 then applies an evolutionary search space technique to the data received. This is the point at which the arrival of new items deviates from a random sample.

In the embodiment of Figure 2 the evolution processor 16 firstly retrieves the attributes of the selected item J stored in the database 13 (step 24). It then uses these attributes of the selected item to bias the random process of choosing the set of attributes for the next item to display. With this bias included, the choice of attributes is then made and the resulting 'evolved' attributes are passed to the selection processor 17 (step 25). In a preferred arrangement the evolution processor 16 uses the history of the last few selections retrieved from the session recording database 15, and not just the current selection, to determine which attributes to influence the biasing of attribute choice. This allows new selections to have more than one "parent".

As an equation:

$$P_{ci} = (1 + n_{ci}) / (N_c + n_{total})$$

where:

P_{ci} is the probability that when choosing the next image to display, the value i will be chosen for category c

n_{ci} is the number of times that the searcher has clicked on an item with the value i for its category c

N_c is the number of different values of category c

n_{total} is the total number of clicks by the searcher in this session

For example, if there are eleven different designers and eight different garment types, there would initially be a 1 in 11 chance that 'Designer Paul' would be the designer chosen for the first item to display and a 1 in 8 chance that the

WO 02/080025

PCT/GB02/01107

10

garment would be a jacket. If the searcher selects a Designer Paul jacket, then the weightings are modified such that the probability that the second item displayed would be another item by the same designer rises from 1/11 (0.091) to 2/12 (0.167), and the chance of getting another jacket would be weighted to rise from 1/8 (0.125) to 2/9 (0.222). (It can easily be seen that the probability of getting another jacket from the same designer is still very small, at $1/27 = 0.037$, but this is greater than the random probability of $1/88 = 0.011$. The increments to the probabilities of choosing particular category values continue to accumulate throughout the user session. Note that the initial weightings do not take account of the number of available items in each category – each designer has an equal chance of being selected initially, however many of his individual products appear in the database.

Once the random selection process has decided on the category values for the next item, the database of items (13) is searched by the selection processor (17) to identify items that match those values. If more than one item satisfies these criteria, one of them is chosen at random with equal probability

The session recording database 15 is consulted to ensure that items that have already been suggested are not repeated (step 27), with another selection having the same criteria being made if possible (step 26). If a predetermined number of attempts to select an item having these criteria fail (because all such items have previously been selected, or if there is no item in the database with the set of category values, a counter (system 271) times out and a new set of category values is generated (return to step 25).

The selection processor 17 next passes the selected items to the output port 18 for onward transmission. At the user terminal, each suggestion offered by the system is added to the display, displacing the item having the greatest "age" (step 28).

This method is simple and computationally efficient, and can readily be extended to a multi-user situation as will be discussed. It also tends to focus the search rapidly because the percentage change of the probability resulting from a reward is largest when that value for the category has not been rewarded much before (change from 1/11 (0.091) to 2/12 (0.167) is bigger than, later in session, 33/217 (0.152) to 34/218 (0.156)). This might or might not be an advantage

WO 02/080025

PCT/GB02/01107

11

depending on what is a good mix of focus versus search. It would be possible to use a different function relating selections to probabilities if required.

However, the search is biased towards showing the searcher items with unusual category value combinations: for example there might be several different
5 Designer Paul grey jackets for adults but only one Designer Paul red trousers for adults. Furthermore, its efficiency will be adversely affected if the space of possible category value combinations is sparsely populated with actual items i.e. if most randomly generated sets of category values do not correspond to any item in the database, resulting in the algorithm having to "re-roll the dice" many times before
10 hitting on a combination of values which does match an item.

For the purpose of the embodiment of Figure 3, links are defined between certain items, as shown in Figure 10. These links may relate to individual attributes by which the items are categorised, or may be determined empirically by research data indicative of searcher preferences. In practice, both methods of determining
15 such associations may be used to define the links illustrated in Figure 10

The processes of Figure 3 (steps 30, 31, 32, and 33) follow a similar procedure to that of Figure 2 (steps 20, 21, 22, 23) up to the point where the evolution of the search space departs from random, as the search strategy employed is different. In the embodiment of Figure 3 a predetermined neighbour list is
20 generated for each item as shown in the right hand column of Figure 8 and indicated by the links between items in Figure 10. It should be noted that a link could relate to any connection that may exist between the two items. For example market research data indicating that purchasers of a given item commonly also buy another item may be used to generate such a link between otherwise apparently unrelated items. The
25 links may all be of unit value, in which case there is an equal probability of choosing any neighbour, or may take real values between 0 and 1, in which case the probability of choosing any neighbour is proportional to the value, or 'weighting', of the link.

In this embodiment an item is selected from the neighbour list of the
30 previously selected item (step 36), and is displayed (step 38). In this way the display is made to "evolve" towards a group of items that are all either selected by the user, or linked to such items. As in the embodiment of Figure 2, a check is made (step 37) to avoid duplication, and a further selection from the neighbour list made if possible

WO 02/080025

PCT/GB02/01107

12

step 371, step 36 repeated) or, if no such item is available, a random selection is made (step 361). The selection processor 17 next passes the items selected (in step 36 or 361) to the output port 18 for onward transmission. At the user terminal, each suggestion offered by the system is added to the display, displacing the item having
5 the greatest "age" (step 38).

In a third embodiment, shown in figure 11, the search space is organised using links between items as in the embodiment of figure 3. However, in this case the next item to display is chosen probabilistically from all items in the database, more like the embodiment of figure 2. The links are used to allow the 'reward' of
10 clicking on one item to spread to neighbouring items and hence increase the probability that those items will be chosen for display by the biased random selector.

On each iteration, one item is selected at random from the database (step 46) for display. Except on the first iteration, when all items are equally likely to be displayed, the probabilities of individual items being selected for display are weighted
15 according to the results of previous iterations. A check is first made (step 47) to ensure the item has not been displayed before (in which case a new selection is made), and the newly selected item is added to the display, displacing the oldest (step 48). The ages of the items on display are then incremented. The user may then select an item from the display (step 42). If such a selection is made, the weightings
20 of each item in the database linked to the selected item are increased (step 45), so that on subsequent iterations the selection is biased towards items in which the user has previously shown interest. The user is also offered the option of buying the selected item (step 49) as in the other embodiments.

Another item is then selected from the database (step 46), using the
25 adjusted weightings. If after a predetermined interval no selection is made by the user (step 42) a selection is made based on the existing weightings (step 46)

Taking the same example as the previous embodiment, in this case an $M \times M$ matrix is created where M is the total number of items. Each row corresponds to an item, and each entry in that row is a number indicating the strength of association
30 between that item and another item.

There is also a vector with M terms, which is updated as the searching session goes on. Each term p_n of the vector represents the probability that the corresponding item n will be selected. Initially, all terms p_n are set equal.

WO 02/080025

PCT/GB02/01107

13

The next item to show is chosen randomly, taking into account the probability factors p_n .

The vector is updated in response to the searcher's clicks as follows:

- i) the searcher clicks on an item
- 5 ii) the row in the $M \times M$ matrix corresponding to that item is found
- iii) the values of the terms in that row are added to the vector
- iv) the vector is normalised so that the sum of terms is equal to M

This process gives a fine granularity of relationships between items. It needn't be tied to categories (if we want to make a grey jacket by one designer
10 highly associated with an olive shirt by a different designer, we can). The relationships need not be symmetrical (people who like the jacket could be shown the shirt but not vice versa). The rapidity with which this method focuses the search can be set by parameterising the function describing the update of the matrix. This allows the system operator or even the searcher to alter the 'exploitation versus
15 exploration' of the algorithm. .

To avoid loss of information about a searcher's clicking behaviour, which would otherwise occur when the single updated vector is used in this method, the history of clicks may be used to make inferences about the main driver(s) of the searcher's search (e.g. looking for red things). This may allow faster focus than
20 relying on this information being implicitly recovered by the item-to-item links method. To carry out this variant, there are three steps

Firstly, a set of 'hypothesis' vectors, each containing M binary terms are produced. A '1' in the M th position means 'the M th item conforms to this hypothesis'. For example, the hypothesis might be 'red items' and the vector would
25 simply have a 1 in each position corresponding to a red item.

Next, a history vector is maintained where the M th term is incremented whenever the searcher clicks on item M

The history vector is compared with the hypothesis vectors to infer which are the most likely explanations for user behaviour.

30 This process allows the extraction of comprehensible information, such as a preference by a particular customer for the colour red. With enough data (a long single session, or many users' short sessions) it may allow the formulation of new

WO 02/080025

PCT/GB02/01107

14

hypotheses, which also retain some explanatory meaning and hence might be useful to retailers.

These embodiments may be developed in a number of ways to make use of correlations between users. Two approaches are discussed here, one linked to
5 purchased items and another relying only on navigation behaviour.

In the first, data is gathered from the searching sessions of many users over time. Information can be retrieved relating to the most popular purchased items and, for each item in the database. Information can also be stored relating to the number of times over the history of user sessions that selecting one specified item at some
10 point in the session correspond to eventual purchase of some other specified item. (e.g. we record 1000 user sessions and find that there were fifteen occasions when a user who eventually bought the Designer Paul grey jacket had rewarded the Designer Peter green sweater. The sweater is now linked to the jacket with a weighting of 15).

15 When a new session starts, this information may be used to preferentially display a 'top selling' item. This could be done at random throughout the session, or particularly when the user has not rewarded any items for a while. The top selling item to display is picked by looking at the history of selections during the current session. Each item selected will have a link of a certain strength (possibly zero)
20 relating it to each of the top sellers. For each top seller the link strengths of rewarded items are summed. The item to display is then chosen from those top sellers with a probability proportional to the sum of link strengths.

An alternative approach relies on the search space being navigated using item-to-item links and can be applied to the embodiments of figures 3 and 12. It uses the
25 selection behaviour of users to directly alter values in the matrix of links between items. Rewarding an item C and then rewarding an item D leads to an increment in matrix value at (c,d). If it is desired that the matrix is symmetrical, it can of course also increment (d,c). If the next rewarded item is E, there will be an increment to (d,e). There may also be a lesser increment for (c,e).

30 In Figure 12 we are to imagine there were selections of items A, B and C which preceded the selection of D so there has already been a sequence of four selections during the user session.

WO 02/080025

PCT/GB02/01107

15

Figure 12 shows the result of the fifth reward i.e. clicking on item E. The primary incrementing function F_1 is applied to the matrix value relating the new rewarded item (E) to the previous (D). Secondary, tertiary, etc. functions F_2 , F_3 , F_4 , etc increment links directly between items further down the chain and the newly selected item. Of course during the clicks on B, C and D there have already been increments to the links among A, B, C and D.

The alterations made to the matrix during a user session would be incorporated into the centrally held matrix and would thus affect the behaviour of the system for subsequent users.

10 In practical terms, it is anticipated that the functions applied to the matrix values would each result in small increments so that the behaviour of a single user does not grossly distort the search space.

The system operator may use the same process to generate new links between items, or alter the strength of existing links. An administrator browses the
15 search space looking for items that are to be linked. For example, if the administrator sees three red items on the screen, clicking on each will make (or strengthen) the links between them.

The displays generated by the systems of any of the embodiments may appear similar to the user, and typical displays are shown in Figures 4 to 7. The
20 "age" value associated with each item is also shown in Figures 4 to 7, although this would not normally be displayed. From the display shown in Figure 4, the user selects item "J", and as previously described items X, K are added. These replace the items F, P with the highest "age" value, as shown by comparison of Figures 4 and 5. The age value of item "J" is re-set to zero. The process then continues: the user
25 selecting, for example item "R" (Figure 5), and items "S" and "L", which each share an attribute with item "R" (if running the process of Figure 2) or which are each on the neighbour list of item "R" (process of Figure 3) are displayed (Figure 6), replacing items "C" and "W" which now have the highest age values, and resetting the age value of item "R" to zero. Item "L" is next selected (Figure 6), causing the display of
30 two further items "N" and "M" (Figure 7), replacing items "J" and "Y" which now have the highest age values.

As the session proceeds, the display will be increasingly populated with items that have either been recently selected by the user, or are descended from

WO 02/080025

PCT/GB02/01107

16

such items. The continued presence of selected items is achieved by the resetting of their ages to zero when they are selected or by moving top choices to a separate window on the display screen. The evolved garden therefore includes a mixture of items sharing some attributes with items the user found interesting, but probably
5 including at least some things which were not in the searcher's mind when the session began. As each new item is selected in accordance with the previous few selections made by the user, the display will gradually evolve to show items likely to be of interest to the user. Suggestions that do not prompt the user to select them disappear from the system as their age value increments, and no similar suggestions
10 are made.

When the user sees an item he wishes to order, he selects the item as before but now transmits a signal 29 (39, 49) indicating he wishes to purchase it, which is directed to the order-processing server 19.

As will be understood by those skilled in the art, any or all of the software
15 used to implement the invention can be contained on various transmission and/or storage mediums, so that the program can be loaded onto one or more general purpose computers or could be downloaded over a computer network using a suitable transmission medium. The computer program may be embodied on any suitable carrier readable by a suitable computer input device, such as CD-ROM, optically
20 readable marks, magnetic media, punched card or tape, or on an electromagnetic or optical signal.

WO 02/080025

PCT/GB02/01107

17

CLAIMS

1. Apparatus for selecting items from a database for display, comprising a data-storage means for storing, for each item, data indicative of the degree of similarity
5 between that item and other items in the database;
input means for receiving a user input identifying a first item in said database
evolution processor means for specifying an evolved specification having a
predetermined degree of similarity to the first item,
identifying from the database one or more variant items meeting the evolved
10 specification,
selection means for selecting a second item from amongst the variant items
and output means for displaying an output identifying the selected second
item.
- 15 2. Apparatus according to claim 1, comprising display means for displaying a
plurality of items, and wherein the input means has means for selecting one of the
displayed items to be the first item.
3. Apparatus according to claim 2, wherein the output means has means for
20 controlling the display means to replace one of the plurality of items by the selected
second item.
4. Apparatus according to claim 3, wherein the selected second item replaces
the item which has been displayed for longest without having been selected.
- 25 5. Apparatus according to claim 4, further comprising means for generating a
display of the selected items, means for allocating to each displayed item an age
value, the age value being initially set at zero, means for periodically incrementing the
age value of each displayed item, means for re-setting the age value of a given
30 displayed item to zero in response to an input received by the input means identifying
that item, and means for deleting from the display items having an age value greater
than a predetermined value.

WO 02/080025

PCT/GB02/01107

18

6. Apparatus according to claim 1, 2, 3, 4 or 5 wherein the data storage means comprises means for allocating to each item specified values for each of a set of attributes, the degree of similarity between any two items being identified by the number of attributes for which the two items have values in common.

5

7. Apparatus according to claim 6, further comprising attribute evolution generation means for generating a set of attribute values differing by a predetermined degree from the set of attribute values of the first item, and selection means for selecting from the database a second item having attribute values corresponding to
10 the generated set

8. Apparatus according to claim 1, 2, 3, 4, or 5, wherein the data storage means comprises means for allocating to each item specified values for each of a set of attributes defining its degree of similarity to each other item

15

9. Apparatus according to claim 8, the evolution specification means comprising means for defining an evolved specification specifying a predetermined degree of similarity to the first item

20 10. Apparatus according to claim 8 or 9, wherein each specified value defines the presence or absence of an association between the two items

11. Apparatus according to claim 7 or 8, wherein the generated set of attribute values is determined according to the attributes of two or more previous inputs.

25

12. Apparatus according to claim 11 comprising recording means for recording the selections made on each cycle of operation of the apparatus, wherein the attribute generation means is arranged to generate a set of attributes related to attributes of items recorded as having been selected within a predetermined number
30 of previous cycles of operation of the apparatus

WO 02/080025

PCT/GB02/01107

19

13. Apparatus according to claim 6, 7, 8, 9, 10, or 11 comprising means for associating with each attribute a weighting value, and means for increasing the weighting values of the attributes associated with a first item on receipt of an input relating to the first item, and wherein the means for retrieving the second item is
5 operable such that items having attributes allocated higher weightings have a greater probability of selection than those with lower-weighted attributes.

14. Apparatus according to any preceding claim, comprising recording means for recording the selections made on each cycle of operation of the apparatus, and
10 wherein the selection means is arranged to constrain the selection of the second item to prevent selection of an item recorded by the recording means as having already been selected within a predetermined number of previous cycles of operation of the apparatus.

15. Method of selecting items from a database for display, comprising the steps of:
generating data indicative of the similarity between each item and other items in the database;
receiving an input identifying a first item in said database;
20 generating an evolved specification, identifying a predetermined degree of similarity to the first item,
selecting an item in the database meeting the evolved specification,
displaying the selected second item.

25 16. A method according to claim 15, wherein a plurality of items is displayed, the first item being one of the items so displayed.

17. A method according to claim 16, being an iterative process wherein the selected second item is added to the display, a further input being received relating to
30 one of the displayed items, for selection of a further item for display.

18. Method according to claim 16 or 17 wherein one of the plurality of displayed items is replaced by the selected second item.

WO 02/080025

PCT/GB02/01107

20

19. A method according to claim 18, wherein each displayed item is allocated an age value, the age value being initially set at zero and incremented periodically, the age value of a given displayed item being reset to zero if that item is selected, and the age values being used to identify items for deletion from the display.

5

20. Method according to claim 19, wherein the selected second item replaces the item that has been displayed for longest without having been selected.

21. A method according to claim 18, 19, or 20 wherein the selection of the second item is constrained to prevent selection of the same item within a predetermined number of iterations of the process.

22. A method according to any of claims 15 to 21, wherein each item has allocated specified values for each of a set of attributes, the degree of similarity between items being identified by the number of attributes for which they have values in common.

23. A method according to claim 22, wherein the evolved specification is selected by generating a set of attribute values differing by a predetermined degree from the set of attribute values of the first item, and the second item is selected from the database from those having attribute values corresponding to the generated set of attribute values.

24. A method according to claim 23, wherein the evolved specification is determined according to the attributes of two or more previous inputs.

25. A method according to any of claims 15 to 21, wherein, for each item, an attribute set is defined, the terms of the attribute set representing the degree of similarity between that item and each other item.

30

26. Method according to claim 25, wherein the evolved specification is generated by specifying a predetermined value for the term of each attribute set relating to the first item.

WO 02/080025

PCT/GB02/01107

21

27. Method according to claim 25 or 26, wherein each specified value defines the presence or absence of an association between the two items

28. A method according to claim 21, 22, 23, 24 or 25, wherein each attribute is associated with a weighting value, and wherein on receipt of an input relating to a first item, the weighting values of the attributes associated with the first item are increased, and wherein the evolved specification is generated such that items having attributes allocated higher weightings have a greater probability of selection than those with lower-weighted attributes.

10

29. A computer program for performing the steps of any of claims 15 to 28.

30. A computer program product directly loadable into the internal memory of a computer, comprising software code portions for performing the steps of any of claims 15 to 28 when said product is run on a computer.

31. A computer program product stored on a computer usable medium, comprising:

computer-readable program means for causing a computer to generate data indicative of the similarity between each item and other items in a database

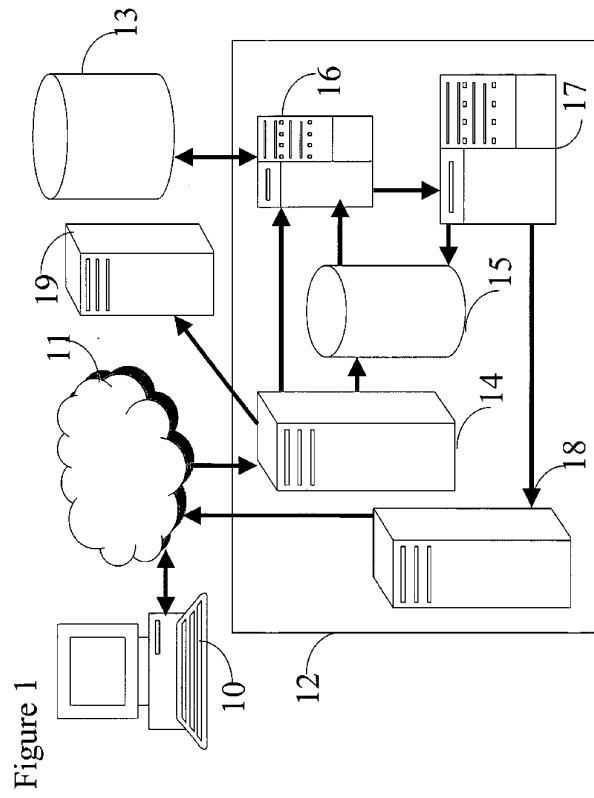
computer-readable program means for causing the computer to receive an input identifying a first item in said database

computer-readable program means for generating an evolved specification, specifying a predetermined degree of similarity to the first item,

computer-readable program means for causing the computer to select an item in the database meeting the evolved specification,

computer-readable program means for causing the computer to generate a display of the selected second item.

1/12



WO 02/080025

PCT/GB02/01107

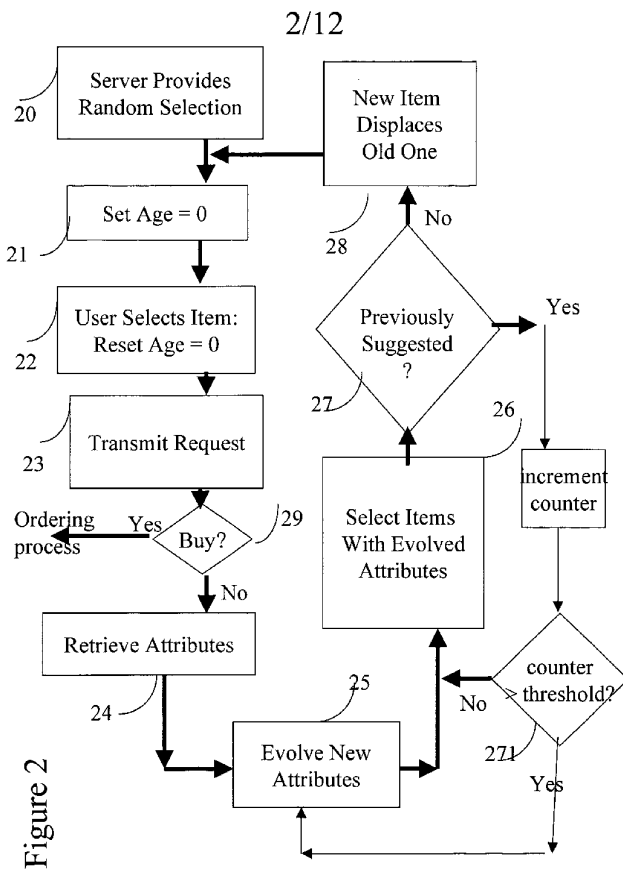


Figure 2

WO 02/080025

PCT/GB02/01107

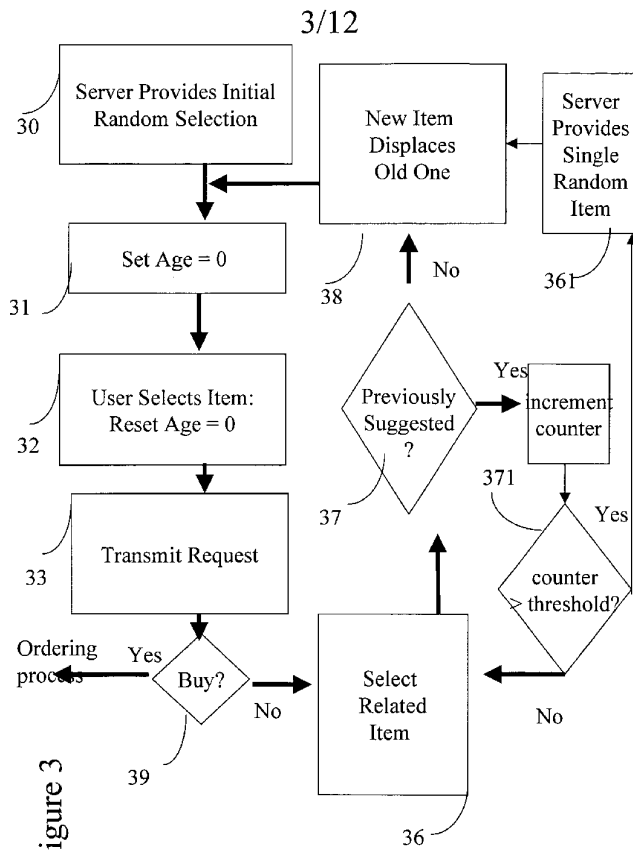
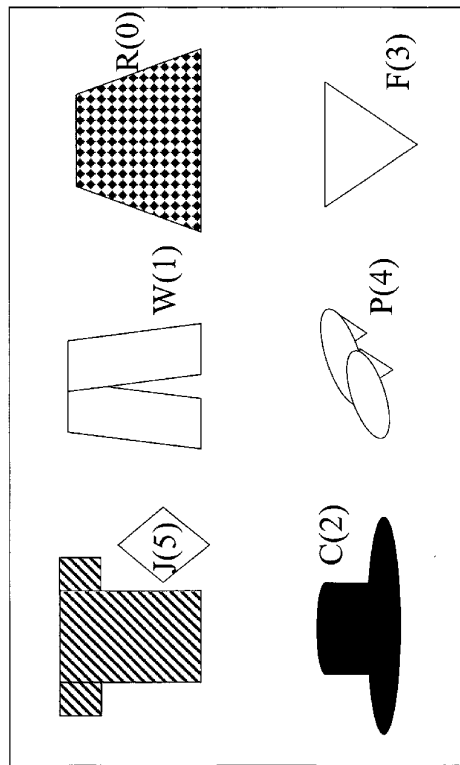


Figure 3

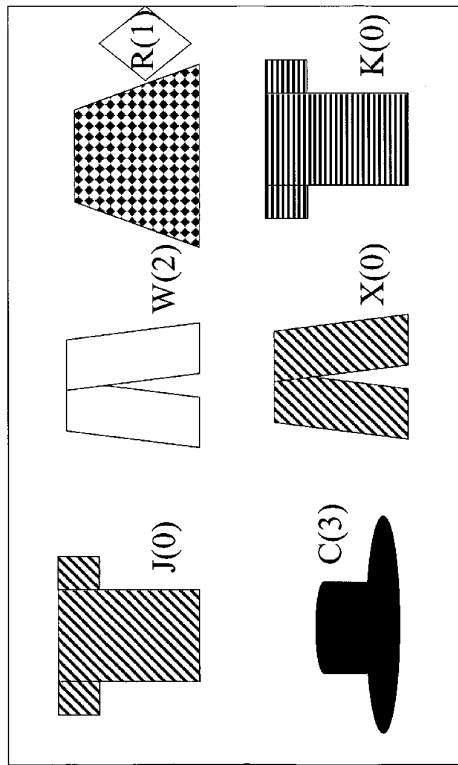
4/12

Figure 4



5/12

Figure 5



6/12

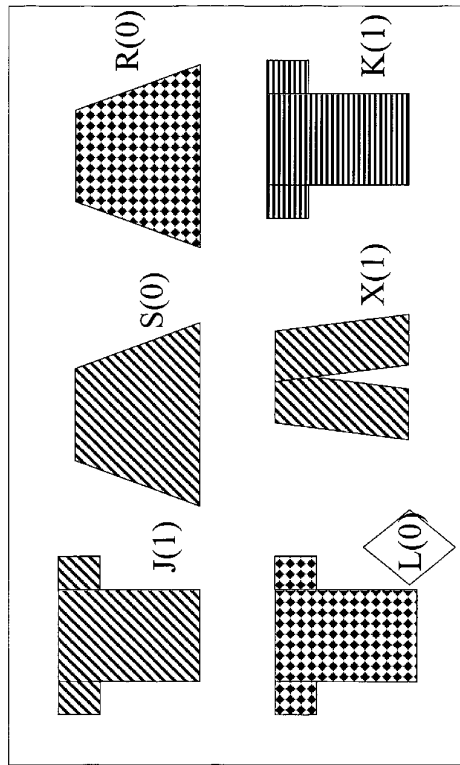
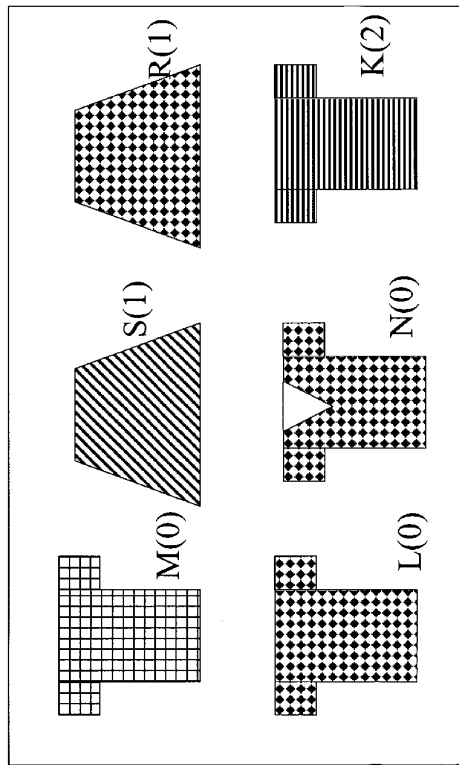


Figure 6

7/12

Figure 7



WO 02/080025

PCT/GB02/01107

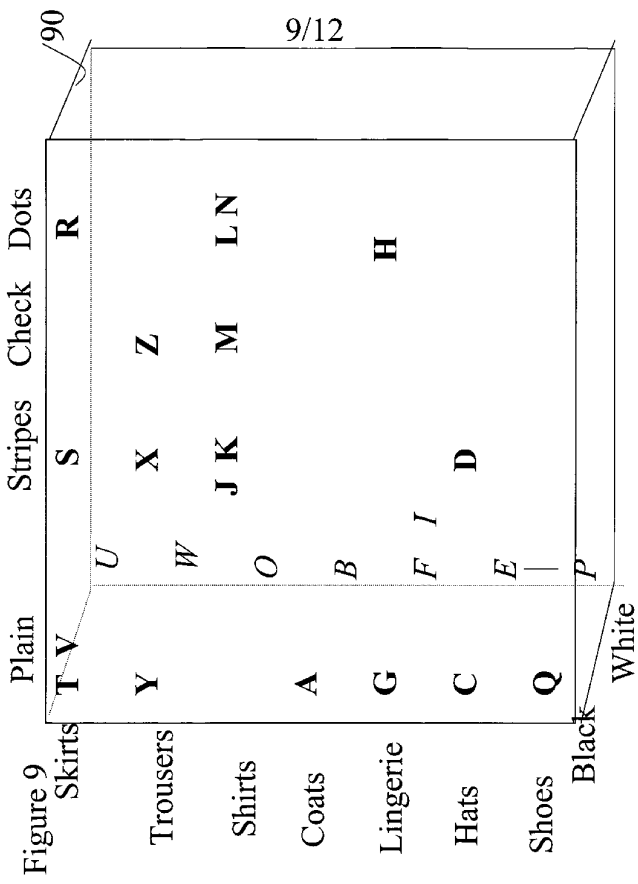
8/12

Figure 8

Item	Type	Pattern	Colour	Price	Designer	Neighbours
A	Coat	Plain	Black			B G W
B	Coat	Plain	White			A I J
C	Hat	Plain	Black			E F G
D	Hat	Stripes				E M
E	Hat	Plain	White			C D P
F	Lingerie	Plain	White			G I
G	Lingerie	Plain	Black			A C F
H	Lingerie	Dots				K N
I	Lingerie	Plain	White			B F
J	Shirt	Stripes				B X
K	Shirt	Stripes				J H
L	Shirt	Dots				M N R Z
M	Shirt	Check				D L
N	Shirt	Dots				L H
O	Shirt	Plain	White			W
P	Shoes	Plain	White			C E
Q	Shoes	Plain	Black			P
R	Skirt	Dots				L S
S	Skirt	Stripes				R
T	Skirt	Plain	Black			V Y
U	Skirt	Plain	White			V X
V	Skirt	Plain	Black			T U
W	Trousers	Plain	White			A O Y
X	Trousers	Stripes				U J Z
Y	Trousers	Plain	Black			T W
Z	Trousers	Check				L X

WO 02/080025

PCT/GB02/01107



WO 02/080025

PCT/GB02/01107

10/12

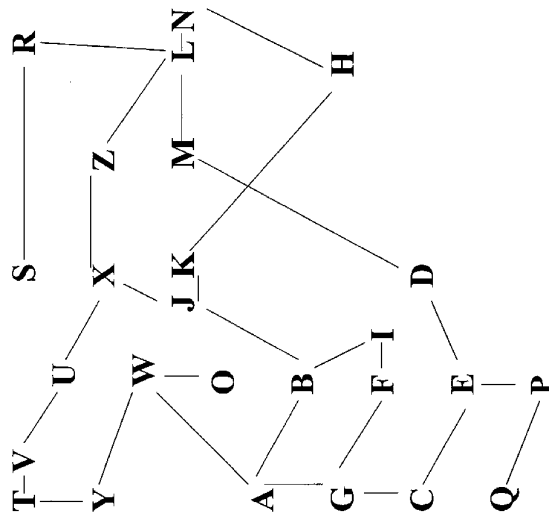
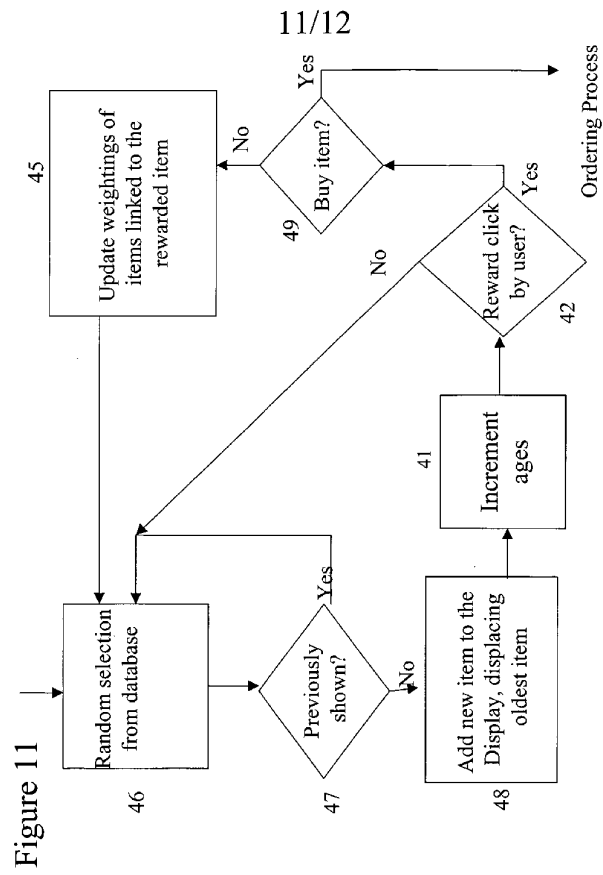


Figure 10

WO 02/080025

PCT/GB02/01107



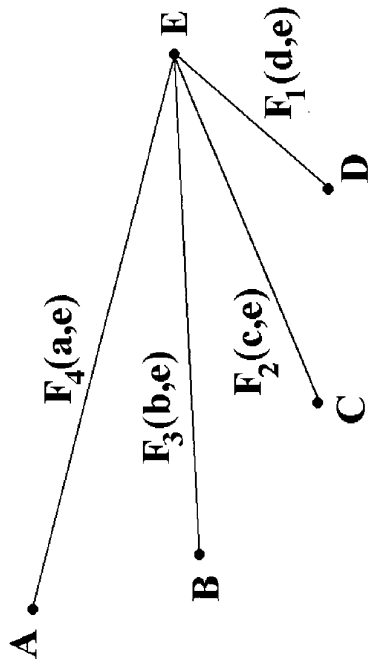
WO 02/080025

PCT/GB02/01107

12/12

Figure 12

Figure 11 11Figure 11



フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,P L,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZM,ZW

(74)代理人 100109830

弁理士 福原 淑弘

(74)代理人 100084618

弁理士 村松 貞男

(74)代理人 100092196

弁理士 橋本 良郎

(72)発明者 タテソン、ジェーン・エリザベス

イギリス国、アイピー 1 3 0 アールディー サフォーク、ウィックハム・マーケット、ハイ・ス
トリート 1 4 5

(72)発明者 タテソン、リチャード・エドワード

イギリス国、アイピー 1 3 0 アールディー サフォーク、ウィックハム・マーケット、ハイ・ス
トリート 1 4 5

F ターム(参考) 5B075 KK38 PQ02 PR06

5B082 GA08